papers and proceedings

the 12th

FEDERAL
FORECASTERS
CONFERENCE

2002

Announcement

The 13th Federal Forecasters Conference (FFC/2003)

will be held

on

September 18, 2003

in

Washington, DC

_____

More information will be available in the coming months.
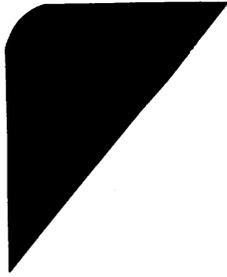
Federal Forecasters Conference - 2002

Papers and Proceedings

Edited by

Debra E. Gerald
National Center for Education Statistics
Institute of Education Sciences
U.S. Department of Education
Washington, DC

Edited and Prepared by

Marybeth Matthews
Kellie Schelach
Healthcare Analysis and Information Group
Veterans Health Administration Office of Policy and Planning
Department of Veterans Affairs
Milwaukee, Wisconsin

Donald Stockford
Health Policy Analysis Group
Veterans Health Administration Office of Policy and Planning
Department of Veterans Affairs
Washington, DC

This page intentionally left blank.

FEDERAL FORECASTERS CONFERENCE ORGANIZING COMMITTEE

Stuart Bernstein
Bureau of Health Professions
U. S. Department of Health and Human
Services

Howard N Fullerton, Jr.
Bureau of Labor Statistics
U.S. Department of Labor

Russell Geiman
Internal Revenue Service
U.S. Department of Treasury

Debra E. Gerald
National Center for Education Statistics
U.S. Department of Education

Karen S. Hamrick
Economic Research Service
U.S. Department of Agriculture

Elliot Levy
International Trade Administration
U.S. Department of Commerce

Stephen A. MacDonald
Economic Research Service
U.S. Department of Agriculture

Jeffrey Osmint
U.S. Geological Survey
U.S. Department of the Interior

Norman C. Saunders
Bureau of Labor Statistics
U.S. Department of Labor

Brian Sloboda
Bureau of Transportation Statistics
U.S. Department of Transportation

Kathleen Sorensen
Office of Assistant Secretary for Policy
and Planning
U.S. Department of Veterans Affairs

Donald Stockford
Veterans Health Administration Office
of Policy and Planning
U.S. Department of Veterans Affairs

Ching-li Wang
U.S. Census Bureau
U.S. Department of Commerce

Peg Young
Bureau of Transportation Statistics
U.S. Department of Transportation

FEDERAL FORECASTERS CONFERENCE
ORGANIZING COMMITTEE

(Front Row) **Ching-li Wang**, U.S. Census Bureau; **Kathleen Sorensen**, Department of Veterans Affairs; **Peg Young**, Bureau of Transportation Statistics; **Debra E. Gerald**, National Center for Education Statistics**; Karen S. Hamrick**, Economic Research Service; and **Elliot Levy**, International Trade Administration.

(Back Row) **Stuart Bernstein**, Bureau of Health Professions; **Norman C. Saunders**, Bureau of Labor Statistics; **Russell Geiman**, Internal Revenue Service; **Jeffrey Osmint**, U.S. Geological Survey**; Howard N Fullerton, Jr.**, Bureau of Labor Statistics; **Stephen MacDonald**, Economic Research Service; and **Brian Sloboda**, Bureau of Transportation Statistics. (Not Pictured**) Donald Stockford**, Department of Veterans Affairs.

FOREWORD

In the tradition of past meetings of federal forecasters, the 12th Federal Forecasters Conference (FFC/2002) held on April 18, 2002, in Washington, DC, provided a forum where forecasters from different federal agencies and other organizations could meet and discuss various aspects of forecasting in the United States. The theme was "Major Shifts: Discontinuity, Uncertainty, and Forecasts."

Two hundred and fifteen forecasters attended the day-long conference. The program included opening remarks by Norman C. Saunders and welcoming remarks from Michael Horrigan, Assistant Commissioner for Occupational Statistics and Employment Projections, Bureau of Labor Statistics.   Following the remarks, a panel presentation was given by Stephen Gallogly, Director of International Energy and Commodities Policy, Department of State; Diane Herz, project manger of American Time Use Survey, Bureau of Labor Statistics; Gregg A. Pane, MD, Chief, Policy and Planning Officer, Veterans Health Administration, Department of Veterans Affairs; Ed Spar, Council of Professional Associations on Federal Statistics (COPAFS); and Herman O. Stekler, The George Washington University.   Stuart Bernstein of the Bureau of Health Professions presented awards from the 2001 and 2002 Federal Forecasters Forecasting Contests. Frederick Joutz of George Washington University and Jeffrey Osmint of the U.S. Geological Survey presented awards for Best Papers from FFC/2000.

In the afternoon, 11 concurrent sessions in two time slots were held featuring 39 papers presented by forecasters from the Federal Government, private sector, and academia. A variety of papers were presented dealing with topics related to agriculture, the economy, labor, population, taxpayers, transportation, and veterans. These papers are included in these proceedings. Another product of the FFC/2002 is the *Federal Forecasters Directory 2002*.

This page intentionally left blank.

ACKNOWLEDGMENTS

This page intentionally left blank.

2002
FEDERAL FORECASTERS CONFERENCE FORECASTING CONTEST

WINNER

Ken Beckman
U.S. Geological Survey

First Runner Up

Betty W. Su
Bureau of Labor Statistics

Second Runner Up

Terry Schau
Bureau of Labor Statistics

Third Runner Up

Douglas Meade
University of Maryland

2001
FEDERAL FORECASTERS CONFERENCE FORECASTING CONTEST

WINNER

John Golmant
Administrative Office of the United States Courts

First Runner Up

Thomas D. Snyder
National Center for Education Statistics

Second Runner Up

Mirko Novakovic
Bureau of Labor Statistics

Third Runners Up
(Tie)

Peggy Podolak
U.S. Department of Energy

Terry Schau
Bureau of Labor Statistics

2000
BEST CONFERENCE PAPER

WINNER

"Accuracy of the U.S. Census Bureau National Population Projections and Their Respective Components of Change"

Tammany J. Mulder
U.S. Census Bureau

First Runner Up

"Assessing the Impact of Government Legislation on BSE in the U.K."

Sandy D. Balkin
Ernst & Young LLP

Second Runner Up

"The Accuracy of Recent Short-Term Employment Forecasts Obtained by Employer Surveys: The State of Illinois Experience"

Roy L. Pearson
College of William and Mary

George W. Putnam
Waleed K. Almousa
Illinois Department of Employment Security

This page intentionally left blank.

The 12th
Federal Forecasters Conference
FFC/2002

Scenes from the Conference

Photos by Norman C. Saunders, Bureau of Labor Statistics

Norman Saunders opens the 12th Federal Forecasters Conference and introduces Michael Horrigan, the Assistant Commissioner for Occupational Statistics and Employment Projections.

Michael Horrigan of the Bureau of Labor Statistics extends a warm welcome to the conference participants and presents past conference highlights.

Stuart Bernstein introduces the 2001 and 2002 winners of the Federal Forecasters Forecasting Contests.

2002 AND 2001
FEDERAL FORECASTERS FORECASTING CONTEST WINNERS

(Front Row) **Betty Su**, Bureau of Labor Statistics; **Peggy Podolak**, Department of Energy; and **Thomas Snyder**, National Center for Education Statistics. (Back Row) **Terry Schau**, Bureau of Labor Statistics; **Mirko Novakovic**, Bureau of Labor Statistics, and **Ken Beckman**, U.S. Geological Survey. (Not Pictured) **John Golmant**, Administrative Office of the U.S. Courts and **Douglas Meade**, University of Maryland.

Fred Joutz of George Washington University announces the winner and runners up for the Best Conference Paper of the Federal Forecasters Conference 2000.

Kathleen Sorensen sets the mood of the conference theme of "Major Shifts: Discontinuity, Uncertainty, and Forecasts" and introduces the morning panelists.

Stephen Gallogly delivers a point on international energy and commodities policy at the State Department.

Diane Herz presents the design and methodology of the American Time Use Survey.

Dr. Gregg A. Pane discusses policy and planning at the Veterans Health Administration.

Ed Spar discusses the work of the Council of Professional Associations on Federal Statistics.

Herman Stekler discusses examples related to discontinuity, uncertainty, and forecasts.

The registration team pauses to smile for the camera. (Left to Right) Harley Carpenter, Department of Veterans Affairs, Linda Diane Felton, Economic Research Service, and Patricia Cleveland, Economic Research Service.

This page intentionally left blank.

CONTENTS

MORNING SESSION

FFC/2002 PAPERS

CONCURRENT SESSIONS I

ISSUES FROM THE BLS 2010 PROJECTIONS

MACROECONOMIC ISSUES

INDUSTRY ISSUES

CONCURRENT SESSIONS II

EVALUATING PROJECTIONS AND DEALING WITH NAICS

COPING WITH ISSUES OF CONTINUITY IN NEW RACIAL CLASSIFICATION

FORECASTING IN TRANSPORTATION

PROJECTING TAXPAYER BEHAVIOR

COMMODITY FORECASTING MODELS

This page intentionally left blank.

This page intentionally left blank.

Panel Presentation

Major Shifts: Discontinuity, Uncertainty, and Forecasts

The panel addresses the issues of discontinuities in past trends, uncertainties about future trends and events, and the challenges of producing forecasts under these conditions. Most forecasts rely on historical trends or past experiences to project or predict the future. When the past cannot be used to predict the future due to major unexpected shifts in economic, political, or social conditions, what can we do as forecasters? The future can be more uncertain due to the occurrence of major events. The measurement of past trends may also be interrupted by changes in policy or data availability. These are challenges that must be faced by agencies responsible for producing forecasts.

Stephen Gallogly
Director, International Energy and Commodities Policy
Department of State

Diane Herz
Project Manager
American Time Use Survey
Bureau of Labor Statistics, U.S. Department of Labor

Gregg A. Pane, MD
Chief Officer, Office of Policy and Planning
Veterans Health Administration, Department of Veterans Affairs

Ed Spar
Council of Professional Associations on Federal Statistics (COPAFS)

Herman O. Stekler
Professor of Economics, The George Washington University

This page intentionally left blank.

The American Time Use Survey
Diane Hertz, Bureau of Labor Statistics

What is the American Time Use Survey?

•A new survey funded by BLS to be conducted by the Census Bureau beginning in January 2002

•A survey designed to measure how people spend their time, including where they are and who they are with

Time-use data will help inform policy debates and business decisions.

•Urban planning
•Transportation planning
•Emergency management
•Child care policy
•Workforce/employer/retirement policy
•Valuation of time (market and non-market) in legal cases
•Determining product audience
•Health care analysis
•Disabled adult care
•Eldercare
•Education policy
•Adolescent time-use patterns
•Services for older adults

Why collect Time-Use Data?

•Time-use data will assist researchers analyzing changes in quality of life

•Time-use data provide broader measures than income and earnings

 –How much time are people spending with their children?
 –How much leisure time do Americans have?
 –How much sleep are we getting?
 –What are the trends?

•The data will help us understand who does non-market work in the U.S.
 This includes things we do for our families, communities, and country for which we
 are not paid.

 –Who does the non-market work in the U.S.?
 –How much time is spent on this work?
 –How does this compare with other countries?
 –How would GDP change if non-market work were included?
 –How can non-market work-time be valued in legal cases?

•Data will help us better understand changes in work patterns.
 This means knowing when and where things are done helps us understand the impact of technology and
 identify overall changes.

 –How many hours are people working?
 –How are they mixing days between work and nonwork?
 –Where are they doing work?
 –How are they combining work and child care?
 –What policies may help workers and their families?

•We will know more about how the U.S. compares to other countries. With time-use data, we can compare market and non-market sectors.

> –At least 25 other countries have done time-use surveys
> –About 25 more are about to conduct them
> –Countries in North America, South America, Africa, Europe,
>   and Australia and New Zealand collect this data
> –Some have used telephone data collection

Estimation objectives

•Quarterly and annual estimates of time spent in variety of activities
•Measures for average weekday, weekend, and week
•Data by demographics and labor force status
•Continuously collected data
•First annual estimates: mid-2004

The American Time Use Survey Sample

•**Household Sample**
> –15 and over
> –Drawn from CPS (after completing last interview)

•**Stratified by various household characteristics**–Race of CPS reference person
> –Ethnicity of CPS reference person
> –Number of adults
> –Presence of children
> –Age and education of CPS reference person

•**Designated person (DP)**
> –<u>One</u> individual <u>randomly selected</u> from sampled households for American Time Use Survey

The American Time Use Survey Operations

•**Advance materials sent to DP**
•**Incentive for no-telephone-number households**
•**Designated Day (DD)**–DP is interviewed only <u>once</u>
> –DP assigned 1 eligible day of the week (Monday, Tuesday, etc.) on which they can be interviewed

•**Computer Assisted Telephone Interviewing (CATI)**–Telephone interview in English or Spanish
> –By Census Bureau interviewers in Jeffersonville, IN
> –Interview focuses on 24-hour time diary about yesterday's activities (4am-4am)
> –8-week CATI "fielding" period (no field visits)

The American Time Use Survey: Structure of the survey

•Introduction
•Household Roster & Employment Status, including non-HH own children
•Core Time Diary
•Summary Questions, including paid work, child care, and missed days
•Labor Force Updates
•Future: Modules on specific subjects to provide complementary information

TURNING INFORMATION INTO INSIGHT:
PUBLIC-PRIVATE SECTOR PARTNERSHIP FORECASTING MODEL

Gregg A. Pane, M.D., M.P.A.
Department of Veterans Affairs

Introduction

At Federal departments and agencies, many forecasting activities are major tools to executive level decision makers who are charged with developing and implementing data-driven policy and budget scenarios. This is no less the case at the Department of Veterans Affairs and, in particular, the Veterans Health Administration (VHA) within VA, which administers all of VA's health care programs and services and delivers health care to eligible veterans. The need for flexible, accurate, and clear modeling of veterans' potential demand for VA health care, their utilization of health care services, and the associated health care costs are cornerstone to strategic management in VHA. In recent years, VHA has developed a collaborative partnership with a major private sector actuarial firm to create an integrated forecasting model that will enhance VHA's strategic management processes now and well into the future. I will discuss various aspects of this model in greater detail, but first let's take a closer look at VHA.

Veterans Health Administration (VHA)

VHA is a major contributor to the Nation's health care system. VHA has a Fiscal Year 2003 budget of over $25B. There are now over 6 million veterans (about 25% of the total veteran population) enrolled in the VHA Health Care System, and VHA, annually, provides health care services to some 4 million (of the 6 million) enrolled veterans. In FY 2001, VHA provided over 700,000 inpatient episodes of care and over 43 million outpatient visits to veterans.

Among Veterans Health Administration assets are included: 21 Veterans Integrated Service Networks (VISNs; the 21 figure reflects the recent administrative consolidation of 2 VISNs), 163 hospitals, 601 community based outpatient clinics, 134 nursing homes, 206 readjustment counseling centers, and 43 domiciliaries.

VHA has a large system of academic affiliations between its VA medical centers and many of the medical schools in the U.S. Indeed, VHA serves as the largest single provider of health professions training in the world. It is a little known fact but true that about half of all physicians in the United States have trained with VA.

VHA also administers one of the largest and most productive research organizations in the country; VHA physicians have been both nominees and winners (e.g., Rosalyn Yallow) of the Nobel Prize for medical research.

VHA is also becoming the principal Federal asset for medical assistance in large-scale disasters. This was one of VA's lesser-known roles before September 11 but, since September 11, VA's responsibilities in this regard have been expanding.

Furthermore, VHA is the largest direct care provider in the world, employing over 226,000 employees and health care providers. It also provides more services to homeless persons than anyone in the country.

VHA Enrollment System

The Veterans Health Administration (VHA) Enrollment System was mandated by Congress in 1996 (Veterans Health Care Eligibility Reform Act of 1996, P.L. 101-262) to help VA stay within its budget, since VA care is not an entitlement like Medicare. As a result of the Act, (most) veterans must be enrolled in order to obtain VA health care. They are assigned to one of seven distinct enrollment priority groups and subsequently enrolled. They have access to a comprehensive range of benefits and services (VHA's "Medical Benefits Package"). Some of the veterans who do not have to enroll include veterans who: (i) have a service-connected compensation rating of 50% or greater, (ii) have been discharged in the past year for a compensable disability that VA has not yet rated, or (iii) want care for a service-connected disability.

Since implementation of VHA enrollment in 1998, participation by veterans has been high and continues to grow. However, annually, VA assesses whether it will have the resources to meet the demand for care by veterans in all priorities. If, based on the Secretary's annual enrollment decision, it cannot, then VA may not continue to enroll veterans in the lowest level of

priorities. However, for the last four years, VA has been able to open the VA health care system to all veterans, even higher income veterans, if they are willing to make co-payments. Other potential management efficiencies that might be achieved are also considered in the Secretary's annual enrollment decision.

As of September 30, 2001, there were some 24,911,226 living veterans in the U.S. and P.R. and as of September 30, 2001, some 5,848,067 veterans (about 23% of all veterans living in the U.S. and P.R.) were enrolled in the VHA Health Care System. As of September 30, 2001, Priority 7 veterans, who include "higher income" non-service-connected veterans, accounted for about 29% of all VHA enrollees. Since the inception of VHA Enrollment, the number of Priority 7 veterans has shown the largest increase, both in absolute numbers and percent, although the number of Priority 5 veterans (about 40% of all VHA enrollees) who are predominantly "low income" has also been increasing. Priority 7 veterans are, however, the lowest cost enrollees since they have other eligibilities and insurance and rely to a lesser degree on VA than enrollees in other priorities. They may be coming to VA to bridge gaps in their insurance coverage or to reduce their out-of-pocket costs. Based on the enrollment projections, developed for the Secretary's annual enrollment decision, enrollee demand shows no sign of decreasing, with a 31% increase in the number of enrollees from 6.1 million in 2002 to 8.0 million in 2010. Most of the increase is due to increases in Priority Category 5 and 7 enrollees. The current VHA enrollment projections show that VHA enrollment will continue to increase and expenditures will also continue to rise over the next decade, if no constraints are implemented and if resources (supply) can meet the projected demand.

Eligibility Reform: VHA *Before and After* the "Veterans Health Care Reform Act of 1996", P.L. 101-262

Prior to the "Veterans Health Care Reform Act of 1996", eligibility rules for VA inpatient and VA outpatient care were different and very complicated, favored care in inpatient settings, and decisions about veterans' access to care were often made locally based upon local resources. After eligibility reform, VA health care came to be provided in the most cost-effective and clinically appropriate manner. Preventive and primary care services were offered. However, enrollment was required for receipt of VA health

care and, once enrolled, all enrolled veterans received the VA's Medical Benefits Package. As a consequence of eligibility reform, a real national system of care evolved. However, with the tremendous growth in enrollment over the past few years, changes in enrollment policies are being considered to better manage the demand.

Veteran Population Trends

The current veteran population of some 25 million veterans is aging rapidly. The current official VA veteran population projections show that the total veteran population count will decrease rapidly between 2000 and 2020. However, in contrast, the number of female veterans in the veteran population will increase rapidly between 2000 and 2010 and beyond. The number of veterans age "65 or over" will reverse their downward trend and increase to peak again in 2013 (due to the aging of the Vietnam era cohort of veterans), as it did around the year 2000 (due to the aging of the World War II cohort of veterans). Also, the actual number of veterans age "85 or over" will increase greatly between 2000 and 2010 and beyond. These veteran population trends are important to VA, since they factor into all of VA's most important modeling processes. Just for example, the increase in the "85 or over" veteran population is significant since this is VHA's population at risk of increasing acute, rehabilitation, and long-term care programs and services.

VHA Demand Model Overview

This is VHA's fifth year of using a VHA Health Care Services Demand Model, developed and refined with the knowledge and capabilities of both VA and the private sector health care actuarial firm, Milliman USA. Few other public agencies have such a robust, flexible projection model, that predicted enrollment to within +/-4% for 2001.

The model integrates, among other things, data on veteran population, historical monthly VHA enrollment, enrollee characteristics from VHA surveys of enrollees, VA actual unit costs, and both VA and private sector workload measures. A summary of the modeling process follows. The results are enrollment, workload, unit costs, and expenditure projections.

Enrollment Projections

1. Obtain baseline actual enrollment by scrambled SSN
2. Develop enrollment rates using historical enrollment and historical veteran population
3. Develop projections of new enrollees using the rates developed in Step 2, the baseline from Step 1, and veteran population projections
4. Apply mortality rates to enrollment projections

Workload Projections

1. Summarize private sector health care utilization averages by geographic area
2. Adjust utilization to reflect Medical Benefits Package and Millennium Bill health care services
3. Adjust utilization to reflect age and gender characteristics of the projected veteran enrollee populations
4. Adjust utilization to reflect the morbidity of the projected veteran enrollee populations relative to the underlying private sector populations (VA patient diagnosis data used to assess relative morbidity levels)
5. Adjust utilization to reflect the estimated degree of health care management observed within the VA health care system relative to the loosely managed level observed in the local community (VA inpatient diagnosis and workload data used to assess Degree of Management)
6. Adjust utilization to reflect the estimated veteran enrollee reliance on VHA for their health care needs (veteran enrollee survey data and HCFA match data used to assess reliance)
7. Adjust utilization to reflect the residual differences between modeled and actual historical VA workload (estimates of unmeasured morbidity, reliance, and degree of health care management differences)

Unit Cost Projections

1. Obtain baseline CDR-based VA unit cost data
2. Unit Cost data adjusted for health care service mix inherent in data
3. Adjust Unit Costs to reflect reconciliation to historical VA total health care obligations

Expenditure Projections

1. Enrollment, Workload, and Unit Cost Projections are combined to produce Expenditure Projections

Results from the VHA Health Care Services Demand Model have been incorporated into other Departmental planning processes, integrated with budget and performance measures, and leverage the ability to perform diverse policy scenarios and forecasts. The model is continually updated with improved methods, new data sources, and additional analyses each year.

Strategic Management Framework

As mentioned in the introduction, the need for flexible, accurate, and clear modeling of veterans' potential demand for VA health care, their utilization of health care services, and the associated health care costs are cornerstone to strategic management in VHA. The strategic management framework in VHA has evolved over the past several years, as VA made many organizational and service delivery changes both before but also in conjunction with the previously described eligibility reforms. There are major continuing efforts to improve access along with a fundamental change in focus upon outpatient, including preventive and primary, care. VHA's continued re-organization reflects the rapid expansion and integration of VA health care programs and services, but also, a more population-focused, community-based, and prevention-oriented system, that ensures timely, accessible, and quality care. There are six guiding principles in all of this:

1. Put Quality First Until First in Quality
2. Provide Easy Access to Medical Knowledge, Expertise, and Care
3. Enhance, Preserve, and Restore Patient Function
4. Exceed Patients' Expectations
5. Maximize Resource Use to Benefit Veterans
6. Build Healthy Communities

These are VHA's "6 for 2006" guiding principles and goals.

Enhancing Strategic Management: Awards and Recognition

VA has received various awards and recognition for enhancing strategic management processes in VHA. In 1999, the Government Performance Project, conducted by Syracuse University and Government Executive magazine, awarded VA the second highest grade of any Federal Department or agency for its FY 1999 Performance Plan. Subsequently, the Mercatus Center at George Mason University rated VA an "A" on its FY 2000 Annual Performance Plan, one of only two Federal agencies to receive this grade.

Enhancing Strategic Performance: "6 for 2006" Linkage with VISN/VAHQ Requirements

There are five basic components to linking the "6 for 2006" guiding principles and goals to VISN and VA Headquarters strategic planning requirements. These are:

1. Network (VISN) Performance plans
2. Network (VISN) Strategic plans
3. VHA Chief Officer Contract
4. Budget and Performance Plan
5. VA Strategic Plan

The linkage of "6 for 2006" principles and goals with VISN and VA Headquarters requirements links strategic goals to operational tactics and provides an accountability framework for driving VHA performance.

Federal Budgets and the Future

Basic things the Federal Budget tells us include: how much will we pay down the national debt; how much will go to Defense spending; how much will our taxes go up or down? For VA, we want to know how much health care spending will there be; and will the budget and resources be enough to provide care to our clientele? The Federal Budget Process is evolving, and so, too, are VA's and the other Departments' and agencies' budget processes.

This is largely the consequence of some of the Statutory Reforms of the1990's. Of greatest significance, the CFO Act of 1990 mandated that Federal Departments and agencies annually prepare audited financial statements, and the Government Performance and Results Act of 1993 (GPRA) mandated annual performance plans. FY 1999 was the first year for agencies to provide both performance reports under GPRA and audited financial statements under the CFO Act. The intent was to make Federal entities results-oriented and accountable. Results and improved management would lead to better decision making on the part of Federal managers, and better Congressional decision making, too, with requested resources linked to results via performance information.

The VA Perspective on the Federal Budget Process, particularly in terms of health care, includes VA Performance-Based Strategic Planning, Budgeting, and Decision Making, under GPRA, the VHA Performance Management System, and VHA Performance-Based Budgeting concepts.

Should a performance budget be included as part of the President's Budget? The President's Management Agenda is part of the FY03 Federal budget; it is a strategy for improving management and performance of the Federal government. All current Federal Budgets and Forecasts are moving headlong into the Federal Performance Budgets of the future. There will be Performance-Based Budgets and there will need to be planning models and forecasts to support the performance-based budgets. For VA, its innovative and powerful public-private sector partnership health services demand model will be a major tool to aid in decision making for results. Flexible, accurate, practical, clear modeling and forecasting is cornerstone to VA's future.

Slide 1

TURNING INFORMATION INTO INSIGHT

Public-Private Sector Partnership Forecasting Model

Gregg A. Pane, M.D., M.P.A.
Chief Policy and Planning Officer
Veterans Health Administration
Federal Forecasters Conference
April 18, 2002

Slide 2

Enhancing Strategic Management Processes in VHA

VHA a major contributor to the Nation's healthcare system

Serves as the largest single provider of health professions training in the world.

Becoming the principal Federal asset for medical assistance in large-scale disasters.

Provides medical care to 4+ million veterans.

One of the largest and most productive research organizations in the country.

Largest direct care provider for homeless persons in the country.

Slide 3

Veterans Health Administration
22 Veterans Integrated Service Networks
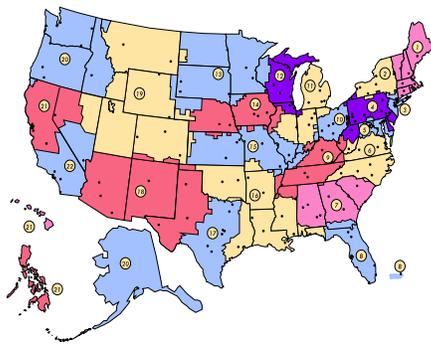
Slide 4

Veterans Health Administration Assets

- 22 Veterans Integrated Service Networks (VISNs)
- 163 Hospitals
- 601 Community Based Outpatient Clinics
- 135 Nursing Homes
- 206 Readjustment Counseling Centers
- 43 Domiciliaries

Slide 5

Eligibility Reform
Before Eligibility Reform

- Different and complicated eligibility rules for inpatient and outpatient care
- Eligibility rules favored inpatient setting
- Access decisions made locally depending upon resources

Slide 6

Eligibility Reform
After Eligibility Reform

- Health care provided in most cost-effective clinically appropriate manner
- Preventive and primary care services
- Enrollment required, with access determined by priority group through annual resource-based enrollment decision
- Costly, inappropriate care minimized

Slide 7

JOURNEY OF CHANGE:
DISCOVERING SIX FOR 2006

OVERVIEW

VHA

STRATEGIC MANAGEMENT FRAMEWORK

6 for 2006

Derived from Domains of Value:

1. Put Quality First Until First in Quality
2. Provide Easy Access to Medical Knowledge, Expertise, and Care
3. Enhance, Preserve, and Restore Patient Function
4. Exceed Patients' Expectations
5. Maximize Resource Use to Benefit Veterans
6. Build Healthy Communities

Slide 8

Veteran Population: Age Trends, 2000 - 2020

o The total veteran population will decrease between 2000 and 2020;

o however, the number of veterans age 65 or over will peak again in 2013;

o and veterans age 85 or over will more than double between 2000 and 2010.

Number in 000s)

Under 45

45 - 64

65 - 84

85 and over

Year

Source: Vet Pop 2000

Slide 9

Veteran Population
by Period of Service
(September 30, 2000)

Period of Service

Gulf War Era   2.9
Vietnam Era   8.3
Korean Conflict   3.4
World War II   5.5
World War I
Peacetime   6.2

Millions

Source: Vet Pop 2000

Slide 10

Veteran Population Age 85 or Over

Number in 000s

Year

Source: Vet Pop 2000

Slide 11

Current Enrollees, as of September 30, 2001

Priority 1 9%
Priority 2 7%
Priority 3 13%
Priority 4 3%
Priority 5 40%
Priority 6 2%
Priority 7 29%

Source: Veterans Health Administration Data
Note: Total Current Enrollees = 5,848,067; excludes "ineligible", "declined", and "deceased".

Slide 12

National Projected Average Enrollment
(Enrollees) by Priority
Fiscal Years 2002 - 2010

Number

FY2002 FY2003 FY2004 FY2005 FY2006 FY2007 FY2008 FY2009 FY2010

7c
7a
6
5
4
3
2
1b
1a

Slide 13

National Projected Expenditures by Priority
FY 2000 - FY2010

Slide 14

Pharmacy Obligations

* Prior to 1998, other than the cost to collect, collections were returned to the Treasury

Slide 15

General Model Overview

- Enrollment Projections
- Workload Projections
- Unit Cost Projections
- Expenditure Projections

Slide 16

FY 2003 Enrollment Decision: VHA Projection Model

- Fifth Year for the Projection Model
- Integration with other Planning Models
- Increasing Integration with Budget and Performance Measures
- Policy and Budget Scenarios and Forecasts

Slide 17

FY 2003 Enrollment Decision: Projection Model

VHA Actuarial Model

- Enrollee Focus - uses private sector utilization benchmarks & adjusted VA unit costs for actual and projected VA enrolled population
- Focuses on actual & historical enrollment trends, with attention to total veteran population and pools of eligible veterans by priority & socio-demographics

Slide 18

FY 2003 Enrollment Decision: Projection Model

Actuary Model Benchmarks

- Adjusted for age, gender, morbidity, mortality, VA reliance
- Adjusted for degree of management in VA vs. community standard
- Incorporates experience gained from actual to expected analyses
- Includes improved methods, new data resources & additional analyses each year

Slide 19

Slide 20

Slide 21

The Future

- Turning Information into Insight

- Data-driven policy and budget scenarios

- On-time, objective, executive-level, decision focus

- The President' s Management Agenda is part of the
  FY 03 Federal Budget: it is a strategy for improving the
  management and performance of the Federal government

- Flexible, accurate, practical, clear modeling is cornerstone

Projecting Data Dissemination
Edward J. Spar, Executive Director COPAFS

Just to be clear, I'm an unrepentant Luddite, who's convinced that civilization ended somewhere around 1905.

My own background in forecasting was really in making some projections many years ago when I had my own demographic research company. I would never grace what we did with the word "forecasting." Indeed, I wonder how many people who say they forecast really just push the data forward with little modeling. Gerber Foods couldn't get enough of five and ten year projections of 0, 1, and 2 year ages, Philip Morris who's market research told them that black teenage girls 15 to 19, were prime candidates for menthol cigarettes, and the Department of Defense loved our five year projections of 17 to 29 year old men and women. All this was for each of the 3132 counties in the United States. Egads, we sure knew how to help do in youngsters either with ghastly food, drugs or bullets. Ah well, no wonder I got out of the business. Well, I still want to project into the future. However, the issue that I've been asked to comment on today is on what data dissemination might look like in the future. And being a cautious projector, I don't think I want to go beyond the near future.

I know that you've heard, in previous meetings, about virtual data and one stop shopping. It all sound great, but I'm not so sure we will get beyond what I call one stop routing for quite a while.

We are still at the stage where someone looking for data on income, for example, is in for a tough ride. Which one do they mean? BEA disposable income? CPS household income? Decennial Census income? SIPP income in kind? SAIPE Poverty income? And there are many more. I remember a great article written by Courtenay Slater many years ago in American Demographics magazine, she's a former Chief Economist at the Department of Commerce, where she pointed out that are quite a few definitions of income in this country. And 25 years later we do a poor job of explaining the differences not only to users, but to ourselves. My point is that we are not even close to the point where a data users, either via a phone call or on the Internet can readily obtain linked information or the documentation (which we now like to call metadata) about the data. Until this intellectual process is thought through and codified, I don't see the Internet and systems such as Fedstats as anything more than routing services. On the other hand I can see potentially a great loss if agencies decide to use the Internet as an archiving medium. It's like having put all your data on 5 1/4 inch floppies ten years ago. Who knows what we will be using for data retrieval in ten years. Most likely something very different from the Internet.

Indeed, in the short term, there's the issue of insuring that data users aren't left behind. The Census Bureau will publish about 15 percent of what they published in 1990 for the 2000 decennial census. And that was only after a lot of in-fighting with users. Originally they were only planning to publish 5 percent. After all, CD's and the Internet are cheaper. But who are we leaving behind? And since I'm trashing the Internet, what about software? There's no standardization or coordination within or among agencies not only on the Internet, but also on CD's. I predict this will get worse.

Moving on to data sharing, although there will be some movement this year in the direction of legislation, we're not even at the point where federal statistical agencies can share their information. And let's abandon the idea that we will ever see an integrated federal statistical agency. So what we have is fairly loose confederacy of agencies working to produce excellent data. Our best hope for data sharing is the possibility that what was known as the Horn bill that actually passed the House, but died in the Senate, could be tacked onto another bill. This would allow about ten agencies, under strict confidentiality guidelines, to share files for research purposes. It might also help solve the previously mentioned issue of software, although I doubt it. There must be something macho about having your own software for your product, even if it's very similar to someone else's data set.

Data sharing would also prove to be full employment bill for those on disclosure review boards. Which is my next concern. I think we may be heading towards a time when DRB's release memos at the end of the year state: a) great news, no individuals disclosed and b) great news, no data products were released this year. I think the argument that faster computers will enable and cause users to break the confidentiality rules through linking more easily is bunk. We've always known how to do it. I did it with retail sales store group data using an IBM 4300 25 years ago. There's really nothing new here.

Just a word about the September 11 and data. If you're not aware, Section 508 of the Patriots Act allows the Justice Department to gain access to National Center for Education Statistics data. This is scary and there's no sunset rules that apply. If confidentiality can be breached at NCES it can happen to anyone At the same time, we're seeing transportation, environmental and other data disappear. I predict that we will see a lot of regulatory and monitoring information disappear in the near future under the rubric of national security. What a convenient way to keep regulatory data from the eyes of environmental groups.

There's also a public/private sector cooperative (and some times not so cooperative) effort to insure that users get what they need. And perhaps with tighter budgets we will even see more of this love/hate relationship. It will range from the private sector vendor either adding value to your data through updating to simply ripping the cover off your books or downloading from the Internet, putting their own logo on, and selling the data for wonderfully outrageous sums of money. Well, when you have no copyright, and you no longer print reports, that's what happens.

To conclude, what we need is a more basic look at what users need and what the agencies are capable of providing. Sure, move ahead with high tech plans. Someday it might even happen. However, what's needed today, in what is most likely your production oriented environment, is more interaction with users to find out what they are doing, what they need and how best to meet these needs.

Improving Our Ability To Predict The Unusual Event

H.O. Stekler, George Washington University

This paper deals with the problems of predicting unusual events. Certainly the events of September 11 were unusual and they definitely were not predicted in advance. In the aftermath of 911 questions have been raised about our government's ability to gather and interpret intelligence information. Certainly there was an intelligence failure, but what has not been recognized is that these "failures" are similar to the errors that have been observed in all fields of forecasting when "big" or "unusual" events occur. In fact, there is some evidence that this inability to forecast these "big" (unusual) events is more pervasive than is customarily recognized.

I.   Examples of Forecast Failures

There are many examples. The inability to predict cyclical turning points has been recognized as one of the biggest failures of macroeconomic forecasting. Meteorologists are frequently surprised by the intensity of some storms. Political pollsters made egregious blunders in predicting Landon and Dewey victories in 1936 and 1948, respectively, and the intelligence community failed to predict the collapse of the Soviet Union. Moreover, there have been a number of military actions where one of the combatants was completely surprised, e.g. Pearl Harbor, the Battle of the Bulge in 1944, the Yom Kippur War in 1973, and the invasion of Kuwait in 1990.

Technological change is definitely associated with structural change. The failure to foresee the impact of new technological developments is legendary. Examples include the predictions made in the 1940s that the world's entire demand for computers would be less than ten and the belief that TV would not be successful because people would prefer to go to the movies rather than sit in front of a box. There are even more recent examples, i.e. the failure to foresee the impact that semiconductors would have on consumer electronics and computers. The use of VCRs was also vastly underestimated.[1]

But what concerns people at this conference most is the inability to predict what happens when there are structural changes resulting from new legislation or deregulation. Here too there are examples of forecast failures. In the airline industry deregulation, for example, few economists predicted that the industry would have the structure that exists today.

II.   Why Can Such Failures Occur?

A.   Model Failure

It is one thing to list the failures and another to provide an explanation for why they occurred. Let us first examine the role of formal models. If the policy or structural changes are small, models that have done a good job in explaining the past are likely to be able to do a good job in forecasting the future. The structural relationships that have been observed in the past are likely to continue in the future. However, when the policy models involve fundamental structural changes, the empirical models estimated from past relationships may not be reliable in predicting the future. In that case we will have to look for different methods for forecasting the future. (I will discuss these in Section III).

B.   Non-model Failures

I classify the non-model failures into three categories: (1) those associated with the interpretation of data or facts, (2) those associated with forecaster bias and (3) forecaster preferences.

1.   The Interpretation Process

---

[1]   On the other hand, the technological forecasting literature is also replete with forecasts of big breakthroughs that never or still have not occurred, i.e. virtually all of our electricity would be generated by nuclear power.

In discussing the interpretation process, I make an assumption about the way that forecasters prepare and then revise their predictions, i.e. a Bayesian approach. This means that the forecasters begin with subjective probabilities about the likelihood of an event and as new information becomes available, these prior probabilities are revised in a Bayesian manner.

I once sought to determine why economists had failed to predict some cyclical downturns. If these economists had considered that a recession was even remotely possible, and if they had examined new information as it became available, they definitely should have forecast an imminent downturn. Because they did not issue this prediction, I concluded that, originally, they had assumed that there was **no likelihood of a recession (zero priors)** and were surprised that it happened.[2] Why did the economists have these zero priors? It has been said that cyclical peaks were often associated with credit crunches and that prior to a crunch, there is no reason to look for the turning point. As a generalization this suggests that forecasters reason by analogy and look for patterns that repeat themselves. Rather they should be attempting to understand the dynamics of the process which might lead to the event that ought to be predicted.

In a similar vein, later analyses of the intelligence data that were available prior to Pearl Harbor indicated that the information about likelihood of such an attack could have been inferred **if** the analyst had considered this a possibility. However, given the mass of data being examined, those clues would not have stood out starkly by themselves. It required an anlyst to search for them among all the available data.

2. Bias

However, an unremitting search for facts to substantiate a particular point of view might, on one hand, lead to false alarms, or be another explanation for the failure to predict unusual events. Bias often leads one to select evidence which is consistent with one's prior views, whether or not they are an accurate reflection of reality. Suppose there is overwhelming evidence that suggests that an unusual event is likely to occur, and that there is a single piece of information that contradicts this evidence. If the analyst focuses on this datum, the event will not be predicted.

3. Forecaster Preferences

Finally, there is the possibility that practicing forecasters have asymmetric loss functions. They do not place the same values on the two types of possible errors: (1) failing to predict the event and (2) predicting an event that does not occur. Thus forecasters' subjective costs may determine whether or not the event is predicted. An individual more concerned about the costs of failing to predict an event, would need less concrete information about the likelihood of the event than would the person who was more concerned about false alarms.

III.   Some Simple Solutions

Some of these failures can be eliminated by developing models and techniques especially designed to predict the unusual event. In economics, techniques based on leading series have been found that lead cyclical turning points. Studies have shown that, if we are willing to accept false positive errors, that these methods can predict or at least detect major turning points. In the area of political forecasting, the emphasis on new procedures has probably eliminated the possibility of egregious errors. The failure to predict Truman's 1948 victory can be explained by the pollsters terminating their interviews too early. Now the electorate is sampled up to the election.

It is also possible to reduce errors that primarily stem from forecasters' behavior. It is important that analysts and their principals recognize observed biases. Forecasts should be evaluated, and feedback should be provided to the analyst whenever bias is observed. This is done in the weather forecasting community with the

---

2   Could it be that our intelligence experts assigned a zero likelihood to the possibility that a plane could be
   used as a weapon for destroying buildings? Any reader familiar with the Clancy novels would have known
   otherwise.

result that the forecasts have become less biased. If a forecaster seems to have an asymmetric loss function, both the practitioner and the user of the forecasts should be made aware of this fact.

How can one avoid such errors? First, all data that might provide insights about the likelihood of a "big" event must be available to the analysts who are preparing the forecast. Second, and even more important, the possible biases of the forecasters must be recognized. Such biases often lead to the selection of evidence that is consistent with one's prior views. Suppose that there is a large amount of evidence that an unusual event is likely to occur, but there is a single piece of information, like no plane has been hijacked in the US for ten years, that contradicts this evidence. If everyone focuses on this single datum, the event will not be predicted. Finally, the preparation of a strategic forecast involving competitors or enemies requires an additional step. It involves predicting the possible actions and reactions of "the other side". There is a forecasting technique that may be used for this purpose. Military planners have long used war games to take into account the possible actions of potential enemies.

In evaluating the performance of the intelligence community, it is important to recognize that failing to predict the "big" event afflicts every field of forecasting. However, individuals who work in the field of forecasting have suggested ways of improving predictive performance and indicated where it is necessary to obtain a further understanding of the forecasting process. The country must make sure that these ideas have been adopted by the intelligence community.

This page intentionally left blank.

CONCURRENT SESSIONS I

This page intentionally left blank.

Issues From the BLS 2010 Projections

Chair: Michael Pilot, Bureau of Labor Statistics, U.S. Department of Labor

The Labor Force Over the Next 50 Years

Mitra Toossi, Bureau of Labor Statistics, U.S. Department of Labor

This paper reviews the past and projects the future of the U.S. labor force over a period of 100 years from 1950. Critical factors that shaped the size, composition, and characteristics of the labor force over the last fifty years have been used to project the composition of the work force in the coming decades. The U.S. labor force is projected to grow at a much slower rate, in part due to stabilization in the growth of women's participation rates. The labor force will become older as the baby boomers exit the work force; and will show more diversity due to higher growth of population and participation of minorities.

The BLS Defense-Related Employment and Projections

Jeffrey Gruenert, Bureau of Labor Statistics, U.S. Department of Labor

This study provides estimates for both direct and indirect industry and occupational employment related to Federal defense spending in the United States. The estimation technique utilizes detailed input-output relationships developed in the Bureau of Labor Statistics in support of the BLS projections program. Representative historical years are examined–the 1977 low following the cessation of Vietnam activities, the peak of the Reagan-era buildup in 1987, and the long slowdown during the Clinton Administration--as well as the BLS projection of defense spending for 2010. Alternative growth paths related to the September 11 atrocities are also examined.

Investment-Related Employment

Eric Figueroa, Bureau of Labor Statistics, U.S. Department of Labor

This study examines employment that is dependent on gross private domestic investment. Over the 1990s, investment spending grew faster than overall GDP, a trend expected to continue through 2010. Much of the growth resulted from robust spending on information processing equipment and software. Using an input-output approach developed in the Bureau of Labor Statistics, employment related to investment spending is estimated for the 1990-2000 period, and compared with that expected for 2010. The impact of changing investment spending is examined, and occupational employment data for 2000 and 2010 are compared.

New and Emerging Occupations

Olivia Crosby, Bureau of Labor Statistics, U.S. Department of Labor

Changing technology, demographics, and regulations can create new occupations and specialties. Job seekers and education administrators hope to plan for these developments. Forecasters want to measure them. New occupations, however, are difficult to distinguish and hard to predict. This presentation offers alternative definitions for new and emerging occupations, discusses ways to identify and track such occupations in the absence of clear measurement tools, and proposes examples of small occupations and specialties that may be poised to become important. Current efforts to gather data on new occupations are described, including possible uses of BLS projections.

This page intentionally left blank.

THE LABOR FORCE OVER THE NEXT 50 YEARS

Mitra Toossi
Bureau of Labor Statistics

This study examines the quantitative changes in the U.S. labor force from 1950 to 2050, with emphasis on the projected numbers for the next 50 years.

Based on the historical data from 1950 to 2000, the labor force in different sex, age, race and ethnicity categories has been projected for the 2000-2050 period.

Our analysis shows that the size, composition and characteristics of the labor force has gone through remarkable changes over the past 50 years and will continue to change in the next 50 years. One can say that the history of the U.S. labor force during this one hundred years is a story of 'constant change'.

The changes in population and the labor force participation rates during the past 50 years have created a steadily growing labor force which has aged, diversified and has witnessed an increasing presence of women. Factors affecting the labor force during this one hundred years can be summarized as follows:

- Women's participation: Rapid growth of women's participation rates for all age groups in the last fifty years, especially during the 70s will be replaced by stable and moderate growth in the next fifty years.

- Diversity: The labor force is expected to become more diverse in the future. With higher population growth and increasing participation rates, the share of the minorities in the work force is projected to expand substantially.

- Baby boomers: The impact of the baby boom generation on the labor force will continue. Just as their entrance swelled the ranks of the labor force in the last three decades, their exit will have a profound effect on the level and composition of the labor force in the next two decades.

The combined effect of the above factors is expected to substantially lower the growth of the labor force especially in the next two decades.

The Bureau of Labor Statistics carries out a medium-term labor force projection every two years covering a period of ten years. The latest projections was released recently and included the 2000-2010 period. This paper is an attempt to provide a long-term perspective of the labor force by looking 50 years into the future on a decennial basis.

The projection method in this study is similar to the one in the ten-year projection.

The future labor force participation rates for 136 different groups, including both genders, 17 age, and four race and ethnicity groups, are estimated based on the labor force participation behaviors of each group in the past. By applying the projected labor force participation rates in each group to the projected population within that group, the size of the labor force is estimated, both for the detailed categories as well as for the economy as a whole.

The detailed labor force participation rates were projected to the year 2015 and held constant thereafter. The uncertainties associated with the factors affecting the decision to participate in the labor force justify the constancy of the participation rates beyond 15 years. However, despite constant detailed participation rates for various population groups, the overall labor force participation rate is projected to decline through 2050 which reflects the impact of the changes in size and composition of the population especially the aging of the population.

Labor force growth

Chart 1 shows the labor force growth during the one hundred years from 1950 to 2050. In the 1950-60 period, the labor force grew 1.1 percent annually, the same rate as the growth of the total population. In the next decade, the labor force growth jumped to 1.7 percent. In that period, the baby boom generation began entering the labor force in massive numbers and overall participation rates increased significantly.

1970-80 period was a unique decade in the history of the labor force. The combined effect of continued absorption of the entire baby boom cohorts into the job market coupled with acceleration in the participation rate of women caused labor force to grow at unprecedented 2.6 percent annually.

In the 1980-90 period, a gradual slowdown occurred in the labor force growth mainly because almost all baby boomers had already entered the labor force. The labor force grew 1.6 percent in that period. The slow down in the labor force growth continued in the next decade when the labor force and population grew 1.1 percent.

It is projected that in the 2000-2010 period, the present growth rate of 1.1 percent will be sustained for both the population and the labor force.

The baby boomers, who will be between 46 to 64 years of age in 2010 and between 56 to 74 in 2020, will increasingly retire and exit the labor force during 2010's and 2020's. The boomers exit from the labor force in large numbers during these two decades decreases the growth of the labor force to a meager .4 percent.

During the 2030s and 2040s, a slight increase in the growth of the labor force is projected. As a response to the dramatic decrease in the availability of potential workers and the retiring baby boomers of the previous two decades, the Census Bureau has an assumption of anticipating an increase in the influx of migrants to the U.S. (about 1 to 1.5 million immigrants annually)

Population growth and participation rates are the main determinants and sources of the labor force growth. As the growth in participation rates subsides, population will be the only source of the labor force growth.

In the 1950s, population growth was solely responsible for the growth of the labor force. During the 1960s, population growth contributed to about 94 percent of the growth in the labor force.

In the 1970s, when the labor force participation of women experienced a rapid growth, 76 percent of the labor force growth was caused by the population growth and the rest was related to the growth of participation rates, mainly of women, at that time.

From 2000 to 2050, as the overall participation rate stops growing and even declines, participation growth exerts even less influence and the growth of the labor force will be mostly due to the impact of the population growth.

Population and labor force pyramids

Population pyramids are a convenient way to show the age and sex composition of the population. In a country with high fertility and high mortality, like the case of developing countries or the U.S. during the 1900, the shape looks like a pyramid. For a population with low fertility and low mortality, the shape is like a skyscraper with a pyramid on the top. The U.S. population is in the process of making a transition to the skyscraper shape.

Chart 2 shows the population and labor force pyramid for 1950. By convention, men are shown on the left and women on the right side of the population pyramids. Although it is hard to see, there are more baby boys born than baby girls. However, the higher mortality of males causes the population of men and women to be the same size around age 24.

The baby boom generation is the population born between 1946 and 1964. In this chart, the surge in the births of the early baby-boom generation is reflected in the long size of the 0-4 age group. The impact of the "birth dearth" which was the decline in the number of births between the late 1920s and early 1930s can be clearly seen in the indent of the population in the 15-19 age group.

The huge gap between the labor force of men and women is clearly visible on the chart. In 1950, the women's share in the labor force is 30 percent, in contrast to 70 percent for men.

The population pyramid in 2000 shows the effects of fifty years of change in the labor force and population composition. The overall shape is more like a skyscraper with a pyramid on the top. The labor force portion has a rectangular shape. (See chart 3)

The baby boom generation will be of ages 36 to 54 in 2000. Both the population and the labor force is older at this time.

In 2000, because of the narrowing of the gender gap in the labor force, the shape and size of the

pyramid for both men and women look very much alike.

And finally the shape of the projected labor force and population pyramid in 2050 is shown in chart 4. The pyramid in 2050 looks rectangular in shape for the higher age brackets, which is indicative of the aging of the population.

The baby boomers are 85 and over at this time and they are all out of the labor force. Women's lower mortality as compared to men's is responsible for the larger share of women in older age brackets.

The shape of the pyramid for both men and women has become even more symmetric which is a reflection of the further narrowing of the gender gap in the work force. At this point in time, the women's share in the labor force is projected to be 48 percent as opposed to 52 percent for men's.

Participation Rates

Overall labor force participation rate for the U.S. labor force was 59.2 percent in 1950. Participation rate reached 67.2 percent in 2000 and is projected to reach its highest point in 2010 at 67.5 percent. This rate is expected to be at 61.5 percent in 2050. Although the total participation rate remained in a tight range over the hundred years under this study, the labor force participation rates for men and women underwent substantial changes over the same period.

Men's participation rate has always been higher than women's at both the aggregate level and every age group. However, The aggregate participation rate of men was at its highest in 1950 at 86 percent. However, it has been on the decline since 1950 and is projected to continue to further decline during the projection years. It is projected that in 2050 it will reach 67 percent. The provision of the Social Security Act contributed to the long-term decline in the participation rates among men by promoting the early retirement due to the increased availability of pensions and disability awards. The Social security Act was amended in 1960 to make individuals under 50 eligible for disability payments. That again helped in lowering the overall participation rates especially among men.

(See chart 5).

In contrast, the women's participation rate which was at 34 percent at 1950 has been on the rise since that year reaching 60.2 percent in 2000. The participation rate will continue to increase until 2010, its highest point at 62.2 percent. It is projected that the participation rate will decline after that and reach 57 percent in 2050. The increase in the share of the higher age groups in the population that have lower participation rates causes the aggregate rate to decline for both men and women. The changes in the aggregate participation rates for women, men and total reflect changes in the age distribution of the population as well as the differences in the participation rate of various age, sex, race, and Hispanic origin groups.

Women's Participation

Among the factors that have contributed to the growth and development of the U.S. labor force, none has been as pronounced as the rise in the participation of women in the labor force during the last fifty years. In 1950, the overall participation rate of women was 34 percent. (See chart 6) This rate rose to 60 percent by 2000 and is projected to attain its highest level in 2010, at 62 percent. From then on, it is anticipated to decline slowly, reaching 57 percent in 2050. The projected decline after 2010 is due to the shifts in population to the older age groups with lower participation rates.

A number of factors were responsible for such a remarkable increase in the participation rate of women in the past. In a larger scale, the Civil Rights movement, legislation promoting equal opportunity in employment, and women's rights movement all created an atmosphere that was extremely hospitable to women working outside the home. Other social, economic, and demographic factors that induced women to join the work force were as follows:

- Women remained single more often.
- Of those who married, many did so late in their lives, and so, the median age-at-first-marriage increased substantially.
- Women elected to stay longer in school, achieved higher educational goals, and pursued better-paying careers.
- Women postponed childbirth to older ages and had fewer children than in previous decades. As a result of the availability of improved childcare, women tended to enter the labor force even before their children

started school, and were able to maintain longer job tenure than in previous periods.

Gender Share

The number of men in the labor force has always been greater than the number of women. However, over the last fifty years, the growth rate of women in the labor force has been significantly higher than that of men's. As a result, women's share in the labor force has increased substantially during the last 50 years. (See chart 7) Women's share in the labor force which was 30 percent in 1950, increased to 47 percent in 2000 in comparison to men's share of 53 percent. It is projected that in 2050 this share will be 48 percent for women and 52 percent for men. Moreover, the difference between the participation rates of men and women has been shrinking between 1950 and 2000 and is projected to further decline in the future. In 2050, the difference will be around 10 percentage points.

Growth of minorities

The growing share of minorities in the labor force has been an important development of the past several decades, and we project a continuation of this trend. The proportion of U.S. population composed of Hispanics and Asians and other increased significantly following the changes in immigration law in the 1960s and subsequent surge in immigration that began around 1970. It is projected that for the 2000-2050 period the two fastest growing groups will be "Asian and other" and Hispanics (chart 9).

Racial and ethnic shares

During the last fifty years of the Twentieth Century, racial and ethnic diversity of the U.S. population has grown tremendously. The greater diversity of the population has also resulted in an increased diversity in the labor force. The population and the labor force are divided into four major race and ethnicity categories: "white", "black" and "Asian and others" for the race group and "Hispanic origin" for the ethnicity category. Although Hispanics can be of any race, most report themselves as White.

As chart 8 suggests the white non-Hispanic group has the largest share of the US labor force, but that share has been on the decline and will continue to decline in the next 50 years. The decline of the white non-Hispanic group is accompanied by the faster growth of other racial and ethnic groups in the U.S. work force. Also, the upcoming retirement of the baby boomers, a group that has a large share of white non-Hispanic men, will lower the share of this group in the total labor force. The low fertility rate and low migration of the white non-Hispanics relative to other racial groups will lead to decreasing share of this group both in the population and the labor force.

Data for the race and ethnic categories were not available prior to 1980. In 1980, Blacks had the second largest share of the total labor force. The increase in the share of the blacks in the total labor force is mostly due to the faster growth of the black population caused by a higher fertility rate and a relatively high participation rate among black women.

The growing share of the Hispanics is mainly due to the high level of immigration of this group during the 1950 to 2000 years. These new immigrants have been mostly in the younger age cohorts with higher than average fertility rates. It is projected that in 2050, the share of the Hispanic-origin group in the labor force will be 23 percent.

Asians have been the fastest growing sector of the labor force in the past and are projected to remain so for the next fifty years. As a result of large number of immigrants in the last 50 years, this group, although small, would be the fastest growing part of the labor force and is projected to compose 10 percent of the labor force.

The future is unknown and often difficult to predict. Any attempt, like this study, to forecast the future, is an exercise in extrapolation of past trends and is based on the main assumption that the future will be subject to the same economic structures as in the past. However, by relaxing or changing any of the assumptions that were made earlier, one can arrive at differing sets of results. In fact, the future is always full of surprises and fifty years is a very long time to project. Therefore one should not be surprised if the future labor force changes in ways just as dramatic as those of the last fifty years.

Chart 1. The labor force will grow more slowly

Annual rates of change

50's	60's	70's	80's	90's	2000-2010	10's	20's	30's	2040-50
1.1	1.7	2.6	1.6	1.1	1.1	0.4	0.3	0.6	0.6
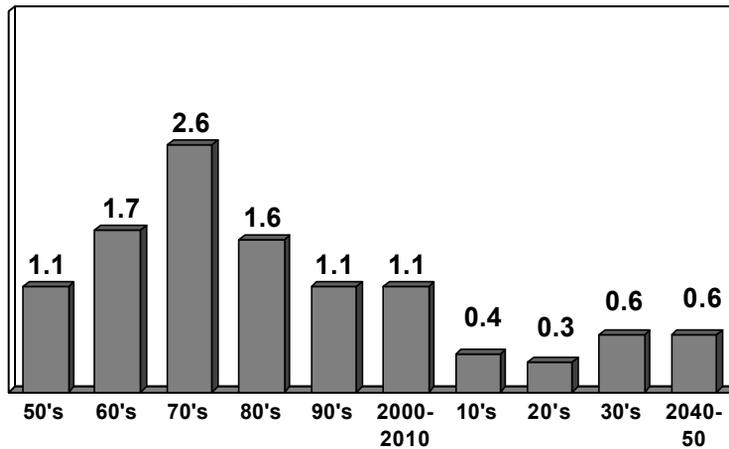
Source: Bureau of Labor Statistics

Chart 2. Population and labor force, 1950
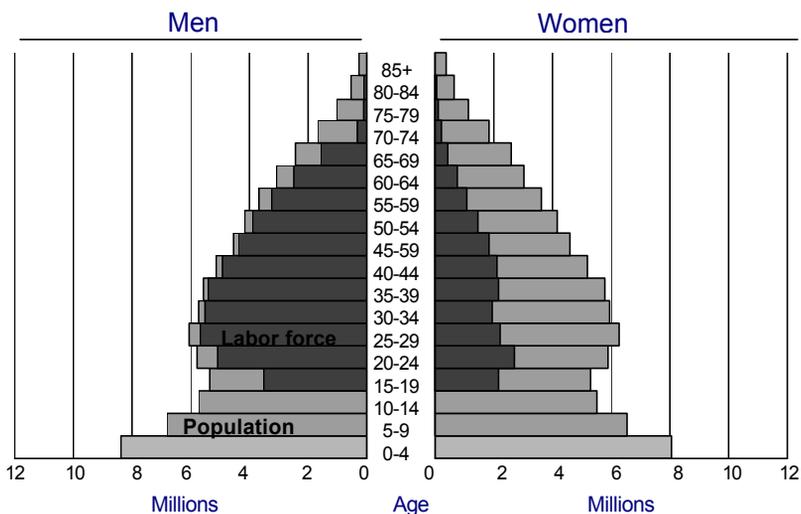
Source: Bureau of Labor Statistics

Chart 3. Population and labor force, 2000

Men **Women**

Labor force

Population

12 10 8 6 4 2 0   0 2 4 6 8 10 12

Millions   Age   Millions

Age groups (top to bottom): 85+, 80-84, 75-79, 70-74, 65-69, 60-64, 55-59, 50-54, 45-59, 40-44, 35-39, 30-34, 25-29, 20-24, 15-19, 10-14, 5-9, 0-4

Chart 4. Population and labor force, projected 2050

Men   Women

Labor force

Population

12 10 8 6 4 2 0   0 2 4 6 8 10 12

Millions   Age   Millions

Age groups (top to bottom): 85+, 80-84, 75-79, 70-74, 65-69, 60-64, 55-59, 50-54, 45-59, 40-44, 35-39, 30-34, 25-29, 20-24, 15-19, 10-14, 5-9, 0-4
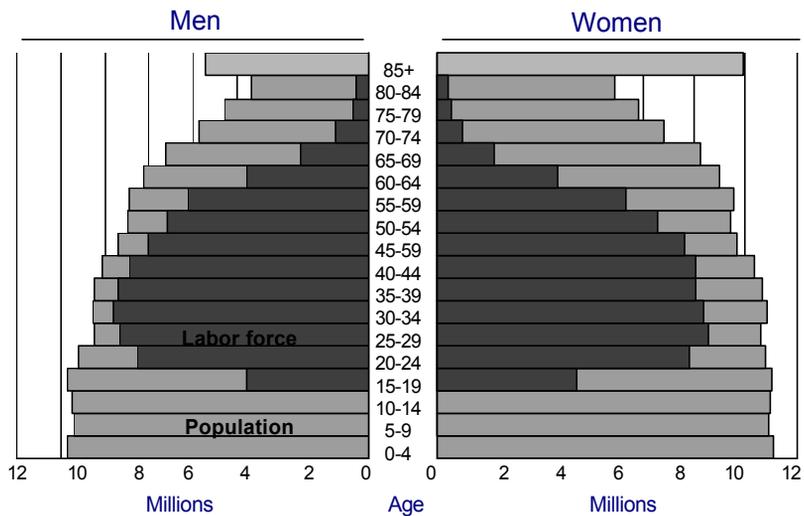
Chart 5. Aggregate labor force participation rates decrease because of aging

Percent

Men

86

Total

Women

34

67

57

1950 1960 1970 1980 1990 2000 2010 2015 2020 2030 2040 2050
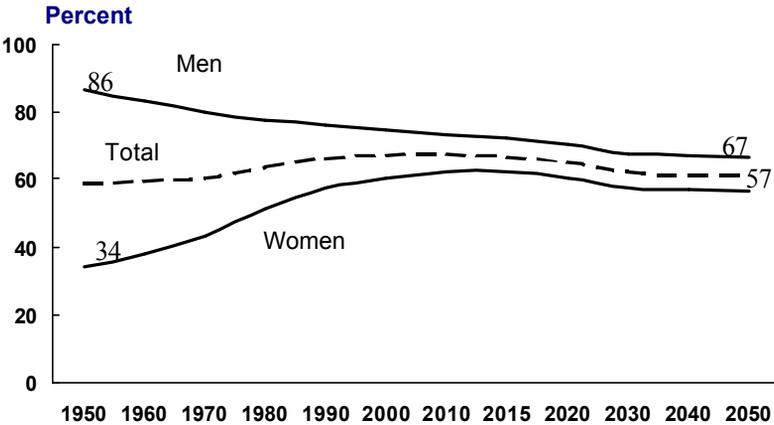
Source: Bureau of Labor Statistics

Chart 6. Labor force participation of women slows because of aging

Participation rate

86

75

67

60

57

34

Women

Men

1950 2000 2050   1950 2000 2050
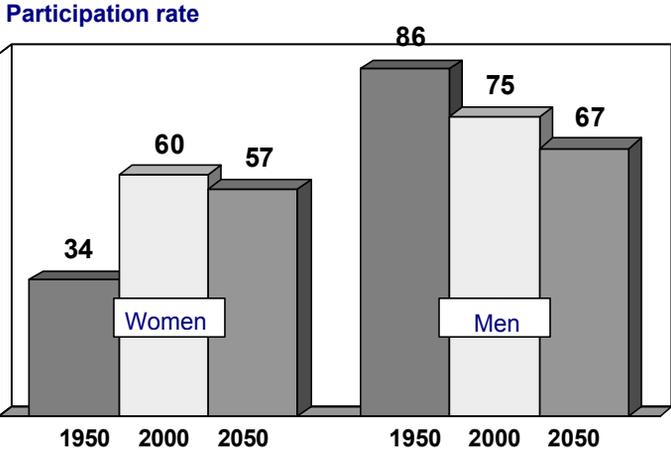
Source: Bureau of Labor Statistics

Chart 7. Women's share of the labor force increases

Percent of labor force
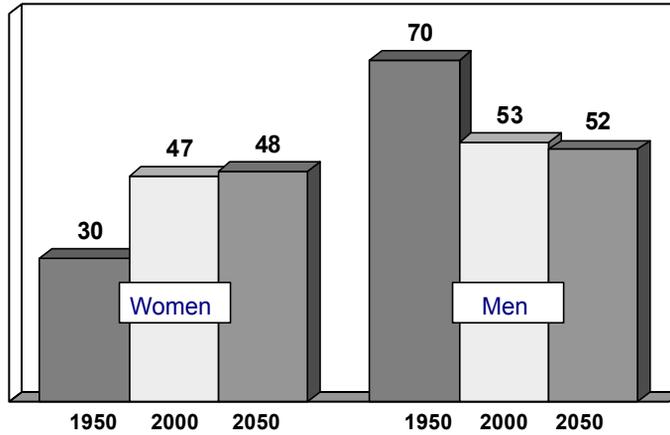
Source: Bureau of Labor Statistics

Chart 8. White, non-Hispanics remain the largest group of workers
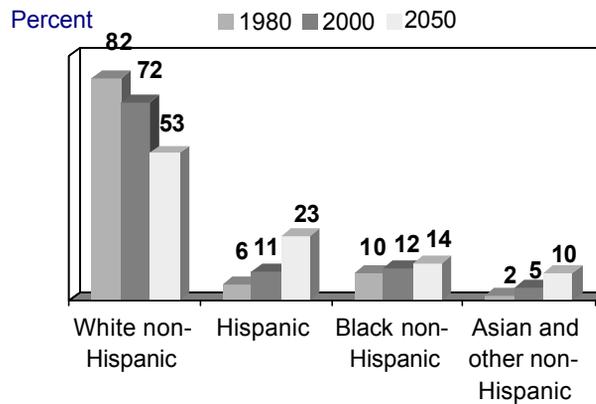
Source: Bureau of Labor Statistics

Chart 9. Labor force growth rates of minorities outpace whites

Percent change, 2000-2050

Category	Value
Total	36
White non-Hispanics	0
Black	63
Hispanic	196
Asian and other	200
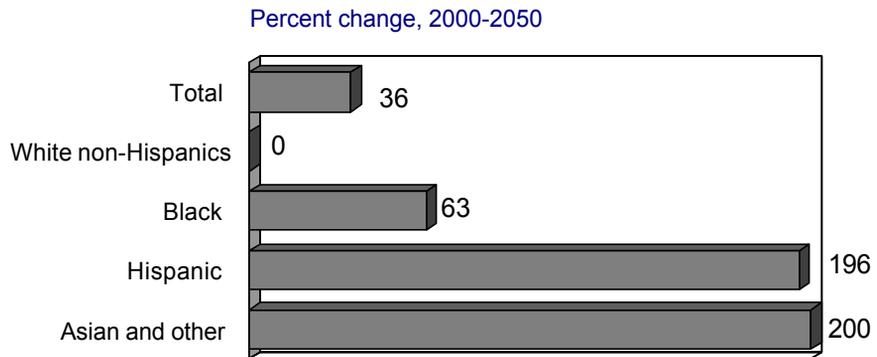
Source: Bureau of Labor Statistics

This page intentionally left blank.

The BLS Defense-Related

Employment Projections

Jeffrey C. Gruenert, Economist

gruenert_j@bls.gov

•In Econ 101 many of us were introduced to the equation Y = Consumption + Investment + Government + net foreign trade.   I study the Government part of the equation for the BLS.

•Today I'm going to discuss research I conducted to derive defense-related employment and project it through the year 2010. This is an offshoot of our 2000-2010 projections cycle, which was completed prior to the events of 9-11-01 and their aftermath, so recent budget supplements and the proposed Fiscal 2003 Budget are not reflected in the study.

•I made use of the same methodology as the previous speaker, Eric Figueroa, and so will not repeat the comments he made earlier due to time constraints.

•One reason that I decided to examine defense-related employment is that historically American defense spending has been a significant proportion of GDP. The US economy has had relatively stable defense spending since WWII, consisting of hundreds of billions of dollars every year creating millions of good, high-paying jobs. We are going to examine the year 2000, our projection for 2010, and both the high and low watermarks since the Vietnam War for perspective.

Office of Occupational Statistics and Employment Projections

- ♦ Defense Spending and
  Defense-related Employment

- ♦ Government / Private Sector

- ♦ Industry Employment

- ♦ Occupational Employment

•I have 4 Handouts which address these points in considerable detail:

•Table 1 shows defense spending and the breakdown of defense-related employment between the Public sector(including the Armed forces) and Private sector for the years 1977, 1987, 2000 and projected 2010 and the division of the private sector into direct and indirect employment.

•Table 2 has defense-related employment by industry sector for the years presented above.

•Table 3 shows defense-related employment by major occupational sector for the years 2000 and 2010. Defense spending creates jobs within the Professional and construction groups while the other groups actually lose related occupational employment .

•Table 4 features the alternative defense spending and related employment that would result from the President's proposed FY2003 defense budget compared to the BLS 2000-2010 projections.

Defense Spending 1977-2000 and Projected 2010

Billions of real $
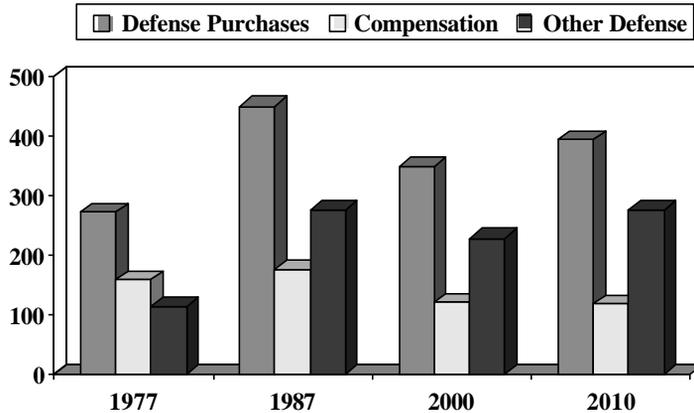
- This material is in the top 2 fields of handout Table1 in chain weighted 1996 $:
- Real defense spending by the Federal Government retrenched in the years following the Vietnam War to a low of $271.1 billion in 1977- 6.0% of GDP.
- Defense spending quickly escalated in the 1980's to a post-Vietnam War high in 1987 due to Cold War tensions between the Soviet Union and the United States. US defense spending culminated in expenditures of $450.2 billion- 7.4 percent of GDP. After the Cold war ended, there was a long defense spending draw-down in the 1990's which climaxed in 2000 with spending at $349.0 billion-3.78% of GDP. We projected defense spending to increase in 2010 to $392.7 billion-3.06% of GDP.
- The numbers show that a considerable amount of Defense spending goes to consumption accounts to cover compensation of Federal civilian employees at DOD and elsewhere, as well as members of the Armed Forces, and that these expenses change relatively slowly over time due to a relatively stable workforce.
- Compensation   • 1977   • 1987   • 2000   • 2010
                    • $157.1   • 176.2   • 120.9   • 116.8

Compensation as a percentage of defense spending
                    • 57.9%   • 39.1%   • 34.6%   • 29.7%

- The private sector benefits from defense spending largely through consumption expenditures which go disproportionately for durable goods and services, primarily benefiting the manufacturing and services industries.

Employment Related to Defense Spending, 1977-2010
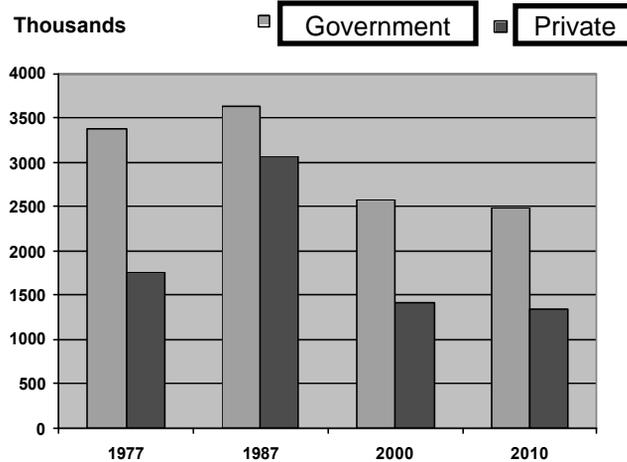
•This material is illustrated in the bottom 2 fields of handout Table 1:

•Each Presidential administration and Congress have different public policy priorities, resulting in increasing levels of Federal employment in some programs and declines in others. A decline in job requirements does not necessarily result in fewer jobs in the economy or into unemployment; other factors of demand, such as exports, investment, or personal consumption, could offset this decline.

•In 1977, defense-related employment reached a post-Vietnam low at 5,120,000, two thirds of whom were employed in the public sector.

•Employment increased steadily and built to its peak in 1987 of 6,694,000, only 54% of whom were employed in the public sector. 1.57 million defense-related jobs were generated.

• In 2000 spending once again declined, to 3.98 mil, 64% of which were in the public sector. A record 2.7 million Americans lost defense-related jobs between 1987 and 2000, and defense- related employment is projected to continue falling through 2010 with projected employment of 3.8 million, 65% of whom are expected to be in the public sector. The greatest reduction in public sector employment from '87 to 2000 occurred among members of the military as troop strength fell by 794,000, a 35 percent reduction. Further reductions among the armed forces and civilian Government employees through 2010 will be tempered by Nations' need to maintain a functional level of readiness

Industry Employment

Defense-related Employment by Major Industrial Sector

Legend: TOTAL — Government — Manufacturing — Services
(X-axis: 1977, 1987, 2000, 2010; Y-axis: 0 to 5000)
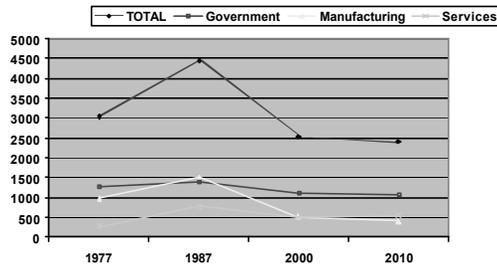
- This material is illustrated by table 2 of the handouts.

- Defense-related employment fluctuated from a low of 3.05 million in 1977 to 4.45 million in 1987. In 2000 it was 2.53 million and it is projected to be 2.41 million in 2010.   The decline over the 2000-2010 period (despite increased defense spending predicted for 2010) is due to our assumption of rising productivity over the projection period.

- The Government, manufacturing, and services industries are, and will continue to be, the most important sources of defense-related industry employment numerically.

- The top 5 industries (those with the greatest share of defense-related employment) are as follows:

	2000	2010
180 Federal General Government	60.06	62.08
37 Ordnance and ammunition	38.48	32.57
61 Search and navigation equipment	32.60	29.21
58 Ship and boat building and repairing	32.31	24.47
57 Aerospace	29.58	22.53

Many defense-related industries in the private sector will see employment continue to decline throughout the remainder of the projection period as a result of increasing productivity.

Direct and Indirect Industry Employment

Industry	2000 Total	Direct	2010 Total	Direct
Construction	84.5	74.9	102.9	93.9
Aerospace	163	156.9	147.7	141.3
Computer	118	115.5	146.5	143.7
Federal Government	1082.2	same	1037.2	same

•Employment generated by defe nse spending can be divided into two categories—direct and indirect. Nearly all industries have some combination of direct and indirect defense-related employment. Direct defense-related employment evolves from the Department of Defense's spending on all final goods and services. Indirect defense-related employment arises from the need to supply inputs to the producers of these final goods and services.

•Examples of total and direct defense-related employment in selected industries:

Industry	2000 Total	Direct	2010 Total	Direct
Construction	84,500	74,900	102,900	93,900
Aerospace	163,000	156,900	147,700	141,300
Computer	118,000	115,500	146,500	143,700
Federal Government	1,082,200	same	1,037,200	same

Defense-Related Occupational Groups 2000-2010

Percent change, 2000-2010

Occupational Group	Percent change
Production	-60.1
Office and administrative	-31.7
Management, business and financial	-15.5
Installation, maintenance and repair	-12.4
Transportation and material moving	-8.
Service	-.4
Sales and related	-0.7
Farming, fishing and forestry	0.1
Professional and related	3.4
Construction and extraction	9.2
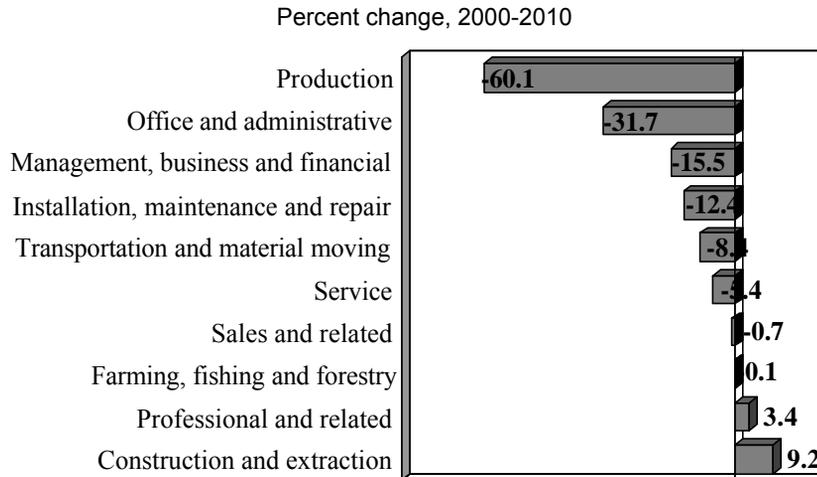
Source: Bureau of Labor Statistics

•This material is illustrated by table 3 of the handouts: Defense-related employment among civilian employees of the Department of Defense and other Federal agencies are, and will continue to be, centered in the Management, Professional, and Office and Administrative support occupations.

	Defense	All industries
Management,	17.2654	18.6
Professional	21.1618	9.5
Office and administrative	19.9723	17.5

•All occupational groups will decline in the number of defense-related employees between 2000-2010 except the Professional and related which are expected to add 3,400 jobs and Construction which will increase by 9,200.
• These numbers reflect the increased spending over the course of the period from 2000-2010. Much of this spending will be for high-technology equipment and machinery manufacturing, where productivity is very high. The increase in professionals reflects a demand for highly trained individuals to operate this equipment and the increase in construction reflects that this group is in a lower productivity classification, requiring more employees.

Real Defense Spending
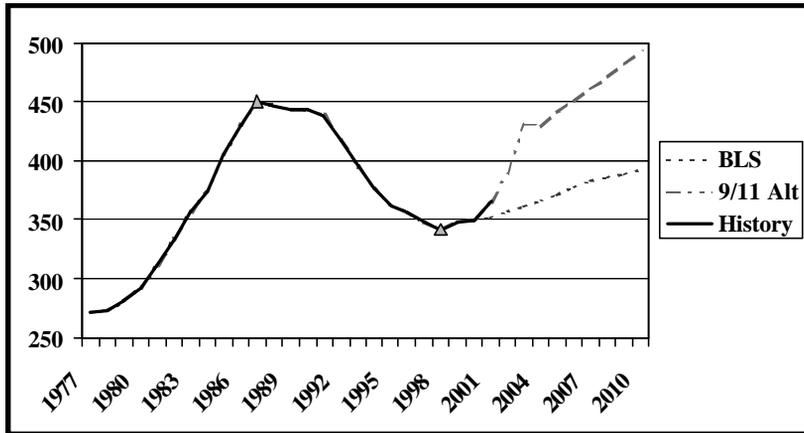
•This material and that of the following 2 slides can be found on handout table 4:

•Since September 11 of last year, President Bush has proposed a sustained 5-year increase of $120 billion in military spending, from $331 billion in 2002 to $379 billion in 2003 and $451 billion in 2007.

•If passed into law, these increased spending proposals would result in an increase in both industry and occupational employment as illustrated above, with the possibility of employment reaching and exceeding the 1987 high watermark.

•According to the Washington Post, the proposed 2003 Budget Plan for that year alone would give the military the biggest increase in two decades, matching the previous Bush administration's budget when adjusted for inflation. These calculations do not include the effect of the administration's recent $27 billion supplemental spending request for fiscal 2002.

•The increases would be significant but they would also take place from very low starting levels of defense spending.

2010 Alternate Employment Growth After 9-11

Thousands of Jobs

■ BLS   □ 2003 Budget

Category	
Government (including armed forces)	
Services	
Wholesale & retail trade	
Transportation, communications, & utilities	
Manufacturing	
Construction	
Agriculture & mining	
Defense-related employment (1,000's)	

0   500   1,000   1,500   2,000   2,500   3,000   3,500   4,000   4,500
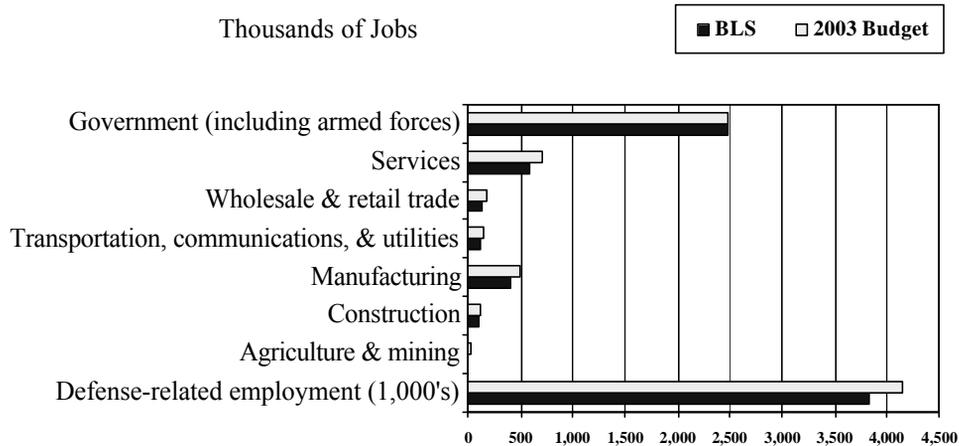
•In 2010 we have projected an increase of 320,100 employees over all industries except government, which was actually expected to decline slightly. We previously expected employment to decline by 121,700. The largest increases are in the Manufacturing and Services Industries which reflects the nature of these gains, many of which will be in high-technology industries.

Employment Difference BLS
Forecast and Proposed Expenditures
under Bush 2003 Budget
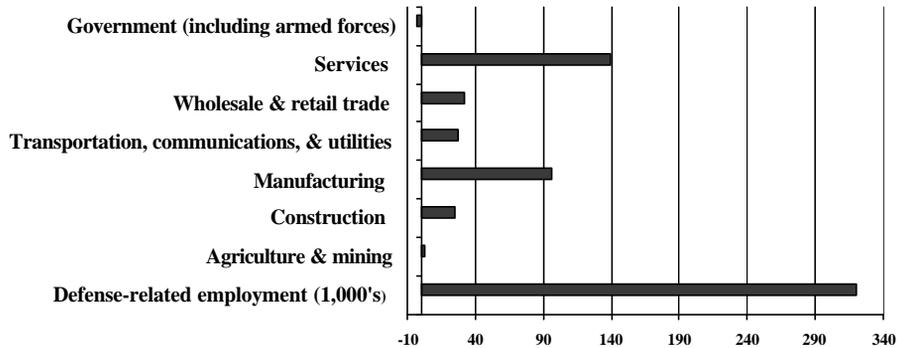
• Employment generated by defense spending can be divided into two categories—direct and indirect. Nearly all industries have some combination of direct and indirect defense-related employment. Direct defense-related employment evolves from the Department of Defense's spending on all final goods and services. Indirect defense-related employment arises from the need to supply inputs to the producers of these final goods and services.

• In terms of where the spending will take place, thereby creating employment, this slide shows that the largest change is in services and manufacturing.

INVESTMENT-RELATED EMPLOYMENT
Eric B. Figueroa, Bureau of Labor Statistics (BLS)
Office of Occupational Statistics and Employment Projections

Introduction

In 2000, 13.4% of all jobs in the US economy were dependent on gross private domestic investment (GPDI), making investment spending responsible for 17.9 million jobs that year (Table 1). The influence of investment spending on employment is expected to increase. In 2010, the share of jobs in the total economy resulting from investment spending is expected to reach 13.9%, or 21.7 million jobs.

This article examines domestic employment related to private investment spending. Investment-related employment for 1990 and 2000 is compared with that expected for 2010, using the most recent economic and employment projections from the BLS Office of Occupational Statistics and Employment Projections. The number and type of jobs dependent on investment were estimated using an input-output model that enables one to trace the purchase of a good or service through the entire production chain. With this approach, it is possible to determine employment required in each industry, including the industries that supply inputs to the production process of a good or service. In addition, an industry-occupation matrix was used to determine the effect of investment spending on occupational employment in 2000 and 2010.

Over the 1990's, investment spending grew strongly due to increased expenditures on software and on computer equipment. As a result, employment in service-producing industries, which include producers of software, also grew strongly. Employment growth was slower in goods-producing industries, which include manufacturers of computer equipment. Despite large investment expenditures on equipment, increased efficiencies among manufacturing industries restrained job growth in the goods-producing sector. Over the 2000 to 2010 period, investment-related employment in service-producing industries will continue growing, while goods-producing employment will decline slightly.

Declining employment in goods-producing industries will result in declining employment in production occupations. Employment gains in service-producing industries will lead to strong growth among professional and related occupations and service occupations. Over the forecast period, the largest employment increases will be in three professional jobs found predominantly in computer and data processing services, the industry, which produces software. These occupations are computer applications software engineers, computer support specialists, and computer systems software engineers.

Table 1. Investment-related employment, by major industry sector

Category	Wage and salary employment (thousands)			Percent of total employment			Change		Average annual rate of change	
	1990	2000	010	1990	2000	2010	1990-2000	2000-2010	1990-2000	2000-2010
Total Employment	111,580.4	133,740.6	155,722.3	100.0	100.0	100.0	22,160.2	21,981.7	1.8	1.5
Investment-related employment	13,163.0	17,910.2	21,659.6	11.8	13.4	13.9	4747.2	3749.4	3.1	1.9
Goods Producing	7,984.9	9,680.3	9,602.2	29.9	34.7	32.4	1695.4	-78.1	1.9	-0.1
Agriculture, forestry, and fisheries	109.9	184.2	201.7	6.1	8.3	7.7	74.3	17.5	5.3	0.9
Mining	273.0	249.7	240.5	38.5	46.0	49.3	-23.3	-9.2	-0.9	-0.4
Construction	3,063.3	4,151.8	4,464.8	59.8	62.0	59.4	1088.5	313.0	3.1	0.7
Manufacturing	4,538.9	5,094.5	4,695.1	23.8	27.6	24.7	555.6	-399.4	1.2	-0.8
Service producing	5,178.0	8,229.9	12,057.4	6.1	7.8	9.6	3051.9	3827.5	4.7	3.9
Transportation	401.1	582.8	706.9	11.4	12.9	12.9	181.7	124.1	3.8	1.9
Communications	116.2	177.9	191.0	8.9	10.9	10.0	61.7	13.1	4.4	0.7
Utilities	67.6	72.3	85.3	7.1	8.5	9.6	4.7	13.0	0.7	1.7
Trade	2,318.1	2,590.9	3,364.0	9.0	8.5	9.8	272.8	773.1	1.1	2.6
Finance, insurance, and real estate	356.1	518.9	655.2	5.3	6.9	7.9	162.8	136.3	3.8	2.4
Services	1,793.6	4,134.2	6,879.2	6.3	10.3	13.0	2340.6	2745.0	8.7	5.2
Government	125.3	153.0	175.8	0.7	0.7	0.8	27.7	22.8	2.0	1.4

Source: Bureau of Labor Statistics

Investment spending

Gross private domestic investment (GPDI) is estimated by the Bureau of Economic Analysis, U. S. Department of Commerce, as part of the national income product accounts. GPDI includes both expenditures on fixed capital goods—the purchase of equipment, software and structures by business and nonprofit institutions—and the value of changes in private inventories.

From 1980 to 1990, the average annual rates of growth for both investment spending and Gross Domestic Product (GDP) were nearly the same: 3.3% for investment spending and 3.2% for GDP (Table 2). However, these rates diverge over the 1990's, as increasing demand for equipment and software drive up the level of investment spending. From 1990 to 2000, the 6.9% growth rate for investment was more than double that of GDP. Over the 2000 to 2010 period, growth in investment spending will remain strong, although at a somewhat slower rate than the previous decade.

Strong growth in investment spending has resulted from the increasing demand for computer equipment and software. In 1980, equipment and software spending was concentrated on the other equipment category.

Purchases in this category—comprising industrial equipment, transportation equipment except light vehicles, and miscellaneous equipment—represented $264 billion in purchases. By contrast, expenditures on information processing equipment, such as computer equipment and software, were much smaller.

From 1990 to 2000, computer and software expenditures grew strongly, as businesses and other organizations sought to improve productivity through the use of computer networks, inter and intra-nets, and electronic commerce. Between 2000 and 2010, spending on computer equipment and software will again rise sharply, although at slower rates of growth than seen in the previous two decades.

Commodity purchases

Investment spending on equipment, software, and structures, involves the purchase a variety of commodities. Certain industries are responsible for producing these goods or services, and others provide the materials needed to produce these commodities. It is in all these industries that investment spending generates employment—a purchase triggers the need for both commodity output and the workers needed to make the good or service.

Table 2. Gross private domestic investment, 1980, 1990, 2000, and projected 2010

Category	Billions of chained 1996 dollars				Average annual rate of change		
	1980	1990	2000	2010	1980-1990	1990-2000	2000-2010
Gross domestic product	$4,900.9	$6,707.9	$9,224.0	$12,835.6	3.2	3.2	3.4
Gross private domestic investment (GPDI)	655.3	907.3	1,772.9	2,953.8	3.3	6.9	5.2
Fixed nonresidential investment	466.4	641.7	1,350.7	2,461.6	3.2	7.7	6.2
Equipment and software	262.2	415.7	1,087.4	2,216.3	4.7	10.1	7.4
Light vehicles	31.9	51.9	125.5	177.3	5.0	9.2	3.5
Computers	1.2	14.2	290.3	1,195.2	28.1	35.2	15.2
Software	10.6	45.9	187.6	613.0	15.8	15.1	12.6
Communication equipment	29.1	43.0	131.4	203.6	4.0	11.8	4.5
Other equipment	264.2	282.2	433.8	655.9	0.7	4.4	4.2
Nonresidential structures	223.2	236.1	272.8	330.1	0.6	1.5	1.9
Public utilities	47.0	33.0	48.5	50.3	-3.5	3.9	0.4
Mining and exploration	36.0	21.3	23.5	26.2	-5.1	1.0	1.1
Buildings and other structures	133.0	181.9	201.8	254.3	3.2	1.0	2.3
Fixed residential investment	210.1	253.5	371.4	464.5	1.9	3.9	2.3
Residential structures	205.9	247.3	361.8	450.1	1.8	3.9	2.2
Landlord durables	4.3	6.2	9.6	15.0	3.8	4.5	4.6
Change in private inventories	-10.5	16.5	50.6	58.6	…	11.8	1.5
GPDI share of GDP	13.4	13.5	19.2	23.0	…	…	…
GPDI Residual[1]	-97.2	-36.0	-91.5	-745.7	…	…	…

[1]The GPDI residual is the difference between total GPDI spending and the sum of the most detailed GPDI categories.
Sources: Historical data, Bureau of Economic Analysis; projected data, Bureau of Labor Statistics

In 1990, purchases of construction commodities were the largest among purchases by major commodity group (Table 3). This reflects the historical importance of investment spending on structures, which in 1990 still exceeded spending on equipment and software. By 2000, purchases of manufactured commodities take the lead, reflecting the shift in investment spending towards purchases of equipment, particularly computer equipment. The gap between manufacturing and construction commodity output is expected to widen over the forecast period.

Wholesale and retail trade was the third largest commodity purchase in 1990. This spending represents margins paid to wholesale and retail trade establishments for the service of supplying and distributing commodities. Although purchases of these trade services will see strong growth over the forecast period, by 2010 they will be exceeded by purchases of services commodities.

Services commodities, which include software, were the fourth largest commodity purchase in 1990. By 2000, increased investment spending on software led to strong growth in the level of these commodity purchases. Purchases of service commodities will continue to grow strongly over the forecast period, in line with rising software investment.

Industry employment

Using an input-output system, analysts can derive the level of industry output necessary to satisfy investment-related demand for commodities. Given these output levels, the system then derives the required level of employment to produce the output. Using these methods, industry employment was calculated by major sector for 1990, 2000, and 2010 (Table 1).

Employment levels are not simply a function of investment demand. From 1990 to 2000, the rate of increase in investment-related employment, 3.1 %, is less than the rate of increase of investment spending, 6.9% (Table 2). This is due to productivity growth—as real investment spending increased, firms met this demand by producing more with a given number of workers and by investing in labor-saving technologies. In addition, some of the difference may have been attributable to changes in the input requirements of producing industries. As new technologies are developed and input requirements change, employment is affected if the labor intensity of the contributing industries vary.

Among manufacturing industries, productivity increases have led to weak or declining investment-related employment growth. From 1990 to 2000, the demand for manufactured commodities grew at an 8.7% annual average rate; however, manufacturing employment only grew at a 1.2% rate over the same period. From 2000 to 2010, employment in this sector is expected to decline by 0.8% despite continued strong demand for manufactured goods, especially computer equipment.

By contrast, investment-related employment in services industries grew strongly between 1990 and 2000, increasing at an annual rate of 8.7%. This trend is expected to continue, and a 5.2% growth rate is expected over the forecast period. Strong employment growth in services industries is largely due to increased demand for software, a product of the computer and data processing services industry. Production in services industries are generally not as easy to automate as manufacturing processes; therefore output growth translates more directly into increased employment.

Table 3. Investment-related commodity purchases, by major commodity sector

Category	Based on billions of chained (1996) dollars			Average annual rate of change	
	1990	2000	2010	1990-2000	2000-2010
All Sectors	893.3	1,695.6	2,886.6	6.6	5.5
Goods Producing	738.9	1,248.4	1,905.0	5.4	4.3
Agriculture, forestry, and fisheries	2.2	4.9	2.1	8.4	-8.1
Mining	21.1	29.5	29.3	3.4	-0.1
Construction	428.6	551.4	671.2	2.6	2.0
Manufacturing	287.0	662.6	1,202.4	8.7	6.1
Service producing	154.4	447.2	981.7	11.2	8.2
Transportation	6.7	14.0	27.6	7.6	7.0
Communications	4.5	13.4	15.5	11.5	1.5
Utilities	0.0	0.0	0.2	--	18.6
Trade	67.6	137.9	303.3	7.4	8.2
Finance, insurance, and real estate	29.9	51.2	69.4	5.5	3.1
Services	61.9	253.8	591.4	15.2	8.8
Government	0.0	0.0	0.0	--	--

Source: Bureau of Labor Statistics

Employment by major occupational group

In 2000, the largest occupational group in terms of investment-related employment was construction and extraction occupations followed by production occupations (table 4). This ranking reflects the historical importance of investment spending on structures and equipment. These purchases generate employment in construction and manufacturing industries where most of these workers are employed.

The third and fourth largest occupational employment groups were, respectively, office and administrative support occupations, and professional and related occupations. Large employment in these occupational groups reflects the growing importance of software purchases from services industries where most of these workers are found.

Over the forecast period, the ranking of these four occupational groups will change due to changes in investment spending. Between 2000 and 2010, professional and related positions will grow at a rate of 4.7%, to become the largest occupational category. This will result from growing employment in services industries as they strive to meet investment demand for software.

Employment in construction and extraction occupations will increase at a slower rate, 0.9%, reflecting the relatively slower growth of investment demand for structures. In 2010, this category will be the second-largest occupational group in terms of investment-related employment.

In 2010, the third and fourth largest groups will be, respectively, office and administrative occupations, and production occupations. Office and administrative

positions will increase at a 1.8% annual average rate of growth. Most of this growth will come from increased employment in services and trade. Overall growth will be dampened, however, by declines in office and administrative employment in construction and manufacturing industries. Declines in manufacturing employment will contribute to declines of 0.2% in employment in production occupations. Despite strong growth in purchases of manufactured commodities, increased efficiencies in manufacturing industries, particularly in computer and office equipment, are expected to result in employment declines in these occupations over the forecast period.

Employment by detailed occupation

Which occupations will contribute most growth to investment-related employment over the forecast period? Not surprisingly, three computer and software-related occupations will see the largest gains (Table 5). These gains are attributable to strong growth in the services sector, particularly the computer and data processing services industry which develops software. Together, these three occupations will add 574,000 jobs, nearly a 15% of all jobs growth generated by investment-related spending.

Computer applications software engineers will gain the most employment from 2000 to 2010. These workers analyze users' needs and design, create, and modify general computer applications software or specialized utility programs. The occupation with the second largest employment gain will be computer support specialists and systems administrators. These workers provide technical assistance, support, and advice to end users of computer systems; and design, install and

Table 4. Investment-related employment, by major occupational group

Category	Employment (thousands)		Average annual rate of change
	2000	2010	2000 -2010
Total, all occupations	17,910.2	21,659.6	1.9
Management, business, and financial occupations	1,881.0	2,358.7	2.3
Professional and related occupations	2,239.2	3,528.6	4.7
Service occupations	844.5	1,158.7	3.2
Sales and related occupations	1,306.2	1,764.2	3.1
Office and administrative support occupations	2,814.5	3,377.4	1.8
Farming, fishing, and forestry occupations	106.7	98.6	-0.8
Construction and extraction	3,132.4	3,432.4	0.9
Installation, maintenance, and repair occupations	912.3	1,034.6	1.3
Production occupations	3,115.0	3,060.5	-0.2
Transportation and material moving occupations	1,558.4	1,845.8	1.7

Source: Bureau of Labor Statistics

Table 5. Investment-related employment, by detailed occupation

Category	Employment (thousands)		Change	Average annual rate of change
	2000	2010	2000-2010	2000-2010
Computer software engineers, applications	163.5	392.5	229.0	9.2
Computer support specialists	133.9	324.9	191.1	9.3
Computer software engineers, systems software	127.7	281.4	153.6	8.2
Retail salespersons	222.9	361.3	138.4	4.9
Computer systems analysts	122.4	242.6	120.2	7.1
Customer service representatives	244.7	352.7	108.0	3.7
Cashiers, except gaming	160.8	264.8	104.0	5.1
Office clerks, general	312.4	403.6	91.2	2.6
Computer programmers	204.8	290.7	85.9	3.6
General and operations managers	396.5	477.4	80.9	1.9

Source: Bureau of Labor Statistics

support an organizations computer network. The third largest employment increase will be found among computer systems software engineers. These workers coordinate the construction and maintenance of a company's computer systems and plan their future growth.

Nearly all of the increase in these three occupations will be found in the computer and data processing industry where rapid growth in response to the growing demand for software investment, has resulted in large employment increases. Employment data for other occupations expected to gain the most employment over the forecast period are found in table 5.

Conclusion

Increasing expenditures on information processing equipment have had varied impacts on investment-related employment.

Strongly rising software purchases are expected to translate directly into employment increases in the services sector. Gains will be greatest in computer and data processing services, which develops software. The work done in this industry requires a high level of skill

and is not easily automated. As a result, increasing demand for software will lead to large employment gains among professional and related occupations found in this industry. Growth will be strongest among computer and software-related professionals, especially computer software engineers and computer support specialists.

Strong growth in computer purchases has not translated into strong employment growth in manufacturing, which produces computer equipment. Increasing efficiencies among that sector's industries, including the computer and office equipment, have restrained employment from rising strongly along with output. As a result, employment in this industry is expected to decline over the forecast period, causing employment among production occupations to decrease.

Rising computer purchases are expected to contribute to employment increases in the trade sector. As more businesses purchase computer equipment through wholesale and retail trade establishments, this sector's output will rise, leading to growth in sales and related occupations over the projection period.

This page intentionally left blank.

NEW AND EMERGING OCCUPATIONS
by Olivia Crosby
Bureau of Labor Statistics
Federal Forecasters Conference 2002

Changing technology, demographics, and regulations create new occupations and specialties. Jobseekers and education administrators hope to plan for these developments. Forecasters want to measure them. New occupations, however, are difficult to distinguish and harder to predict.

This paper offers alternative definitions for new and emerging occupations and explores why occupations change. Then, it reviews past attempts to identify and measure new occupations and describes additional data that may become available. Examples of possible new occupations are given along the way.

What is an occupation?

Before they can decide if an occupation is new, researchers must decide what they mean by "occupation." The definitions researchers choose affect the number and types of new occupations they identify.

For example, the 1970 definition cited by sociologist Eliot Freidson—"A set of jobs in which workers perform the same tasks using the same methods and pass those methods onto recruits"—would lead to an alarming number of new and emerging occupations. Even the smallest change in job duties or skills will push a job out of an existing occupation.

A more useful definition is the one provided in the latest version of the Standard Occupational Classification Manual (SOC): "A group of jobs in which the workers perform similar tasks at similar skill levels."

This latter definition is more flexible. But it also leaves a question: how similar must the tasks of an occupation be? In other words, when are jobs different enough to be in separate occupations?

The answer depends on who is classifying the occupations. Jobseekers and educators want detailed information about new trends in jobs and skills. They hope to identify any important occupational change, so they want to identify many new occupations. Labor market statisticians, on the other hand, want fewer occupations so their classification systems are manageable and lasting and their data comparable from year to year.

The occupation of geriatric social workers illustrates the difficulty of deciding whether tasks are unique enough to be a new occupation. Within the broad occupational group of social worker are a few well-established detailed occupations, including mental health social worker and medical and public health social worker. Geriatric social workers combine some of the tasks of both of these occupations and add some other duties. They perform counseling, referral services, and case management and usually help their patients deal with chronic medical problems. Geriatric social workers treat older patients exclusively, addressing concerns unique to that age group. But are geriatric social workers' tasks unique enough to warrant classification as an occupation? Or are these workers part of an existing occupation? Education planners and jobseekers want to be aware of both the new skills geriatric social workers need and new job possibilities. Thus, lists aimed at jobseekers often cite geriatric social worker as a new occupation, although classification systems usually do not.

Researchers not only have to decide how distinct job duties must be, they also must choose which duties are important. If core tasks are the same for a group of jobs, the jobs are in the same occupation—even if less important tasks are different. Trying to classify geographic information systems (GIS) specialists brings this issue to the fore. GIS specialist is a common title in job postings and often appears in lists of new and emerging occupations. But job descriptions suggest that these jobs may be part of a number of existing occupations. Many GIS specialists program or maintain databases of geographic information, so they might be programmers, software engineers, or database administrators. Other specialists concentrate on creating maps and charts, acting as mapping technicians. And others use GIS while planning cities, designing marketing campaigns, or conducting geographic research, possibly making them urban planners, market researchers, or geographers. If workers perform several of these tasks in one job, or if working with GIS is their most important task, then perhaps GIS specialist is its own occupation. Otherwise, these workers will be classified in existing occupations.

Differences in education and earnings are sometimes an indication that what appears to be one occupation

is actually more than one. Variations in education and earnings can reveal differences in tasks, acting as a kind of proxy for job duties. Bioinfomatics specialists, for example, are workers who design ways to collect and analyze biological data, usually for biotechnology firms looking for new treatments, genes, and proteins. Their data management tasks might mean that bioinfomatic specialists are database programmers, but their education suggests they might be a unique occupation. Most have advanced degrees in chemistry, biology, or a health profession, and they use that education in their work. Their earnings, too, tend to be higher than those of typical database programmers, suggesting that their work is different.

Usability specialists are another example. These workers design and test websites to make them easy for visitors to navigate. Most have training in cognitive science, psychology, or human factors engineering. This training is one indication that their work is different from that of webpage designers and programmers.

Using education to suggest differences in tasks can be misleading, however. Variations in education level alone might be caused by supply factors or overall educational upgrading in a single occupation. Differences in educational subject matter are more telling. If workers study an entirely different subject, it is more likely that they are in different occupations.

Another indication that a group of jobs might be a distinct occupation is the existence of professional associations and private certifications. These signal that workers have a common professional identity. This is a component of many definitions of occupation, especially those proposed by E.C. Hughes, Eliot Freidson, and other sociologists.

This criterion is almost never included in measurements of new occupations, but professional identity affects how employers advertise for jobs and, thus, how jobseekers will identify occupations.

What makes an occupation new?
After choosing a definition for occupation, researchers must decide how they will categorize an occupation as new, emerging, evolving, or static.

Each study defines these terms differently. In general, however, a new occupation is one that materialized recently. This could be in the last year or in the last decade, depending on the purpose of the study. Often, a new occupation is defined as one that is not listed in the current occupational classification system.

An emerging occupation is a small occupation that is expected to grow large in the future. Most novel occupations cited in the literature fit this category. This is because nearly any occupation that comes to a researcher's attention has already existed for a relatively long time. Massage therapists, for instance, created a professional association in the 1940's. Decades later, massage therapist was identified as emerging, and in 2000 it received its own explicit title in the SOC.

An evolving occupation is an existing occupation whose tasks are changing significantly. This definition leaves vague the meaning of significant. To some extent, all occupations are changing. The evolving occupation category tries to capture the occupations changing most dramatically. An animator moving from two-dimensional pen and paper work to three-dimensional computer modeling is one example of an evolving occupation. Computer programmers who are learning and implementing artificial intelligence is another. Corporate trainers developing desktop and network-based training is a third.

Computers are not the only drivers of occupational evolution. Other frequently-cited changes include robotics, laser, and optics advances affecting electrical engineering tasks and creating electrical engineering specialties and real-time inventory practices affecting warehouse management.

Why do occupations emerge?
There is considerably more consensus about what causes an occupation to change than there is about occupational definitions. Most researchers agree that the driving forces behind new occupations are technology, law, demographic and social trends, and business developments. These factors are the same as those that cause employment to change in existing occupations. To create a new occupation, they must be more dramatic—giving rise to entirely new business needs.

Technological change is the most obvious trigger for new occupations. The development of new communications technology, for instance, creates a need for wireless communications technicians and videoserver technicians to install and maintain communication systems.

The technological changes driving most new occupations today include

- improvements in data management capabilities leading to new programming, records

management, information architecture, security, and other specialties;

- improved graphics, leading to new printing, publishing, design, and animation specialties;

- increasingly sophisticated manufacturing automation and robotics, giving rise to new types of engineering, technician, and material-working, and machine operating jobs;

- computer and communications networking and the Internet, driving demand for new types of programmers, content developers, technicians, and business specialists; and

- medical advances, leading to new technicians and machine repairers.

Occupations also emerge because of changes in the law. Welfare-to-work laws, for example, prompted the creation of new types of job coaches and human service workers. Telecommunications laws that require closed captioning of television programs have spurred growth of closed captioners, or stenocaptioners, a stenography specialty.

Similarly, changes in Medicaid regulations have created demand for new types of record keepers and record makers, including assessment specialists who test the mental and physical functioning of residents in assisted-living institutions and report their findings to government agencies. In addition, stronger environmental regulations during the last 30 years have led to the emergence of environmental monitoring technicians and compliance officers.

Finally, changes in criminal laws have led to occupations such as restitution specialist and victim, witness, and children's advocate.

Demographic shifts and social developments are another source of new occupations. An aging population has created geriatric specialties in medical and social service occupations.

The increase in the women's labor force participation rate, household income, emphasis on nutrition, and prevalence of in-home delivery have led to new contract service occupations such as personal chef and corporate concierge.

New occupations also result from changes in business practices. The increase in health management organizations, for example, drives growth of utilization review coordinators, who examine patient

records to ensure that patient care was in line with an organization's standards. The increase in international trade and international manufacturing standards has led to the demand for workers to ensure that products meet those standards.

Most new occupations result from a combination of these causes. For example, employee wellness managers and work-life experts have developed in response to business trends, social trends, and increases in telecommuting made possible by technology.

Measurements

Futurists often choose new occupations by examining the causes cited above and extrapolating future developments. The occupations they identify—such as space mechanic—might not yet exist. Most researchers, however, are interested in finding extant occupations. Several governments and organizations have conducted studies of new and emerging occupations in the current economy.

Education planner studies

Education planners identify new occupations in order to establish new vocational training programs. Educators are seeking well-established occupations that are only relatively new. In particular, educators usually look for occupations that have materialized in the last 10 years and that meet 2 criteria: there is growing demand for the occupation and specialized training is required to enter the occupation.

The latest study of this type was published in 1999 by the National Council for Occupational Education. To conduct the study, researchers first surveyed community colleges about vocational programs they had added in the past 2 years and programs they planned to add in the next 2 years. The researchers then studied changes in classification systems and lists of new and emerging occupations developed by a variety of sources. Next, they scanned national and local job postings.

Using this information, researchers developed their own inventory of potentially new occupations. They culled this list by checking to see whether other schools already trained students for the occupations and by calling employers to verify demand.

The survey of new vocational programs completed for this study could help identify new occupations. The most commonly reported new programs were in computer networking, computer information systems, and Web design and multimedia. Within each of these areas were training programs for several

occupations—some new and some new only to a particular school.

Less commonly reported programs included bereavement counseling, geriatrics, and teleconferencing. Because of their rarity, these programs might be indications of newer occupations.

Other education planner studies have used projections developed by the Bureau of Labor Statistics (BLS). One definition for an emerging occupation is a small occupation that is expected to become large. Using BLS projections, researchers can find occupations that meet this definition. One could select occupations with fewer than 50,000 workers in 2000 that are also expected to grow twice as fast as the average for all occupations over the 2000-10 projections decade. About 15 occupations meet these criteria. Although some are relatively novel, most are not new. Examples are biomedical engineers, cardiovascular technicians, massage therapists, and desktop publishers.

BLS projections also can be used to identify large occupations that are growing quickly, including some with specialties that could spin off to new occupations. Also, fast-growing residual occupations, including "all other computer specialists", might indicate new occupations that are not yet measured in BLS projections.

These educator and BLS projection studies yield well-established occupations, but the results are not satisfying. Researchers usually hope to identify newer, revolutionary occupations. Occupational surveys are one way to do this.

Occupational surveys
The BLS Occupational Employment Statistics (OES) survey offers some information on new occupations. The survey provides employers a list of occupations common in the employers' industry and asks employers how many of their workers are in those occupations. At the end of the survey, on a supplemental sheet, employers are asked to list and describe occupations in their establishments that are not on the main form. They are asked to give particular attention to occupations that are numerically important or emerging due to technology. Analysts study these forms to identify recurring or important responses.

The OES supplemental forms have two great advantages. First, the large sample generates many responses, and second, the description of job duties

that establishments provide can help analysts determine if a job is in a new occupation.

The supplemental forms cast a wide net, however. Some of the occupations given are new, but many are merely unusual. Electronic commerce specialists are reported, for example, but so are grocery cart washers. Also, many occupations are already classified but not listed on a particular industry's form. One example is aerobics instructor in the hospital industry, an occupation reported on the 1999 supplemental form.

Supplemental forms also can present the opposite problem: too few emerging occupations identified. Respondents may cram their establishments' jobs into occupations listed on the main survey even if the jobs are actually quite different.

Below is a list of occupational titles from the 1993, 1996, and 1999 OES supplements. These titles were selected by OES analysts. The titles in the first group now appear in the new SOC 2000; the those in the second group are either alternative SOC titles or do not appear in the SOC.

SOC 2000 titles
•Aerobics instructor (hospital)
•Administrative assistant
•Customer service representative
•Desktop publisher
•Environmental engineer
•Webmaster

Alternative titles or not in SOC
•Bereavement counselor
•Bus aid
•Credentialer
•Geographic information systems specialist
•Kidney dialysis technician
•Quality assurance director
•Utilization review coordinator
•Volunteer coordinator
•Emissions technician
•International standards organization compliance specialist

The decennial census is another source of new occupational titles, although it does not include job descriptions as the OES survey does.

Between censuses, new alternative titles for census occupations are gathered. Titles are added to the census database at the request of experts or because coders reading and recording census forms bring titles to their managers' attention. Only a few of the

many new titles that occur are added. Still, looking at the titles added for residual occupations could offer clues about occupations not currently classified. And many of these might be new. A few of the titles added between 1990 and 2000 for residual occupations are

- Artificial intelligence specialist
- Information technology (IT) specialist
- Employee wellness coordinator
- Ethics officer
- Human factors engineer

Only two titles were added to the list for "all other computer specialists" between the 1990 and 2000 censuses: IT specialist and artificial intelligence specialist. This suggests that coders from the Bureau of the Census try to fit responses into existing occupations whenever possible.

Surveys specifically designed to identify new occupations might yield better and more complete information than the general surveys like the Census and OES. A few States have conducted this type of specific survey. One of the most recent was Minnesota, which surveyed new and emerging occupations in 1998. In the survey, new occupations were defined as occupations with "work activities, skills, and knowledge that are so new that they cannot be classified under the existing system." Evolving occupations were defined as "established occupations with a rapid change in skill set requiring new knowledge."

The survey was mailed to a sample of the State's employers. Then, analysts checked and compiled responses. Minnesota identified the following new occupations:

New high tech occupations in Minnesota
- GIS specialist
- Interactive specialist
- Internet specialist, administrator, or webmaster
- Management information specialist and assistant
- Network administrator
- Technical information systems manager
- Y2K software analyst

Other new occupations in Minnesota
- Curriculum integration specialist
- Grants specialist
- International standards organization coordinator
- Resident assessment specialist
- Restorative therapy coordinator
- New business venture project manager
- Quality assurance analyst

Most were computer-focused occupations, including 11 titles for Internet specialist, administrator, or webmaster. Interactive specialist is another title for usability specialist described earlier in this paper.

Evolving occupations identified include safety director and warehouse manager.

Literature review and interview studies
More common than surveys were literature reviews and interviews with experts. Texas is undertaking one of the most comprehensive of these studies.

First, Texas analysts identified industries that were likely to have emerging occupations, especially occupations with high wages. To choose industries, they looked at wage data, employment projections, productivity, and labor-to-capital ratios from the U.S. Department of Commerce.

The analysts decided to study information technology, biotechnology, health, and education and training. The information technology study has been completed, as has the first report on biotechnology.

For each industry, analysts have reviewed or will review of trade publications, job postings, and job titles from the State's database of community college graduates. Analysts also have interviewed or will interview major employers.

Some of the results of the IT study follow. In this study, emerging occupations were defined as occupations not in the 1980 SOC. Evolving occupations were in the SOC, but their duties and skills had changed significantly.

Emerging IT occupations in Texas
- Computer network administrator

- Electronic commerce specialist
- Electronic research technician
- Multimedia specialist
- Webmaster
- Direct broadcast satellite services technician
- Global positioning system specialist
- Videoserver technician
- Wireless communications technician

Evolving IT occupations in Texas
- Geographic information systems technician
- Utilization review coordinator
- Warehouse manager

Many of these occupations were identified by the Minnesota study as well.

Another State that researches new and emerging occupations is California. Conducting literature reviews and interviews, California analysts focused on occupations resulting from new technology. The State publishes descriptions of these occupations, together with earnings and employment estimates.

Focus groups
Some researchers ask employers and industry groups to identify new occupations. One example of this type of research is a study of IT jobs by Personnel Decisions Research Institutes, Inc. and commissioned by the U.S. Department of Labor's Employment and Training Administration (ETA). The study included reviews of trade literature and a mass mailing to employers, but the majority of the research occurred during workshops with major employers.

After the study, researchers suggested adding three occupations to the O*NET classification system: quality assurance engineer, computer industry, information technology technical trainer, and database developer.

A similar study was performed in California. A group of multimedia companies met to identify new multimedia occupations. After the occupations were identified, employers were sent an early version of O*NET survey forms for each occupation. The form asked them if they employed a particular occupation and to identify the major skills it required.

Coming data
Some new data on new and emerging occupations might soon be available. Data sources include the following:
- OES summary pages from 2000. These will be the first summary pages after conversion to the new SOC.
- Employment numbers for Census 2000, including the size of residual occupations.
- New Classification of Instructional Programs and new transcript analyses from U.S. Department of Education. These may provide information on new vocational training programs.
- Possible transaction analysis of job postings to America's Job Bank and Monster.com from the ETA. Job postings by employers—and the way they are classified—might be analyzed.

Risks and rewards for jobseekers
New occupations are, by their nature, small, offering few opportunities for jobseekers. It often takes decades for a new occupation to show large employment numbers. And some new occupations never grow. Horticulture therapy aid, for example, was identified as emerging in 1976. Nearly 30 years later, in 2002, it was again listed as new and emerging in popular literature. It is obviously not new and shows no signs of rapid emergence.

Some jobseekers mistakenly assume that a new and emerging occupation will have fast growth, when in reality most studies and surveys offer no projections of future employment.

The specialized nature of some new occupations is another risk. If occupations center on a new technology, they can become obsolete as other workers learn to use the technology and integrate it into their existing occupations. Because of these risks, jobseekers are often encouraged to train for a range of tasks and specialties.

New occupations might also be fleeting for other reasons. If organizations have functioned without an occupation for many years, they may decide they can do without the occupation if budgets tighten.

Still, a jobseeker need only find one open job in an occupation. And new and emerging occupations will continue to be attractive. They are often exciting, offering opportunities to work on the cutting edge. An entrant into a new occupation has the chance to direct the occupation's development, becoming a founding member and veteran. Workers may also be able to take advantages of labor shortages before

other workers are trained. This could lead to high earnings and good opportunities for advancement.

With these advantages in mind, researchers will continue to study and discuss—and voice cautions and caveats about—new occupations.

References

Definitions and causes

Capeli, Peter. "Conceptual Issues in Developing a System for Classifing Occupations," *Standard Occupational Classification Revision Policy Committee, Seminar on Research Findings*. U.S. Department of Labor, 1995.

Drase, M., Dubinsky, O., Flaming, D, and Law, R. "Future Uses of the SOC," *Proceedings of the International Occupational Classification Conference*. U.S. Department of Labor, 1993.

Freidson, E. "The Profession of Medicine: A Study of the Sociology of Applied Knowledge." Dodd and Meade, 1970.

Freidson, E. "Conceiving of Divisions of Labor." *Proceedings of the American Sociological Society Conference*, August, 1980.

Hughes, E.C. *Men and Their Work*. New York: Free Press, 1958.

Levine, Chester; Salmon, Laurie; and Weinberg, Daniel. "Revising the Standard Occupational Classification system." *Monthly Labor Review*, May, 1999. U.S. Department of Labor.

McElroy, Michael P. "The occupational employment statistics program occupational classification," *Proceedings of the International Occupational Classification Conference*. U.S. Department of Labor, 1993.

McGregor, Elizabeth. "Emerging Careers," *Occupational Outlook Quarterly*. Fall 1990. U.S. Department of Labor, 1990.

Tennessee General Assembly. TN Job Skills Terms Defined: Emerging Occupation. Tennessee Department of Economic and Community Development, 1998.

U.S. Department of Labor. "Economic Change and Structures of Classification," *Report on the American Workforce, 1999*. Pp. 101-10.

U.S. Department of Labor. "Emerging Occupations: Not Every Acorn Becomes an Oak," *Occupational Outlook Quarterly*. Fall 1982. U.S. Department of Labor, 1982.

Studies

Adelman, C. *The New College Cource Map and Transcript Files: Changes in Course-Taking and Achievement, 1972-1993*. U.S. Department of Education, 1999.

Anderberg, M and Edwards, D. *Biotechnology: Emerging Occupations in Te*xas. Career Development Resources, Texas Workforce Commission, 2002.

Anderberg, M and Edwards, D. *Emerging and Evolving Occupations in Texas*. Career Development Resources, Texas Workforce Commission, 1999.

Blum, Roman, and Tootle. "The Emergence of an Occupation," *Work and Occupations*, Vol. 15. Sage Publications, February 1988.

Bureau of Labor Statistics. "New and Emerging Occupations," *Issues in Labor Statistics*. U.S. Department of Labor, 2001.

Bureau of Labor Statistics. "New Occupations Emerging Across Industry Lines," *Issues in Labor Statistics*. U.S. Department of Labor, 1998.

California Economic Development Department. *California Occupational Guides, Emerging Occupations*. California Economic Development Department, 1993-2002.

Feingold, S. N. and Miller, N. R. "28 Emerging Careers and Job Opportunities of the Next 10-25 Years," *Forecasts for the Next 25 Years: A Special Report*. World Future Society, 2002.

Greybill, B., Owens, D., Achtelik, L., Hax, Linda, et all. *Results From a Study of Occupations in the Entertainment and Multimedia Industries Using the O*NET Prototype*. State of California Employment Development Department, 1998.

Meeleen, P., Suzanne, G., Baker, K., et all. Identifing and Planning for New and Emerging Occupations: A Suggested Guide. Contract Research Corporation, 1976.

Mishler, Carol. *A Model for Identifying New Occupational Markets for Community and Technical Colleges*. National Council for Occupational Education, October 1999.

O*NET and Information Technology Occupations. Personnel Decisions Research Institutes, Inc. 1999.

What Could Your Next Job Be?: Findings From the 1998 Minnesota Survey of New and Evolving Occupations. Minnesota Department of Economic Security, February 2000.

Slide outline

New and emerging occupations
- Definitions
- Causes of occupational emergence
- Past measurements
- Coming measurements
- Draws and dangers for jobseekers

This page intentionally left blank.

Issues and Strategies on Uncertainty in Population Projections

Chair: W. Ward Kingkade, U.S. Census Bureau, U.S. Department of Commerce

Discussant: Peter Johnson, U.S. Census Bureau, U.S. Department of Commerce

Evaluating Forecast Error in State Population Projections Using Census 2000 Counts

Paul R. Campbell, U.S. Census Bureau, U.S. Department of Commerce

In order to determine if a popular summary statistic–the mean absolute percentage error (MAPE)–is a valid measure of forecast error for the Census Bureau's 1995 to 2000 State population projections, statistical tests were used to determine if the error distribution is strongly influenced by outliers. It was found that the absolute percentage error distribution is skewed and asymmetrical. Since the MAPE understates accuracy, MAPE-R–a variant of MAPE derived from the transformed absolute percentage error distribution–was accepted as more accurate. Using simple extrapolated projections, the findings suggest that the Census Bureau's projections are fairly accurate over a short projection horizon.

Impact of 1990 Census Undercount on the Accuracy of Population Projections

Ching-li Wang, U.S. Census Bureau, U.S. Department of Commerce

This paper examines the accuracy of the latest U.S. Census Bureau's population projections by taking into account the 1990 census undercount. The estimated births, deaths, and migration data from administrative records and the population estimates based on the 1990 census are used to evaluate the projections' performance. The levels of accuracy vary dramatically depending on whether the census counts or the population estimates are used. Results show that the undercount contributed a large proportion of errors in the population estimates, and consequently, the undercount has a considerable impact on the accuracy of the projections. Thus, the uncertainty of the projections is not only the issue of projecting uncertain future, but also the issue of uncertainty in the base year population for projections.

Forecasting Uncertainty in Upcoming Census Bureau Population Projections

Frederick W. Hollmann, U.S. Census Bureau, U.S. Department of Commerce

U.S. Census Bureau population projections have always been viewed by their producers as the working out of a set of deterministic assumptions about fertility, mortality, and migration, imposed on a base population age distribution via the cohort component method. Such projections have generally been defended as illustrative, as opposed to predictive. In the coming year, we are seeking to produce projections that represent forecasts in a more formal sense. With this comes the need to address more systematically the issue of forecast uncertainty. This paper is an introduction to the problems of projecting uncertainty in a highly detailed population matrix. While it does not propose tested solutions to all of the many problems that arise, it clarifies the problems and proposes a method to address them.

Optimization of Population Projections Using Loss Functions When the Base Populations Are Subject to Uncertainty

Charles Coleman, U.S. Census Bureau, U.S. Department of Commerce

Population projections are the product of the base population and a growth ratio. Both of the latter variables are subject to risk and Knightian uncertainty. That is, they can have probability distributions containing both known and unknown components. The expected loss function is used as a criterion for developing "optimal" population projections. Projections can then be made which minimize their expected total loss.

This page intentionally left blank.

EVALUATING FORECAST ERROR IN STATE POPULATION PROJECTIONS USING CENSUS 2000 COUNTS

Paul R. Campbell, U.S. Bureau of the Census

Users of the Census Bureau's 1990 census-based state population projections for 1995 to 2025 are interested in the Bureau's forecast error. Even counties or smaller geographic areas are dependent upon state level accuracy, since projections for these areas often are prorated to the state figures. In this study, Census 2000 counts are used to measure forecast error in projections for April 1, 2000. This is the first opportunity to evaluate the 1995 to 2025 projections with the 'truth' assuming that the Census 2000 results are correct. This evaluation specifically examines the forecast error of the 2000 state population projection totals (including the District of Columbia) for Series A and B, and a simple extrapolated projection to identify the more accurate set of projections and the state outliers. The basic statistical method used to detect state population projection error is to examine the Percentage Error (PE) and the Absolute Percentage Error (APE), while the overall accuracy of the set of projections is measured by the Mean Absolute Percentage Error (MAPE). Additionally, this study examines the skewness and asymmetry of the APE distribution for states to determine if there is a need to use MAPE-R, derived from the transformed APEs, as recommended in the literature in order to correct for the influence of outliers on the mean.

Several questions are addressed in this state projections evaluation. How accurate are the Census Bureau's state population projections for the year 2000? Which Census Bureau state population projection series is the most accurate? Are the Census Bureau's projections as good as or better than the results obtained from simple extrapolated projections? To answer these questions, a discussion on evaluating the Census Bureau's state population projections is presented in several sections. First, the "Prior Research" section reviews recent literature on the evaluation of subnational population projections. The "State Projections for 2000" section discusses the methodology of the Census Bureau's state population projections for 1995 to 2025. The "Extrapolated Projections" section identifies the procedures used to produce a simple extrapolated projection from the enumerated 1990 census counts and the post-1990 census estimates for 1995. Next, the "Calculation of MAPE" section explains how the basic statistical summary measures are derived. The "Transformed APEs and Test for Symmetry" section (1) describes how the APE distribution is examined to see if it is asymmetrical and needs correcting, and (2) presents

the necessary transformation and conversion formula used to correct for the influence of extreme outliers. Finally, the results are discussed in the "Findings and Conclusions" section.

Prior Research. Since literature on the evaluation of state-level population forecasts is not extensive, a broader review of research on the evaluation of subnational population estimates and projections provide useful guidelines for measuring forecast errors. Most frequently subnational population estimates or projections were evaluated using the mean absolute percentage error (MAPE), as indicated in evaluations by Swanson, Tayman, and Barr (2000), Tayman and Swanson (1999), Campbell (1997), Davis (1994), Smith and Sincich (1992). The MAPE is a measure of the central tendency of errors calculated by averaging the sum of the percentage differences between the projections and the census for states, ignoring the plus or minus sign.

There are other alternative statistical measures used to evaluate errors (or differences) in subnational population estimates and projections. Davis (1994), evaluating post-1980 county populations estimates with the 1990 census used several summary measures. In addition to the MAPE, he used the mean algebraic percentage error (MALPE), weighted mean absolute percentage error (WMAPE), root mean square error (RMSE), index of dissimilarity (INDISS), median percent difference, and 90th percentile (or percent error at which 90 percent of the observations are lower). A complete description of these and other statistical measures is available in his study and in Armstrong (1977). Davis (1994) limits his discussion of findings to the "familiar" MAPE, since this summary statistic is highly correlated with the other summary measures for most of his tabulations, except for the MALPE and the 90th percentile.

In questioning the validity of the MAPE measure, Tayman and Swanson (1999) use the MAPE, the Symmetrical MAPE (SMAPE), and a class of measures known as Minimization-Estimators (M-estimators) to evaluate county forecasts for selected states. The M-estimators are described by Tayman and Swanson (1998:303) as "minimizing a more general objective function using maximum likelihood procedures rather than the sum of squared residuals associated with the sample mean, a sum that is highly sensitive to outliers."

Their findings suggest that a robust M-estimator like the Tukey-M statistic is a suitable alternative summary measure of forecast error.

Swanson, et al., (2000) in evaluating subnational estimates, argues that the MAPE is reliable, easy to interpret, and clear in its presentation. On the other hand, they acknowledge that the MAPE in many cases lacks validity, since the APEs used in its calculation are right-skewed, such that extreme values can unduly influence the MAPE. To reduce the effects of outliers and asymmetrical distributions on the arithmetic mean, the alternatives are the median, the geometric mean, the weighted mean, the M-estimator, and data transformation.

In their evaluation of county estimates, Swanson, et al., (2000) and Tayman and Swanson (1999) suggest that some extreme outliers may influence the MAPE by pulling its value upward so that it is not valid for a data set. They suggest validating the MAPE with tests for skewness and symmetry. If the MAPE is not valid, then a variant of the MAPE is calculated using a data transformation that corrects for the error inflation due to the outliers.

This paper follows guidelines found in evaluation literature that recommend performing a data transformation to obtain a variant of the MAPE whenever the original distribution of APE is not symmetrical. In order to test for the effect that outliers have on the summary measure, Swanson, et al., used a modified Box-Cox (1964) transformation to obtain a symmetrical distribution of the original APEs, such that very large errors (or outliers) are compressed. The original distribution can be tested for asymmetry using graphic devices like histograms and boxplots; or statistical measures like the skewness coefficients (Snedecor and Cochran 1980:78); and the D'Agostino skewness test (D'Agostino, et al., 1990). When the original APEs are biased upward, then Swanson, et al., (2000) recommend calculating a summary measure from the transformed APEs that they refer to as MAPE-T. Since the MAPE-T, the average of the transformed APEs, is not represented in a familiar scale, they recommend using a nonlinear power function to statistically map the scales of the original error observations to the transformed observations. Swanson, et al., (2000) suggest calculating a re-expressed average (MAPE-R) in metrics, which is solved using linear regression results and the logarithm of MAPE-T. They argue that MAPE-R is a measure of central tendency of the error that is not influenced by the asymmetry and outliers that characterize the untransformed absolute percentage error distribution. Evaluating county

estimates, Swanson, et.al. (2000:199) validated the MAPE-R statistic by finding consistent results using Tukey-M.

Swanson, et al., (2000) identified some of the shortcomings of the alternative summary measures that may more accurately describe APEs as follows: M-estimators are not easy to explain, the median is "influenced by changes in the centermost observations resulting from grouping," the geometric mean is affected by the logarithmic transformation which sometimes fails to yield a distribution that has optimal symmetry, and the loss function lacks a standard weighting scheme.[1] Additionally, they acknowledge that MAPE-R was cumbersome to calculate and required the use of different statistical software packages.

From a different perspective, economists Kolb and Stekler (1993) recommend using mean square error (MSE) to test whether a set of economic forecasts are statistically significant or better than "naïve" or "no change" forecasts. Using the simple extrapolated population difference as a standard, they used Theil's U statistic to determine if the more complex projection model performs at least as well as the simplest model. Furthermore, evaluating state population projection models that ranged from simple to complex, Smith and Sincich (1992) concluded that the simple trend projections derived from linear extrapolation are just as accurate as the more complex models like the Census Bureau's cohort-component projections. This analysis goes a step further than Smith and Sincich (1992) or Campbell (1997), the MAPE or MAPE-R from the Census Bureau's 2000 state projection and the census population is compared to the same summary statistics from a standard, the simple extrapolated population projection and the census population.

State Projections for 2000. The Census Bureau's state population projections use detailed demographic accounting procedures and professional judgement in developing projection assumptions. The Census Bureau's state projections were prepared for July 1 of each year from 1995 to 2025 using the cohort-component projection method. The cohort-component method is based on the traditional demographic accounting system:

$$P_1 = P_0 + B - D + DIM - DOM + IM - EM$$

where: P_1 = population at the end of the period, P_0 =

---

[1]For a discussion on the merits of the loss function versus MAPE / MAPE-R and other summary measures see Coleman (2000 and 2002).

population at the beginning of the period, and the following events during the period: B = births, D = deaths, DIM = domestic in-migration, DOM = domestic out-migration, (both DIM and DOM are aggregations of the state-to-state migration flows), IM = immigration, and EM = emigration.

Each component of population change -- births, deaths, internal migration (domestic or state-to-state migration flows), and international migration (immigration and emigration) -- utilizes separate projection assumptions for each birth cohort by single year of age, sex, race, and Hispanic origin. The race and Hispanic origin groups projected separately were: non-Hispanic White; non-Hispanic Black; non-Hispanic American Indian, Eskimo, and Aleut; non-Hispanic Asian and Pacific Islander; Hispanic White, Hispanic Black, Hispanic American Indian, Eskimo, and Aleut; Hispanic Asian and Pacific Islander. The detailed components used in the state population projections were derived from vital statistics, administrative records, 1990 census data, state population estimates (U.S. Bureau of the Census, 1996c), and the middle series of the national population projections (Day, 1996). Detailed assumptions and procedures by which these data were generated by single year of age, sex, race, and Hispanic origin are described in detail in the report, "Population Projections for States, by Age, Sex, Race, and Hispanic Origin: 1995 to 2025," (Campbell, 1996). Overall, the assumptions concerning the future levels of fertility, mortality, and international migration are consistent with the assumptions developed for the national population projections (Day, 1996).

Once separate data components were developed, the cohort-component method was applied, producing the detailed demographic projections. For the start of each projection year, the beginning population for each state was disaggregated into race and Hispanic origin categories (the eight groups previously identified) by sex and single year of age (0 to 84, and 85 plus). Components of change were individually applied to each group to project the next year's population. For the mortality component, each age-sex-race/ethnic group was survived forward one year using the pertinent survival rate. The internal redistribution of the population was accomplished by applying the appropriate state-to-state migration rates to the survived population in each state. The projected out-migrations were subtracted from the state of origin and added to the state of destination (as in-migrants). Next, the number of immigrants from abroad was added to each state, while the number of emigrants leaving each state was subtracted. Applying the appropriate age-race/ethnic-specific birth rates to females of childbearing age

created the populations less than one year of age. The number of births by sex and race/ethnicity were survived forward and exposed to the appropriate migration rate to yield the population less than one year of age. The results were adjusted to be consistent with the national population projections by single years of age, sex, and race/ethnicity. Both the state and national population projection reports indicate that 1994 was the base year or the most recent year estimates were used to begin the forecast. However, the first year of the states projection horizon, 1995, was also adjusted to be consistent with a set of preliminary 1995 state estimates only available by age and sex. The entire process was then repeated for each year of the projection.

Two sets of state population projections were prepared and the only component specified differently in each projection model was the domestic migration component. The dynamic possibilities of change in state-to-state migration make it the most difficult component to forecast. Migration trends in the Census Bureau's state projections are based on matched Internal Revenue Service (IRS) individual income tax return data sets containing 19 annual observations (from 1975-76 to 1993-94) on each of 2,550 state-to-state migration flows. The two projection series provide users with different domestic migration scenarios since one set includes the rate of change in employment. Both sets of state projections were summed and adjusted by age, sex, race and Hispanic origin to agree with the middle series of the national population projection.

The Census Bureau refers to Series A, which uses a time series model, as the "preferred series." The first five years of the projection horizon (1995 to 2000) use the time series projection exclusively. The next ten years on the projection horizon (2000 to 2010) are interpolated toward the mean of the series, while the final 15 years (2010 to 2025) use the mean of the series exclusively. Series B is the economics model. State-to-state migration flows are derived from the Bureau of Economic Analysis projected rate of change in employment in the origin and the destination states. The "preferred series" was accepted as the projection model most likely to be the more accurate series based on results from ex-post facto evaluations.[2]

The current sets of state population projections were previously evaluated in the Census Bureau using post-1990 census estimates for July 1, 1996. The MAPEs

---

[2] Ideally, the "preferred series" should be based on the evaluation of preliminary projections where the most recent launch year estimates are withheld from the projection so that estimates can be compared against the preliminary projections and prior evaluation results.

calculated for 1996 were found to be fairly accurate (at or below the U.S. total of 0.40 percent for Series A and 0.30 percent for Series B, for all regions except the West). That earlier evaluation also looked separately at the components of change and found that the birth component was the most accurate followed by the mortality component. The study concluded that both domestic and international migration components were more difficult to forecast accurately, and domestic migration was the least accurate component in the projections (Campbell, 1997).

An important first step in this evaluation was to obtain projections that are consistent with the Census 2000 reference date. While projections are available annually for July 1 of each year, there are none readily available for the target date centered on April 1, 2000. To match the projections to the census date, the solution was to linearly interpolate between the July 1, 1999 and July 1, 2000 state population projection totals to obtain projections for the census date April 1, 2000 using Waring's formula (see Shryock and Siegel, 1976:533).[3]

Extrapolated Projections. A simple extrapolated total population projection for each state was used as a standard to evaluate the forecast error in the state's total population projections. The extrapolated state population projections were derived by linearly extrapolating from the enumerated April 1, 1990 census population and the July 1, 1995 populations[4] to April 1, 2000 for every state. Smith and Sincich (1992), using several techniques in an evaluation which ranged from extrapolating growth rates and ratio shares to time series models, concluded that the linear extrapolation and ratio share models performed the best. Based on their recommendation, the following formula was used to extrapolate to April 1, 2000:

$$P_t = P_0 + X / Y(P_0 - P_b)$$

where the P_t is the state population projection for the target year (April 1, 2000), P_0 is the state population size on July 1, 1995, P_b is the state population size in the base year (April 1, 1990), X is the number of years in the base period (5.25 years between April 1, 1990 to

July 1, 1995) and Y is the number of years in the projection horizon (4.75 years between July 1, 1995 to April 1, 2000). No attempt was made to control the extrapolated state totals to independently derived national projection totals, which results in simple extrapolations for states not affected by inflation/deflation errors (from states being forced to sum prorata to the nation).[5]

Calculation of MAPE. Forecasters tend to treat the terms projection, extrapolation, prediction, and forecast as synonymous (Armstrong, 2001:39). In this study, the initial forecast error refers to the percentage difference or error between a state's total population projection and the Census 2000 population enumerated for the same date.[6] Calculating the percentage error (PE) is useful in examining the error magnitude, direction of error, and identifying outliers in the evaluation of the state projections with census counts (see Table 1). The absolute percentage error (APE) is calculated without regard to the direction of error. The statistical measure that summarizes the APE distribution is the mean absolute percentage error (MAPE). The formula for the APEs and MAPE, ignoring the ± sign, are as follows:

$$APE_i = |P_i - C_i| \div C_i *100$$

$$MAPE = \Sigma(APE_i) \div N$$

where N refers to the number of states (in the U.S., a region, or a division), P is the projected or extrapolated population, C is the census population, and i refers to the state.

MAPEs were developed for the United States (the states and the District of Columbia), where N equals 51; and for each census region or division, where N equals the number of states in each region or division. All data evaluated are from unrounded state population projection figures reported in U.S. Bureau of the Census (2000, 1996a, 1996b, and 1996c) and Campbell (1996).

Transformed APEs and Test for Symmetry. After APE and MAPE are calculated, Swanson, et al., (2000:194) suggest using transformed APEs that are

---

[3] Waring's two-point interpolation formula was used for each state: f(x) = [f(a)*(x-b)/(a-b)]+[f(b)*(x-a)/(b-a)], where f(x) = April 1, 2000 population; f(a) = July 1, 1999 population; f(b) = July 1, 2000 population; and the proportions of the year were x = 2000+92/366; a = 1999+182/365; and b = 2000+183/366. Wang (2002) using geometric interpolation to calculate the April 1, 2000 projections reported slightly different results.

[4] The July 1, 1995 totals rather than July 1, 1994 totals were used since the state projections for the first target year 1995 were inflated/deflated to the preliminary 1995 state estimates.

[5] Comparison of the April 1, 2000 U.S. extrapolated state population totals with the Census 2000 total indicated that the extrapolated projections underprojected the U.S. total population by 2.12 percent. This is more accurate than the Series A and B projections, which were underprojected by 2.62 percent.

[6] Projection evaluations based on estimates are ephemeral, since estimates may be corrected several times during the intercensal decade, only to be finalized after incorporating the latest census results.

Table 1. Percentage Error in Census Bureau's Projections Series A and B, and Extraplated Projections, 2000

Region, Division and State	Census	Projection Series A Number	Percentage Error	Projection Series B Number	Percentage Error	Extrapolated Projections Number	Percentage Error
NORTHEAST							
New England							
Maine.....................................	1,274,923	1,258,277	-1.31	1,249,996	-1.96	1,253,593	-1.67
New Hampshire....................	1,235,786	1,220,918	-1.20	1,213,932	-1.77	1,183,513	-4.23
Vermont...............................	608,827	615,398	1.08	606,141	-0.44	604,642	-0.69
Massachusetts.....................	6,349,097	6,192,883	-2.46	6,216,274	-2.09	6,125,219	-3.53
Rhode Island........................	1,048,319	997,150	-4.88	988,592	-5.70	977,458	-6.76
Connecticut..........................	3,405,565	3,283,684	-3.58	3,285,888	-3.51	3,263,436	-4.17
Middle Atlantic							
New York..............................	18,976,457	18,144,490	-4.38	18,171,587	-4.24	18,267,830	-3.73
New Jersey...........................	8,414,350	8,167,064	-2.94	8,174,604	-2.85	8,139,921	-3.26
Pennsylvania........................	12,281,054	12,196,852	-0.69	12,213,689	-0.55	12,243,940	-0.30
MIDWEST							
East North Central							
Ohio......................................	11,353,140	11,311,801	-0.36	11,343,222	-0.09	11,425,043	0.63
Indiana..................................	6,080,485	6,033,717	-0.77	6,047,746	-0.54	6,038,106	-0.70
Illinois..................................	12,419,293	12,040,250	-3.05	12,057,487	-2.91	12,191,274	-1.84
Michigan...............................	9,938,444	9,673,709	-2.66	9,703,959	-2.36	9,779,185	-1.60
Wisconsin.............................	5,363,675	5,317,128	-0.87	5,314,560	-0.92	5,331,994	-0.59
West North Central							
Minnesota.............................	4,919,479	4,819,724	-2.03	4,812,141	-2.18	4,821,686	-1.99
Iowa......................................	2,926,324	2,897,210	-0.99	2,888,920	-1.28	2,900,595	-0.88
Missouri................................	5,595,211	5,530,291	-1.16	5,535,993	-1.06	5,510,330	-1.52
North Dakota.........................	642,200	660,765	2.89	656,275	2.19	643,632	0.22
South Dakota........................	754,844	774,908	2.66	768,379	1.79	758,890	0.54
Nebraska..............................	1,711,263	1,702,280	-0.52	1,697,153	-0.82	1,690,255	-1.23
Kansas..................................	2,688,418	2,663,372	-0.93	2,669,488	-0.70	2,644,715	-1.63
SOUTH							
South Atlantic							
Delaware...............................	783,600	765,340	-2.33	756,486	-3.46	763,341	-2.59
Maryland...............................	5,296,486	5,264,225	-0.61	5,251,268	-0.85	5,278,573	-0.34
District of Columbia..............	572,059	524,104	-8.38	530,118	-7.33	506,624	-11.44
Virginia.................................	7,078,515	6,979,511	-1.40	6,949,653	-1.82	7,008,299	-0.99
West Virginia........................	1,808,344	1,840,407	1.77	1,832,857	1.36	1,859,488	2.83
North Carolina......................	8,049,313	7,750,347	-3.71	7,761,108	-3.58	7,707,711	-4.24
South Carolina......................	4,012,012	3,849,116	-4.06	3,843,811	-4.19	3,842,137	-4.23
Georgia.................................	8,186,453	7,843,520	-4.19	7,859,752	-3.99	7,854,730	-4.05
Florida..................................	15,982,378	15,181,072	-5.01	15,197,736	-4.91	15,276,275	-4.42
East South Central							
Kentucky...............................	4,041,769	3,988,406	-1.32	3,983,394	-1.44	4,018,483	-0.58
Tennessee............................	5,689,283	5,638,679	-0.89	5,648,707	-0.71	5,598,854	-1.59
Alabama................................	4,447,100	4,441,087	-0.14	4,426,738	-0.46	4,445,124	-0.04
Mississippi............................	2,844,658	2,810,204	-1.21	2,819,891	-0.87	2,809,485	-1.24
West South Central							
Arkansas...............................	2,673,400	2,624,492	-1.83	2,615,811	-2.15	2,604,110	-2.59
Louisiana..............................	4,468,976	4,420,298	-1.09	4,439,675	-0.66	4,453,055	-0.36
Oklahoma..............................	3,450,654	3,367,555	-2.41	3,365,485	-2.47	3,397,212	-1.55
Texas....................................	20,851,820	20,050,906	-3.84	20,106,138	-3.58	20,296,038	-2.67
WEST							
Mountain							
Montana................................	902,195	946,058	4.86	933,518	3.47	934,667	3.60
Idaho....................................	1,293,953	1,338,081	3.41	1,324,245	2.34	1,304,869	0.84
Wyoming...............................	493,782	522,521	5.82	517,023	4.71	504,236	2.12
Colorado...............................	4,301,261	4,149,286	-3.53	4,135,073	-3.86	4,155,703	-3.38
New Mexico...........................	1,819,046	1,852,019	1.81	1,849,867	1.69	1,839,492	1.12
Arizona.................................	5,130,632	4,770,861	-7.01	4,807,423	-6.30	4,717,977	-8.04
Utah......................................	2,233,169	2,194,955	-1.71	2,202,945	-1.35	2,158,202	-3.36
Nevada.................................	1,998,257	1,856,460	-7.10	1,847,723	-7.53	1,827,109	-8.56
Pacific							
Washington...........................	5,894,121	5,838,090	-0.95	5,810,813	-1.41	5,941,423	0.80
Oregon..................................	3,421,399	3,385,241	-1.06	3,384,972	-1.06	3,410,454	-0.32
California..............................	33,871,648	32,462,610	-4.16	32,377,783	-4.41	33,244,071	-1.85
Alaska...................................	626,932	650,925	3.83	630,985	0.65	652,081	4.01
Hawaii...................................	1,211,537	1,253,667	3.48	1,234,930	1.93	1,257,884	3.83
UNITED STATES	281,421,906	274,061,914	-2.62	274,061,954	-2.62	275,462,964	-2.12

Resident Population for April 1, 2000. Series A and B reflect different interstate migration assumptions and do not sum to the same total due to rounding. See text for explanations. Source: Population Division, U.S. Bureau of the Census.

symmetrically distributed to produce an average that reflects more accurately the error represented by most of the observations. The MAPE in most instances is based on a right-skewed, asymmetrical distribution of absolute percentage errors where outliers are likely to pull the summary measure of error upward, thereby overstating the error represented by most of the observations.

Examining data spread. Emerson and Stoto (1983) and Swanson, et al., (2000:194) recommend first looking at the spread of the data to determine if a distribution of APEs

Figure 1. Box-Cox Maximum-Likelihood Values for Series A, B, and Extrapolated Projections, 2000

Source: Population Division, U.S. Bureau of the Census

appears to be unduly dominated by outliers. They suggest applying a transformation when the ratio of the largest value to the smallest value exceeds 20.

Power transformation. Swanson, et al., (2000) have described a modified Box-Cox power transformation procedure to determine the most symmetrical transformed distribution of APEs, which was defined as:

$$Y = (X^\lambda - \lambda) / \lambda, \text{ for } X \neq 0; \text{ or } Y = LN(X), \text{ for } X = 0,$$

where X is the untransformed APE, Y is the transformed APE, and λ (lambda) is the power transformation constant. Lambda is determined by finding its value that maximizes the function:

$$ML(\lambda) = -(N/2)(LN[(1/N)\Sigma(Y_i - \overline{Y})^2] + (\lambda - 1) (\Sigma LN(X_i))),$$

where N is the number of states, Y_i is the transformed APE, $\overline{Y}$ is the mean of the transformed APEs, X is the untransformed APE, and Σ represents the sum over all observations. A "coarse grid" search, set up in a Microsoft Excel spreadsheet, was used to solve for values of λ from –2 to 2 inclusive, using increments of .10.

Figure 1 shows the nonlinear relationship of the Box-Cox maximum-likelihood values associated with λ for each set of projection APEs. The optimal value of λ (0.3) corresponds to the largest maximum-likelihood value (the smallest negative value in the graph) for Series A, Series B, and extrapolated projection APEs.

Test for skewness. The original APEs and transformed APEs can be compared for skewness using graphic devices and the D'Agostino skewness test. Boxplots are graphic devices used to identify location, spread,
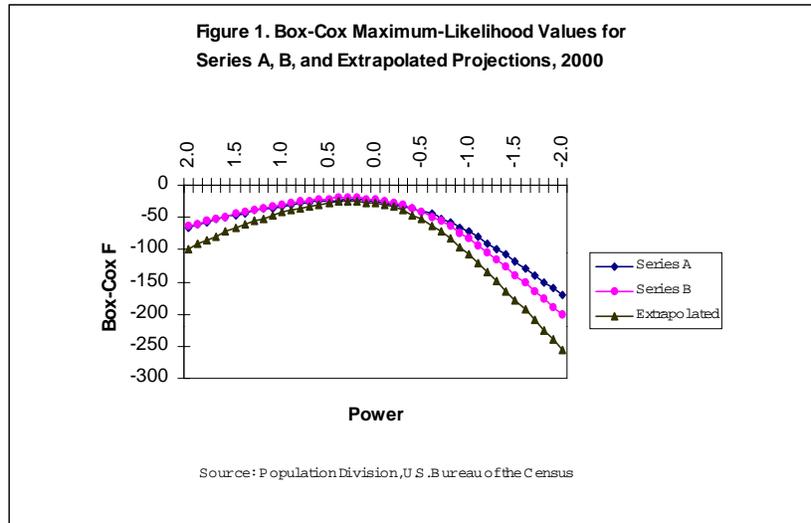
skewness, tail length, and outliers. The spread of the box represents 50 percent of the values between the first and third quartiles (qt). The boxplot for the untransformed APEs indicates a right-skewed distribution whenever the median (the crossbar in the box) is closer to the lower end of the box with a long upper tail. Similarly, the histogram can be used to visually identify asymmetrical and right skewed distributions (see Figures 2 to 4).

Identifying extreme outliers. Swanson, et al., (2000:196) and Emerson and Strenio (1983:59-60) suggest that extreme population outliers for the original APEs should be mathematically identified using information from the boxplot. They suggest calculating extreme outlier cutoff points by multiplying the fourth spread or width of the middle half of the data by 1.5, adding that product to the third quartile value, and subtracting the resulting sum from the first quartile value.

Calculating MAPE-T and MAPE-R. Once it is established that the transformation of APEs was necessary to correct for skewness and asymmetry, the MAPE-T is calculated from the transformed APE distribution using the APE and MAPE formula discussed above. The next step is to calculate MAPE-R, the re-expressed average that matches the original metric distribution, since it is not easy to interpret MAPE-T, the average of the transformed observations, which is in a different unit of measurement. Swanson, et al., (2000:199) recommended using a nonlinear power function to map the scales of the transformed and original observations such as:

$$X_i = A * Y_i^b,$$

where X is the original APE, Y is the transformed APE, and A and B are estimated parameters. The estimated parameters from the linear regression expressed in logarithm form:

$$LN(X) = LN(A) + B \, [LN(Y)]$$

can be used with MAPE-T to estimate:

$$MAPE\text{-}R = e^{\{A + (B * [LN(MAPE\text{-}T)])\}}.$$

The resulting MAPE-R is reported to be a better measure of the central tendency of the error that is not influenced by the asymmetry and outliers that are found in the original absolute percentage error distribution.

Figure 2. Distribution of Absolute Percentage Error, Projection Series A, 2000

Source: U.S. Bureau of the Census

Findings and Conclusions. The findings below suggest the need for a more robust summary measure of forecast error, such as MAPE-R, to evaluate the Census Bureau's state population projections. Additionally, the use of 1) statistical cutoffs to identify extreme outliers, and 2) simple extrapolation as a standard to evaluate state population projections, provides statistical guidelines, rather than subjective conclusions for identifying forecast errors. Data issues associated with comparing the 1990-based projections with the 2000 census are also presented below.

Descriptive analysis. Clearly, the first step in this evaluation was to identify outliers by looking at the percent error for the magnitude and direction of forecast error (see Table 1). There appears to be some overall consistency in the direction of forecast error for states. All three sets of projections underprojected nearly four-fifths of the same states. The few states that were consistently overprojected for all three sets of projections were mostly in the West; i.e., Montana, Idaho, Wyoming, New Mexico, Alaska, and Hawaii,

while the remainder were West Virginia in the South, and South Dakota and North Dakota in the Midwest. For three states, the direction of error was not consistent across all of the projections. Vermont was overprojected only on Series A, while Ohio and Washington were only overprojected on the extrapolated series. The most accurate state projections were those for Alabama on Series A (-0.14 percent) and the extrapolated projections (-0.04 percent), and Ohio on Series B (-0.09 percent).

The range of error was smallest for Series B. The Series B projections ranged from an underprojected population of –7.5 percent for Nevada to an overprojected population of 4.7 percent for Wyoming. In comparison, error in the Series A projections ranged from –8.4

Figure 3. Distribution of Transformed Absolute Percentage Error, Projection Series A, 2000

Source: U.S. Bureau of the Census

percent for the District of Columbia to 5.8 percent for Wyoming. The range of variation for the extrapolated projections was much wider than either Series A or B, ranging from –11.4 percent for the District of Columbia to 4.0 percent for Alaska.

Among the three sets of population projections, two states, Nevada and Arizona, plus the District of Columbia, consistently stand out as extremely low outliers (see Table 1). Both sets of the Census Bureau projections were more accurate than the extrapolated projection for these three outliers.

Clearly, the simple descriptive review so far suggests that Series B appears to be the most accurate. All three sets of projections had trouble accurately forecasting the up and down swings in population growth that occurred in the West during the 1990's. Additionally, the quality of the 1990 census and post census estimates probably contributes to error in the projections; however, these issues were not the focus of the current study.

MAPEs. A comparison of the MAPEs in Table 2 suggests that Series B is the most accurate for the U.S. and the regions. The MAPE for Series B (2.44 percent) is slightly lower than the extrapolated projections (2.54 percent) and Series A (2.63 percent). Furthermore, Series B tends to underproject the actual population (42 of the 51 states which include the District of Columbia were too low; see Table 1). Nearly half (25 out of 51 APEs) were within 2.5 points of the MAPE for Series B.

Forecast error for the regions and divisions varied greatly, but tend to be consistent across the three sets of projections. The MAPEs for regions were highest for the West and lowest for the Midwest. Two-thirds of the division MAPEs were lower for the extrapolated projections than for the Series A and B projections. Due to the smaller number of observations for MAPEs at the region and division level, no attempt was made to validate the region or division results. The next step is to measure the variation in the APEs and determine if they are asymmetrically distributed.

Spread and Asymmetry. Review of the data found a wide range of variation in the APEs which warrants the application of the Emerson-Stoto spread ratio to the original distribution. The transformation of the original distribution of errors was performed since the spread ratio of the highest original APE to the lowest original APE exceeds 20 for each of the three projections. In Series A, the highest APE, 8.38, is for the District of Columbia, while the lowest APE, 0.14, is for Alabama which results in a spread ratio of 60 (8.38/0.14). For Series B, the highest APE, 7.53 in Nevada, and the lowest APE, 0.09 in Ohio, results in a ratio of 84. The widest range occurs in the extrapolated projections where the highest APE, 11.44 for the District of Columbia, and the lowest APE, 0.04, for Alabama, results in spread ratio of 286.

In contrast, transformed APEs for Series A and B, and the extrapolated projections had Emerson-Stoto spread ratios below 20. For example, the log-percentage errors for the transformed APEs (not shown) for Series A ranged from a high of 5.31 percent for the District of Columbia to a low of 0.83 percent for Alabama, which results in a spread ratio of 6. Similarly, the transformed APEs for Series B ranged from 5.11 percent for Nevada to 0.60 percent for Ohio, with a spread ratio of 9. The transformed APEs for the extrapolated projection ranged from a high of 5.92 percent for the District of Columbia to a low of 0.31 for Alabama, with a spread ratio of 19.

After calculating transformed MAPEs (MAPE-T) using the modified Box-Cox method, histograms and boxplots

are created to evaluate the shape of both the original and the transformed distributions. Histograms of the original APEs show data that are asymmetrical and slightly right-skewed (see Series A in Figure 2), while

Table 2. Mean Absolute Percentage Error in State Population Projections, By Region And Division From Series A and B, And Extrapolated Projections, 2000

Region and division	Series A	Series B	Extrapolation
United States	**2.63**	**2.44**	**2.54**
Northeast	**2.50**	**2.57**	**3.15**
New England	2.42	2.58	3.51
Middle Atlantic	2.67	2.55	2.43
Midwest	**1.58**	**1.40**	**1.11**
East North Central	1.54	1.36	1.07
West North Central	1.60	1.43	1.14
South	**2.60**	**2.58**	**2.69**
South Atlantic	3.50	3.50	3.90
East South Central	0.89	0.87	0.86
West South Central	2.29	2.21	1.79
West	**3.75**	**3.13**	**3.22**
Mountain	4.41	3.91	3.88
Pacific	2.69	1.89	2.16

Mean absolute percentage error (MAPEs) are results for 5 years-out from the 1995 population. Based on the enumerated 2000 census counts, the 2000 population for Series A, B, and Extrapolated Projections derived from the Absolute Percentage Errors calculated for the states and the District of Columbia, see text for detailed explanation. Source: U.S. Bureau of the Census.

the histograms for the transformed APEs (TransAPEs) are symmetrical (see Series A in Figure 3). The same patterns were found for Series B and the extrapolated projections, but the histograms are not shown here. The histograms for the three sets of projections validated the need for the data transformations.[7]

Similarly, the boxplots in Figure 4 confirm that original APEs are asymmetrical and right-skewed for all three sets of projections. The median (the crossbar in the box) appears between the middle and bottom of the box, with a long upper tail for the three original APE distributions. The box spread is narrower for the transformed APEs (TransAPE or T-APE in the graph) and symmetrical with the median in the middle of the box (the same location as the mean), with a lower and upper tail of equal length. All of the calculations and graphs were derived using Microsoft Excel, which does not easily facilitate showing the asterisks for extreme outliers in the boxplot graphs.

---

[7] The histogram sorts the APE and TransAPE distributions in a Microsoft Excel spreadsheet using bin range values of 1 to 7. The lowest bin values include APE or TransAPE values less than one, while the highest values were included with 7. A rule of thumb for choosing the APE bin range is one minus the highest value and one plus the lowest value.

The skewness coefficients of 1.02 for Series A, 1.05 for Series B, and 1.74 for the extrapolated projections imply that each set of projections was asymmetrical and right-skewed. A symmetrical distribution would have a skewness coefficient of zero. The D'Agostino skewness test suggests that the null hypothesis (the data are not skewed) should be rejected (p = 0.000) for all three of the original APE distributions.

Extreme outliers. The Emerson-Strenio "fourth spread" procedure was used to identify extreme outliers among the original APE distributions. The "fourth spread" upper cutoff values were 7.80 for Series A, 7.38 for Series B, and 7.93 for the extrapolated projections. The District of Columbia and Nevada were identified as the only extreme outlier in Series A and B, respectively. In the extrapolated projections, three extreme outliers with APEs above the cutoff value were Arizona, the District of Columbia, and Nevada.

MAPE-R results. While the individual transformed APEs are of no interest in the evaluation of the state projections, the summary statistic for the transformed APEs is useful. Additionally, the MAPE-T which equals 3.21, 3.10, and 3.04 for Series A, B, and the extrapolated projections, respectively, is difficult to explain, since the results are log-based. In order to explain the transformed summary statistic in the original metric format, the next step is to re-express MAPE-T as MAPE-R using the logarithm regression results (see formula discussed earlier).[8] The MAPE-R derived for Series B projections at 2.06 percent is slightly more accurate than Series A at 2.24 percent. Additionally, the MAPE-R for the extrapolated projection at 2.00 percent was about the same as the Series B projections and slightly more accurate than the Series A projections.

Table 3 shows MAPE overstating the forecast error in comparison to the MAPE-R. The ratio of the MAPE to

the median (absolute percent error) is another useful descriptive tool that shows the overstated forecast error (Tayman and Swanson 1999:307). In Table 3, the MAPE-to-median ratios confirm that MAPE overstates

FIGURE 4. BOXPLOTS OF APE & TransAPE FOR SERIES A, SERIES B, AND EXTRAPOLATED PROJECTIONS, 2000
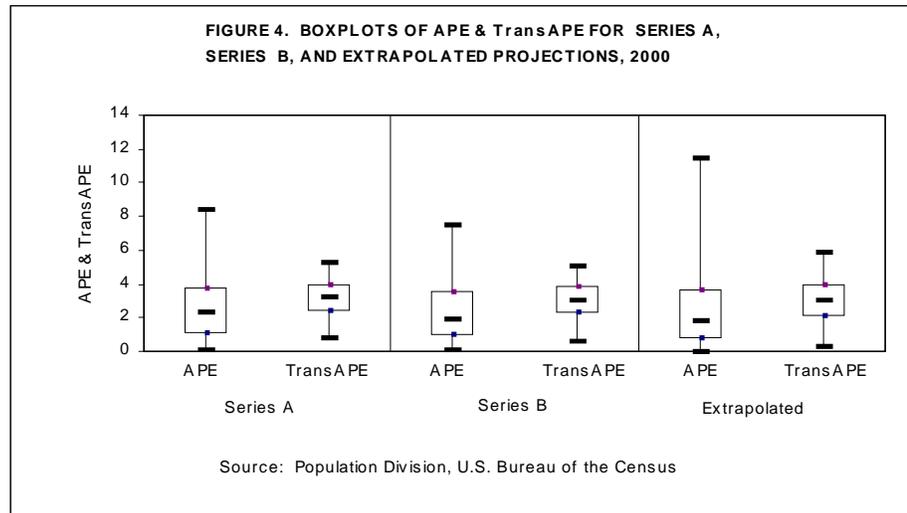
Source: Population Division, U.S. Bureau of the Census

forecast error, since the ratios are greater than 1.0 for all three projections. A different conclusion would have been drawn if the original error distribution were not corrected for skewness and asymmetry.

Initially, Theil's U was considered as a potential summary measure to determine if the Census Bureau's forecast models were more accurate than the extrapolated projections. However, it was not accepted as a valid measure since the distributions of APEs were found to be skewed and asymmetrical. Armstrong and Collopy (1992:77) reported that RMSE (used to derive Theil's U) is unreliable due to its poor protection against outliers.[9] Additional issues related to the guidelines used for choosing appropriate forecast error measures are discussed by Ahlburg (1992).

To summarize, MAPE-R was used to replace the summary measure MAPE in the evaluation of the Census Bureau's projections since the data distributions were skewed and asymmetrical. The results show that the Census Bureau's state population projections for April 2000 (Series B - the economic model) had the least forecast error, with an average absolute percentage error of 2.06 percent. This is slightly better than Series

---

[8] The logarithm regression results used to derive MAPE-R were (1) Series A: $A = -1.950$, $B = 2.366$, $R^2 = 0.992$, standard error (SE) = 0.074; (2) Series B: $A = -1.850$, $B = 2.275$, $R^2 = 0.986$, and SE = 0.104 and (3) the extrapolated projections: $A = -1.649$, $B = 2.104$, $R^2 = 0.968$, and SE = 0.194.

[9] Theil's U, interpreted as the RMSE of the projections model divided by the RMSE of the extrapolated or no-change model, is derived from the formula: $U = [\Sigma(P_i - E_i)^2]^{1/2} / [\Sigma E_i^2]^{1/2}$, where P_i refers to the projection error for each state, and E_i is the corresponding extrapolation error used as the standard for each state (see Kolb and Stekler, 1993). Using either absolute or percent change or the transformed percent change, Theil's U coefficients found both Series A and B with fewer errors than the extrapolated projections.

A, with an average absolute percentage error of 2.24 percent. The forecast error in the Census Bureau's Series B projections was the same as that found in the extrapolated projections (2.00 percent), while the extrapolated projections slightly out-performed Series A - the preferred series. All three projections consistently underprojected approximately four-fifths of the states (40 states in the extrapolated projections, 41 states in Series A, and 42 states in Series B) out of a total of 51 states (including the District of Columbia). The widest range of variation and the most extreme outliers were found for the extrapolated projections.

Table 3. Comparison of the MAPE, Ratio to Median, and MAPE-R for Series A, Series B, and Extrapolated Projections, 2000.

	MAPE	Ratio of MAPE to Median	MAPE-R
Series A	2.63	1.13	2.24
Series B	2.44	1.24	2.06
Extrapolated	2.54	1.38	2.00

Summary Statistics for projections evaluation using enumerated Census 2000 results, see text for detailed explanation.
Source: Population Division, U.S. Bureau of Census.

An added feature of the extrapolated projection is that base period (1990-1995) growth trends are held constant over the projection horizon (1995-2000). This information is useful for identifying changes in trends (or error) between the base period and the projection horizon. Ideally, when the extrapolated projection error is zero, there is no evidence of change in the pattern of growth between the base period and the projection horizon. In this study, nearly a third of the states (16 states with error ranging from 1.0 percent to -1.0 percent) showed little change in the 1990-95 pattern of population growth extrapolated to 2000.

Several issues or differences between the 1990 and 2000 censuses not examined in this study probably affect the accuracy of the state projections. First, adjusted 1990 census counts were not used as the base year and any undercoverage in the 1990 census is carried throughout the post-1990 estimates and projections.[10] Second, this evaluation only examines the aggregated population totals and does not evaluate the separate component totals, such as births, deaths, state-to-state, and international migration, by age, sex, and race/Hispanic

origin. The domestic migration and international migration components are the most difficult to adequately baseline or project. Additionally, retrospective census information on place of residence during 1985-90 used in the projections may not reflect changes in the age pattern of migrants during the 1990's. Third, the race/Hispanic origin categories are quite differently defined in each of the censuses, the vital statistics, and administrative records. Fourth, the state projections use national data as a proxy in the absence of detailed demographic components. Mulder (2001) evaluating the Census Bureau's national population projections produced between 1947 and 1994 has documented the inability of past projections to accurately forecast turning points, particularly for the immigration and fertility components of the projections. Finally, there is the issue of the multi-dimensional raking, in other words the state projection results are aggregated pro-rata to the national estimates and projections for consistency at the national level by age, sex, and race/Hispanic origin.

The 2000 state population projections appear to be slightly more accurate than vintage projections produced decades earlier. Wetrogan and Campbell (1990) calculated MAPEs ranging from 3.0 percent to 5.2 percent for a 5-year projection horizon in their evaluations of 1970's and 1980's Census Bureau projections using corresponding 1970's and post-1980 census estimates.[11] They reported U.S. MAPEs from the Census Bureau's 1987 projection at 0.5 percent, 1.1 percent, and 1.6 percent for one-, two-, and three-year projection horizons, respectively. MAPEs of 0.5 percent per year appear to be a reasonable level of accuracy to expect for state population projections over a short term or 5-year projection horizon.

This study found that the Census Bureau's 2000 state population projections are as accurate as simple extrapolated projections and have fewer extreme outliers. Further evaluation of the detailed demographic components should aid in identifying areas of the projection model that needs to be improved. It appears that tests for skewness and asymmetry are necessary to validate the use of the popular summary measure, such as the MAPE or its variant MAPE-R.

The advantage of using MAPE-R in conjunction with the original absolute percentage error is that users are more familiar with interpreting this summary measure and MAPE-R resolves the central tendency issues

---

[10] An evaluation of the factors affecting the accuracy of the state projections, such as census undercount, estimates error, and error in the projected components of change have been addressed by Wang (2002).

[11] An evaluation of the 1970's and 1980's state population projections using the 1990 post census estimates, final intercensal estimates, and MAPE-R would probably yield lower forecast errors.

whenever MAPE is found to be invalid. Clearly, a drawback to its widespread use is the cumbersome statistical calculation needed to carry out its application; nevertheless, all of the results for this evaluation were carried out in Microsoft Excel spreadsheets.[12] With a few modifications, the spreadsheets can be used to evaluate error in other small subnational estimates or projection data sets.

Acknowledgments and Disclaimer. I would like to thank Greg Spencer and Edwin Byerly for inspiring this research, and Ching-Li Wang for his comments and sharing evaluation results.

This paper reports the results of research and analysis undertaken by Census Bureau staff. It has undergone a more limited review than official Census Bureau Publications.   This report is released to inform interested parties of research and to encourage discussion.

References
Ahlburg, D.A. 1992. "A Commentary on Error Measures: Error Measures and the Choice of a Forecast Method." *International Journal of Forecasting* 8:99-11.

Armstrong, J.S. 2001. "The Forecasting Dictionary" in *Principles of Forecasting: A Handbook for Researchers and Practitioners*. Norwell, M.A., Kluwer Academic Publishers.

Armstrong, J.S. and F. Collopy. 1992. "Error Measures for Generalizing about Forecasting Methods: Empirical Comparisons." *International Journal of Forecasting* 8:69-80.

Armstrong, J.S. 1977. *Long-range forecasting, from Crystal Ball to Computer*. New York: John Wiley & Sons.

Campbell, P.R. 1996. "Population Projections for States, by Age, Sex, Race, and Hispanic Origin: 1995 to 2025." PPL-47. U.S. Bureau of the Census, Population Division. Data files: PE-45 on Internet http://www.census.gov/population/www/ projections/stproj.html.

Campbell, P.R. 1997. "An Evaluation of the Census Bureau's 1995 to 2025 State Population Projections – One Year Later." Presented at the Population Association of America Meeting: Washington, DC.

Coleman, C.D. 2000. "Evaluation and Optimization of Population Projections Using Loss Function." in Papers and Proceedings of the 11th Federal Forecasters Conference - 2000. edited by D. E. Gerald. U.S. Department of Education, Office of Educational Research and Improvement. Washington, D.C.

Coleman, C.D. 2002. "Measures of Estimates Quality." Presented at the Population Association of America Meeting. Atlanta, GA.

Davis, S.T. 1994. "Evaluation of Postcensal County Estimates for the 1980s." Working Paper 5. U.S. Bureau of the Census, Population Division.

Day, J.C. 1996. *Population Projections of the United States by Age, Sex, Race, and Hispanic Origin: 1995 to 2050*. Current Population Reports. Series P25-1130. U.S. Bureau of the Census. Washington, D.C. Government Printing Office.

Emerson, J. and M. Stoto. 1983. "Transforming Data." Pp. 97-128 in *Understanding Robust and Exploratory Data Analysis*, edited by D. Hoaglin, F. Mosteller, and J. Tukey. New York: Wiley.

Emerson, J. and J. Strenio. 1983. "Boxplots and Batch Comparisons." Pp. 58-96 in *Understanding Robust and Exploratory Data Analysis*, edited by D. Hoaglin, F. Mosteller, and J. Tukey. New York: Wiley.

Kolbs, R.A. and H.O. Stekler. 1993. "Are Economic Forecasts Significantly Better than Naïve Predictions? An Appropriate Test." *International Journal of Forecasting* 9:117-120.

Mulder, T. 2000. "Accuracy of the U.S. Census Bureau National Population Projections and Their Respective Components of Change." In Papers and Proceedings of the 11th Federal Forecasters Conference - 2000. edited by D. E Gerald. U.S. Department of Education, Office of Educational Research and Improvement. Washington, D.C.

Shryock, H.S. and J.S. Siegel. 1976. *The Methods and Materials of Demography*. Condensed Edition. San Diego, California: Academic Press.

Smith, S.K. and T. Sincich. 1992. "Forecasting State and Household Populations, Evaluating the Forecast Accuracy and Bias of Alternative Population Projections for States*." International Journal of Forecasting* 8: 495-508.

Swanson, D.A., J. Tayman, and C.F. Barr. 2000. "A Note on the Measurement of Accuracy for the Subnational Demographic Estimates." *Demography* 37:193-201.

Tayman, J. and D.A. Swanson. 1999. "On the Validity of MAPE as a Measure of Population Forecast Accuracy." *Population Research and Policy Review* 18:299-322.

U.S. Bureau of the Census. 2000. "Table 2. Resident Population of the 50 States, the District of Columbia, and Puerto Rico: Census 2000." December 28. Internet site: http://blue.census.gov/ population/cen2000/tab02.pdf.

U.S. Bureau of the Census. 1996a. "Estimates of Population and Demographic Components of Change for States: Annual Time Series, 1990-96." ST-96-1. Population Distribution Branch.

---

[12] A copy of the Microsoft Excel spreadsheet used to evaluate the state projections can be obtained from the author by e-mail addressed to Paul.R.Campbell@census.gov.

U.S. Bureau of the Census. 1996b. "Population of States by Single Years of Age and Sex: 1990 to 1995." PE-38. Population Division.

U.S. Bureau of the Census. 1996c. "Estimates of the Population of States by Age, Sex, Race, and Hispanic Origin: 1990 to 1994." PE-47. Population Division.

Wetrogan, S.I. and P.R. Campbell. 1990. "Evaluation of State Population Projections." Presented at the Population Association of America: Toronto, Canada.

Wang, C. 2002. "Evaluation of Census Bureau's 1995-2025 State Population Projections" Population Projections Branch. U.S. Bureau of the Census, Population Division. Forthcoming.

Impact of 1990 Census Undercount on the Accuracy of Population Projections

Ching-li Wang, U.S. Census Bureau

I. Introduction

One of the major concerns in preparing population projections is the accuracy of the projections. When the new census results become available, people want to know how close the projections are to the census counts. The availability of the 2000 census counts prompts us to examine the accuracy of the latest U.S. Census Bureau's projections.

To evaluate the accuracy of the projections, we need a "true" population to compare. Most studies compare the projections with the census count for the census year or with most recent population estimates available for the inter-censal or post-censal years (Smith and Sincich, 1990, 1992; Wetrogan and Campbell, 1990; Campbell, 1996a, 1997). However, the precise "true" population may not exist because of census undercount and errors in population estimates. According to the Accuracy and Coverage Evaluation (ACE) survey and the Demographic Analysis (DA) conducted by the Census Bureau, the net undercount rates in the 2000 census are significantly lower than in the 1990 census (Robinson, 2001a, 2001b). Since the latest population projections were based on the 1990 census-base population estimates, we would expect that the projected 2000 population based on the 1990 census would understate the 2000 population considerably as compared with the 2000 census counts due to a lower beginning population. Changes in net undercount between the two censuses affect the validity of measuring the accuracy of the projections. The accuracy of the census affects not only the accuracy of base year population for population estimates and projections but also the validity of measuring their accuracy later.

The purpose of this paper is to evaluate the latest Census Bureau's projections for the nation, 50 states and District of Columbia to demonstrate the impact of the 1990 census undercount on the accuracy of the projections. The paper also examines errors in the population estimates and projected components of change - births, deaths, and migration to determine the importance of 1990 census undercount affecting the accuracy of population projections.

The results show that the levels of projection accuracy vary dramatically depending on whether the census counts or the population estimates are used to evaluate the accuracy. The majority of the errors in the latest national projections are due to the 1990 census undercount. At the state level, when the percent errors of projected components of change (births, deaths, domestic migration, and international migration) are held constant in a regression analysis, the percent errors in population estimates and the 1990 Census undercount account for most of the errors in the state population projections. The 1990 census undercount also contributed a large proportion of errors in the state population estimates. Since the population estimates were used as the base population for developing the latest projections, the 1990 census undercount has a considerable impact on the accuracy of the projections through its impact on the estimates. Thus, the uncertainty of the projections is not only the issue of projecting uncertain future, but also the issue of uncertainty in the beginning population itself.

II. Methodology of Census Bureau's Population Projections

The Census Bureau's latest population projections for the United States from 1999 to 2100 were released in January, 2000 (Population Division Working Paper No. 38). The latest state population projections from 1995 to 2025 were released in October 1996 (Population Division PPL-47). Both national and state population projections use the cohort component method to prepare the projections. The components of population change - births, deaths, and migration are projected separately. It requires separate projection assumptions for each birth cohort by single year of age, sex, race and Hispanic Origin. The race and Hispanic origin groups were non-Hispanic White, non-Hispanic Black; non-Hispanic American Indian, Eskimo, and Aleut; non-Hispanic Asian and Pacific Islander; Hispanic White, Hispanic Black, Hispanic American Indian, Eskimo, and Aleut; and Hispanic Asian and Pacific Islander. For the national projections, the foreign-born population was also projected separately.

The national projections were launched from an estimated

resident population by age, sex, race, Hispanic origin, and nativity as of January 1, 1999, which are based on the 1990 census. The component method for the nation is expressed by the following formula:

$$P_1 = P_0 + B - D + M$$

Where,

P_1 = Population at the end of the period

P_0 = Population at the beginning of the period
B = Births during the period
D = Deaths during the period
M = International Migration: legal immigration, refugee movements, emigration (of natives and foreign-born combined), net migration from Puerto Rico, and net undocumented migration.

Three series of projections were produced as middle, lowest, and highest based on the variant assumptions of "extreme" lowest and highest values of the three major components (Hollmann, Mulder, and Kallan, 1999). In this paper, we use the middle series to examine the impact of census undercount on the national projections.

The state population projections were launched from the estimates of state population as of July 1, 1994, which are also based on the 1990 census population. The first projected 1995 results were later adjusted to agree with the 1995 state population estimates when they were available. The final results were consequently controlled to agree with the middle series of the national projections (P25-1130, 1996). The cohort component method used to prepare the state population projections is based on the following formula:

$$P_1 = P_0 + B - D + DIM - DOM + IIM - IOM$$

Where,

P_1 = Population at the end of the period
P_0 = Population at the beginning of the period
B = Births during the period
D = Deaths during the period
DIM = Domestic in-migration during the period
DOM = Domestic out-migration during the Period
IIM = International in-migration during the Period
IOM = International out-migration during the Period

Two sets of state population projections were prepared based on different models used in projecting the domestic migration component. The migration trends data used in both projections were based on state migration flows data,

extracted from Internal Revenue Service(IRS) individual income tax returns. The data contain 19 annual observations from 1975-76 to 1993-94 on each of the 2,550 state migration flows (51 x 50 matrix). Two models were used to project these migration flows into the future:

(1) Series A, as preferred series, used a time series model - regression of changes in the natural logarithms of the migration rates. The first five years of the projections used the time series projections exclusively. The next ten years of projections were interpolated from the time series projections toward the mean of the series. The final 15 years used the series mean exclusively.

(2) Series B is an economic model. Changes in state-to-state migration rates were derived from the relationship between changes in the migration rates and Bureau of Economic Analysis projected changes in employment in the origin and the destination states. Detailed assumptions and procedures used in the projections are described in the Census Bureau's report, PPL-47 (Campbell, 1996b).

For the analysis in this paper, we use only the preferred series A to examine the impact of census undercount on the accuracy of projections. The state projections series were prepared as of July 1, each year. Thus, the projected April 1, 2000 populations are derived from the geometric interpolation of projected populations between 1999 and 2000 in order to compare with the 2000 census counts as of April 1.

III. Assessment of the Accuracy of Population Projections

Based on the methodology we just described, several factors need to be considered in order to evaluate the accuracy of the projections, such as the accuracy of based year population, the accuracy of projected components, and methodological appropriateness. For this paper, we will focus on the evaluation of the performance of the projections, not the methodological procedures used in producing the projections.

1. Census undercount

The Census Bureau has used two approaches to measuring the undercount. One method uses birth and death records, and immigration records to estimate the true population. This estimate is compared to the census count to measure the difference. This method is called Demographic Analysis.

Another method is to conduct special surveys to measure the undercount. A scientific sample of census blocks are re-interviewed independently of the census enumeration. The results of these interviews are checked against the census records on an individual basis to see who was missed and who was counted in error. The survey used in the 1990 census is called the Post-Enumeration Survey (PES). The survey used in the census 2000 is called the Accuracy and Coverage Evaluation (ACE) survey. (For description of the surveys and Demographic Analysis, see Hogan and Robinson, 1993, and Robinson, 2001a, 2001b)

The net undercount rates for the nation based on Demographic Analysis has decreased over the past few decades (5.4% in 1940, 4.1 % in 1950, 3.1 % in 1960, 2.7 % in 1970, 1.2 % in 1980) until 1990 when the undercount rate increased again to 1.8 percent. (Hogan and Robinson, 1993). However, the net undercount rate decreased dramatically to 0.12 percent in the 2000 census (Robinson, 2001b). The net census undercount decreased from more than 4 million in 1990 to only 0.3 million in 2000. Such a big difference in the net census undercount between 1990 and 2000 would have a tremendous effect on the evaluation of population estimates and projections. For the analysis in this paper, we use the 1990 PES estimates of net undercount by state available from the Census Bureau's (http://www.census.gov/dmd/www/pdf/understate.pdf.)

2. Accuracy of population estimates

The 1999-2100 national projections are based on January 1, 1999 national population estimates. The 1995-2025 state projections are based on July 1, 1994 state population estimates as the first base year population and then are adjusted to agree with the 1995 state population estimates for the first projection year. The accuracy of the population estimates definitely affects the base year population for projections. To assess the accuracy of the projections, we need to examine the accuracy of the estimates against the 2000 census population. The 2000 vintage population estimates based on the 1990 census are used for analysis.

3. Accuracy of projected components of change

Since the projections are derived from the demographic accounting of births, deaths, and migration, the quality of input data and methodologies for deriving projection assumptions for each component will definitely affect the accuracy of the projections. To assess the accuracy of the projected components of change, we use the most recent available statistics for 1999 to 2000 to evaluate the national

projections, and use the available statistics by state between 1995 to 2000 to evaluate the accuracy of the components for state projections.

4. Impact of national projections on state projections

The results of state population projections were controlled to agree with the most recent national population projections as the final stage of procedures. The accuracy of the national projections will eventually affect the accuracy of the state projections. For example, the national projections, to which the current series state projections were controlled, showed 274.7 millions people in 2000 while the 2000 census showed 281.4 million. A difference of 6.8 million between projected national population and the census count will definitely affect the accuracy of the state projections when the state projections are controlled to agree with the national projections.

5. Uncertainty of demographic changes

Most projections are based on the assumption that population change can be predicted if the current or historical demographic trends continue in the future. However, it is not always the case. Therefore, we can anticipate that the projections for the areas which experience dramatic socioeconomic changes will not be as accurate as the areas with stable socioeconomic conditions. The population change between 1990 and 2000 can be used to measure whether the states have experienced dramatic changes or not.

6. Measurement of accuracy

The most commonly used measurement of accuracy of the projections is the percent difference or absolute percent difference between the projected population and "true" population for a geographic area. When measuring the magnitude of errors among a specific number of geographic areas (such as 50 states or the number of states in each region or division), Mean Absolute Percent Error (MAPE), which is the average error when the direction of error (positive or negative) is not used. However, when the direction of errors is taken into account, the Mean Algebraic Percent Error (MALPE) has been used as a measure of forecast bias, whether under-projected or over-projected (Smith and Sincich, 1990, 1992).

It has been argued that the MAPE overstates the error of projections or estimates because a few extreme outliers would make the average (arithmetic mean) higher than reality (Tayman and Swanson, 1999; Tayman, Swanson,

and Barr 1999, Swanson, Tayman, and Barr, 2000). Therefore, the use of the measurement of accuracy also affect the levels of accuracy. However, in order to compare the results with previous studies using the MAPEs, and cross-comparison of errors in different variables, we will rely on the Percent Difference and MAPE to discuss the accuracy of the projections.

IV. Census Undercount and National Projections

The latest national projections to year 2100 released in January, 2000 show that the projected population of 274,659,000 under-projected by 6.8 million as compared with the census 2000 of 281,421,906 as shown in Table 1. It is a 2.4 percent error for the first two years. However, if the projections are compared with the national estimates, the projections over-projected the population by only 42,000, a very small marginal error of 0.02 percent. Thus, it is very clear that the accuracy of the national projections depends on the use of census counts or population estimates.

Table 1 also shows that if the 1990 census undercount rates are applied to the projected 2000 population, the under-projection of the U.S. population are reduced dramatically from 6.8 million to 2.4 million using the 1990 PES (Post-Enumeration Survey) undercount rate, and to 2.2 million using the DA (Demographic Analysis) undercount rate. The percent errors for the projections are reduced from 2.4 percent to 0.9 percent with PES undercount rate adjustment and to 0.8 percent with DA undercount rate adjustment. This suggests that if the projections were based on the undercount adjusted 1990 census population, the Census Bureau's latest U.S. projections would be very accurate.

As Figure 1 shows, the amount of reduction in difference between projections and the census count from 6.8 million to 2.2 or 2.4 million with adjustment of the 1990 census undercount accounts for 64.1 percent to 66.9 percent of the projection error. In other

words, the 1990 census undercount contributes about 2/3 of the discrepancy between the projections and the census count. The remaining 1/3 of the discrepancy can be attributed to other source of errors such as projected births, deaths, and migration. Therefore, the 1990 census undercount is far more important than other source of errors in explaining the errors in the national population projections.

In terms of other source of errors, as Table 2 shows, the projections under-projected the number of births by 214,000 (2.7%) for 1999-2000 based on the provisional NCHS report. The projections under-projected the number of deaths by 31,000 (0.6%) for the same period. If the projections were based on the 1990 undercount adjusted population, the projected births and deaths should have increased to some extent due to a larger population base.

Table 1: U.S. Population Projections, Census 2000 Count, and Vintage 2000 Estimates: 4/1/2000

Projections/Census/Estimates	Population	Projections Minus Census or Estimates		Difference Reduced After Undercount Adjustment	
		Number	Percent	Number	Percent
Census 2000	281,421,906	--	--	--	--
Projections	274,649,908	-6,771,998	-2.41	--	--
Adjusted projections based on					
PES undercount rate**	278,989,377	-2,432,529	-0.86	-4,339,469	64.08
DA undercount rate**	279,181,631	-2,240,275	-0.80	-4,531,723	66.92
Vintage 2000 Estimates	274,608,346	41,562	0.02	--	--

* Population Projections of the United States: 1999 to 2100 (Population Division Working Paper No. 38)
**The adjustment for 1990 census undercount is based on the following information.

1990 Official U.S. Population	248,709,873
Undercount Adjusted (PES)	252,730,369
PES undercount rate	1.58
DA undercount rate	1.65

Source: U.S. Census Bureau, ESCAP II: Demographic Analysis Results, October 13, 2001.
and http://www.census.gov/dmd/www/pdf/understate.pdf.

Figure 1
Difference Between Projected U.S. 2000 Population and Census Count
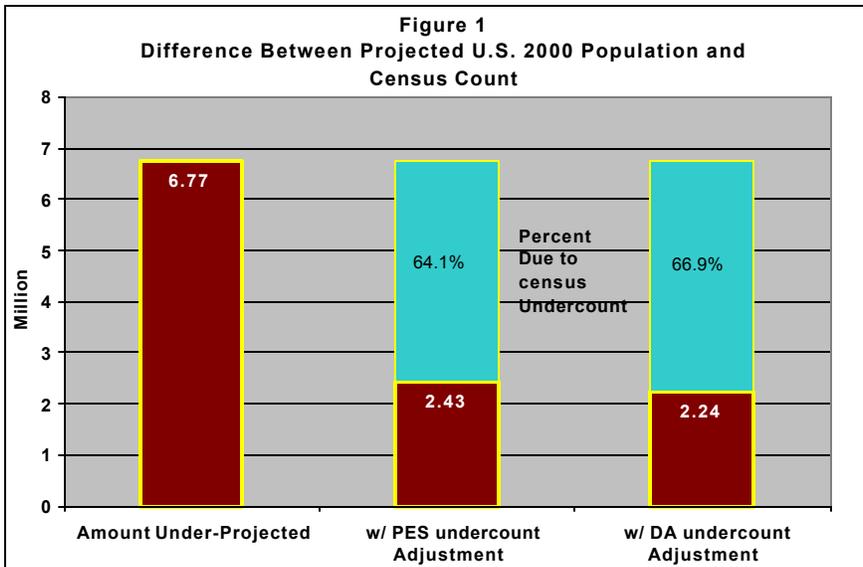
The percent errors of projected births and deaths should

also be reduced. Therefore, we can conclude that the projected births and deaths for the first two years are quite accurate.

Table 2 also shows that the projections of net international migration in 1999 and 2000 are higher than the estimated figures by 10.6 percent. However, the projections of the international migration are conditioned by the estimates of

Table 2: Projected and Estimated Components of Change of the U.S. Population: 1999 - 2000

Components of Change	Projections 1999-2000	Estimated ** 1999-2000	Projections - Estimates Number	Percent
Births	7,813,864	8,028,000	-214,136	-2.67
Deaths	4,769,367	4,800,000	-30,633	-0.64
Net International Migration	1,930,583	1,744,963	185,620	10.64

Source: U.S. Census Bureau, U.S. Population Projections 1999 - 2100;
 ** NCHS, National Vital Statistics Report, Vol. 49, No.6, August 22, 2001.
 **Net International migration is derived from the annual estimates in the Census Bureau's
   Population Estimates Program

international migration as the base for developing assumptions for the migration component. Since the projected births and deaths are relatively accurate, the remaining 1/3 of discrepancy between the national projections and census 2000 is due mostly to the under-projected international migration, caused by the under-estimated international migration. Even if the remaining 1/3 of discrepancy between the projections and census 2000 is assumed totally due to the international migration, the 1990 census undercount is still far more important in explaining the errors in projections.

VI. Accuracy of the State Population Projections

The census undercount has the similar effect on the accuracy of the state projections. The level of accuracy also depends on whether the population estimates or the census counts are used to evaluate the accuracy. First of all, let us look at the overall performance of the state projections for 50 states and the District of Columbia.

1. Comparison with Census 2000

As shown in Table 3, for the first five years, the state projections produced a mean absolute percentage error (MAPE) of 2.6. The Mean Algebraic Percent Error (MALPE) was -1.4 percent. This indicates a general tendency for the projections to under-project the state populations as expected due to higher undercount rates in the 1990 census.

The MAPEs vary from region to region. The projections

are more accurate in the Midwest and less accurate in the West. The MAPEs for the West vary dramatically from state to state. Generally, the projections are less accurate in Mountain states and South Atlantic states with a wide range of levels of accuracy.

Despite the errors we just described, the current set of projections tends to be more accurate than in the earlier projections produced before the 90s. According to Smith and Sincich (1992), the MAPEs for the Census Bureau's state projections after 5 years ranged from 3.1 to 5.0 percent for earlier versions of the projections (1955 through 1980). Wetrogan and Campbell (1990) analyzed the Census Bureau's previous series of state projections from 1965 (P25-375) to 1980 (P25-937) and found the MAPEs for the first five years of projections ranged from 5.2 to 3.0 percent. The MAPEs for the 1986 Series (P25-1017), 1988 Series (P25-1053) and 1992 Series (P25-1111) are calculated to compare with the current series. The overall accuracy of the state population projections has improved since the 1986 Series (P25-1017) with an MAPE of 2.6. The first projections series after 1990 (P25-1111) was even more impressive with an MAPE of 1.6 for the first 5 years. Then, the MAPE for the latest series PPL-47 returned to the same level of 2.6 as previous two series in the late 80s. (See Wang, 2002).

2. Comparison with population estimates

However, if the 2000 population estimates, which are based on the 1990 census, are used to evaluate the accuracy of the state projections, the MAPEs are generally lower than when the 2000 census was used. As Table 3 shows, the MAPE was reduced to 1.73 percent if the population estimates are used to measure the accuracy. The reduction of projections errors can be seen for all regions, especially in the Northeast and South.

3. Undercount Adjusted Projections

As mentioned above, the 2000 census had a higher coverage rate than the 1990 census. The projections based on the 1990 census will certainly tend to under-project the population. If we use the 1990 census undercount rates to adjust the state projections, we should see a reduction of percentage errors.

As Table 3 shows, the MAPE for all states was reduced from 2.6 to 2.2. The MAPEs for all regions were reduced

Table 3: Mean Absolute Percentage Error (MAPE) and Mean Algebraic Percentage Error (MALPE) of of State Population Projections Using Census 2000 and Population Estimates for Calculation and Undercount Adjusted Projections

Region and Subdivision	Number of States	Compared with Census 2000		Compared with Population Estimates		Undercount Adjusted Projections	
		MAPE	MALPE	MAPE	MALPE	MAPE	MALPE
U.S. Total	51	2.64	-1.40	1.73	1.19	2.21	0.13
Northeast	9	2.50	-2.26	0.84	0.36	2.06	-1.57
New England	6	2.42	-2.06	0.82	0.55	2.14	-1.40
Middle Atlantic	3	2.67	-2.67	0.89	-0.03	1.90	-1.90
Midwest	12	1.58	-0.65	1.47	0.80	1.27	0.01
East North Central	5	1.54	-1.54	1.11	-0.41	0.98	-0.85
West North Central	7	1.60	-0.01	1.73	1.67	1.47	0.63
South	17	2.60	-2.39	0.87	0.44	1.58	-0.38
South Atlantic	9	3.50	-3.11	0.78	0.29	2.24	-1.05
East South Central	4	0.89	-0.89	1.04	1.04	0.95	0.95
West South Central	4	2.29	-2.29	1.01	0.28	0.73	-0.17
West	13	3.75	-0.18	3.66	3.05	4.00	2.06
Mountain	8	4.41	-0.44	4.03	3.76	4.68	1.90
Pacific	5	2.69	0.22	3.08	1.92	2.91	2.33

domestic migration, and international migration) derived from vital statistics and the administrative records for 7/1/1995 to 6/30/2000 are used to compare the projected components of change in the same period.

As Table 4 shows, the projected births are more accurate than other components with the lowest Mean Absolute Percentage Errors, followed by the MAPE for deaths. The net domestic migration is the worst component in the projection - the MAPE reached 193.3 percent. The MAPE of net international migration was 31.5.

after adjusting undercount except the West. The reason for the West to have higher MAPE after the adjustment was made is that many states in that region were over-projected initially. For example, Idaho, Montana, Wyoming, Hawaii, and Alaska have a higher projected population over the 2000 census count. Once their projected populations were inflated by the undercount rates, the Mean Absolute Percentage Error for the region becomes higher.

4. Accuracy of Projected Components of Change

Since the Cohort-Component Method was used to produce the state projections, the accuracy of every component should affect the accuracy of the projections.

Table 4: Mean Absolute Percentage Error of Projected Components of Change by Region and Divison: Mid Year 1995-2000

Region/ Division	Births	Deaths	Net migration	
			Domestic	International
All States	6.3	7.0	193.3	31.5
Northeast	7.5	10.5	104.5	35.9
New England	9.1	8.6	139.7	49.1
Middle Atlantic	4.4	14.3	34.0	9.6
Midwest	2.6	7.9	186.1	33.5
East North Central	2.4	7.4	213.6	29.8
West North Central	2.7	8.2	166.4	36.2
South	5.7	6.0	71.9	22.4
South Atlantic	6.5	7.1	44.1	22.6
East South Central	4.3	4.0	147.4	17.6
West South Central	5.3	5.5	58.7	26.5
West	9.6	5.1	420.1	38.5
Mountain	9.9	4.1	606.6	50.9
Pacific	9.2	6.7	121.8	18.6

The estimated components of change (births, deaths,

V. Multiple Regression Analysis of Factors Affecting the Accuracy of State Projections

The relationship between the 1990 census undercount and the accuracy of the state population projection discussed above seems to indicate that the impact is not very dramatic - a reduction of MAPEs from 2.6 to 2.2 (see Table 3). This may be due to other source of errors in the projections - such as errors in the projected components - births, deaths, domestic migration, and international migration, along with the errors in population estimates.

To identify the impact of census undercount on the accuracy of state projections, we need to hold the other source of errors constant in a regression analysis. The dependent variable for the analysis is the absolute percent error of state population projections. The independent variables include - 1990 census net undercount rates, absolute percent error of state estimates, absolute percent error of projected births, deaths, net domestic migration, and net international migration. In addition, the absolute percent population change between 1990 and 2000 is used to measure the uncertainty of the projections in predicting future trends. The units of analysis are 50 states and the District of Columbia.

The reason to use the absolute percent errors for analysis is that the accuracy of the state projections is measured by the absolute percent difference between projected population and census count. The same measurement of the accuracy of independent variables is also used to be consistent with the dependent variable, which does not take into account the bias of the projections.

(1). Correlation between Projection Error and Dependent Variables

Before presenting the results of multiple regression analysis, we need to present the correlation between dependent and independent variables - the gross relationship between two variables without holding other variables constant. Table 5 shows the simple correlations among these variables. As expected from the discussion above, the projection errors are highly correlated with

of tax returns addresses overtime, and different levels of coverage rates of population by tax returns among states may contribute to the uncertainty of this variable. The migration flows used in the projections may not reflect the true migration and the estimated net domestic migration to evaluate the projected domestic migration may not be accurate.

(2). Multiple Regression of Factors Affecting the Projection Accuracy

The simple correlation between two variables may include the effects of other variables on the specific variable. For example, the correlation between errors in projected births and errors in projected population may be due to the impact of state estimates and census undercount on the projected births because the census undercount and errors in the state population estimates affect the accuracy of population base to derive fertility rates for the projections. In other words, the impact of errors in births on projection errors is also due to the effects of errors in state estimates or census undercount on projections at the same time. We need to hold other variables constant. The results of the multiple regression analysis in Table 6 show the importance of each variable contributing independently to the projection errors while holding other variables

Table 5: Correlation Matrix of Absolute Percent Projections Errors and Independent Variables: 50 States and the District of Columbia

Variables	Absolute Percentage							
	Projections Error	Undercount Rate	Estimates Error	Births Error	Deaths Error	Domestic Mig Error	International Mig Error	Pop Change 1990-2000
Projections	1							
Undercount	0.474*	1						
Estimates	0.724*	0.411*	1					
Births	0.565*	0.394*	0.591*	1				
Deaths	0.334*	-0.051	0.353*	0.332*	1			
Domestic Mig	-0.057	0.068	-0.018	0.160	-0.205	1		
International Mig	0.272*	-0.002	0.255*	0.182	0.147	0.272*	1	
Pop Change	0.419*	0.493*	0.595*	0.338*	-0.273*	0.092	0.124	1

* Significant at 0.05 level

percent error in state estimates (correlation coefficient of .72), and also related to census undercount rates (0.47). The projection error is also associated with population change (0.42) -- a dramatic change in population would usually produce a larger error in projections.

The general perception is that the percent errors in the projected components should be the primary source of errors in the projections because the projections were based on the cohort component method. As expected, the error in projected births is significantly correlated with the projection errors (0.57). However, the percent errors in projected deaths and international migration only correlate moderately with errors in population projections. Surprisingly, the percent error in domestic migration has no correlation with percent projection errors. This indicates that a state with higher percent error in projected domestic migration may not necessarily have a higher percent error in projections. This may also reflect the problems of measurement of domestic migration based on IRS data. Changes in tax laws, problems in the geo-coding

constant in two conditions and how much all the variables together can explain the projection errors.

Table 6 shows the standardized regression coefficients of the independent variables on percent projection error in 2 models. Model 1 includes only percent errors in births, deaths, domestic migration, and international migration. Model 2 include census undercount rates, state estimates errors and population change between 1990 and 2000, in addition to the variables in model 1. The errors in the projected components as shown in model 1 explain about over 40 percent of projection error (R-square of 0.40). The percent error in projected births accounts for most of the weight (coefficient of 0.52), followed by international migration (0.21). The errors in projected deaths and domestic migration do not explain the variation in percent projection errors in the 50 states and District of Columbia. Surprisingly, when other components are held constant, the domestic migration tends to have a slight negative impact on projection accuracy. This further indicates that

Table 6: Standardized Regression Coefficients of Independent Variables or Absolute Percent Error of State Projections

Independent Variables	Standardized Coefficients	
	Model 1	Model 2
Absolute % error in projected births	0.52*	0.17
Absolute % error in projected deaths	0.09	0.07
Absolute % error in projected domestic migration	-0.18	-0.11
Absolute % error in projected international migration	0.21*	0.14
1990 census undercount rate	–	0.24*
Absolute % error in state estimates	–	0.48*
Absolute % Population change:1990-2000	–	-0.03
R	0.63	0.78
R-Square	0.40	0.61
Adjusted R-Square	0.34	0.55
Residual	0.78	0.62
F	7.52	9.67
Significance	<0.001	<0.001

* Significant at 0.05 level.

problems with the measurement of domestic migration in the state population estimates and population projections.

When the census undercount rate and percent errors in state estimates are included in the regression, all the variables explain over 60 percent of variation in projection errors. Most of the projection errors originally explained by the projected components of change are taken over by the percent errors in state population estimates and the census undercount. The standardized coefficient of percent errors in births is reduced from 0.52 to 0.17. The percent error in state estimates stands out as the most powerful variable in explaining errors in the state population projection -- 0.48, followed by the census undercount (0.24).

Table 7: Difference between State Estimates and Census 2000 Count and Mean Absolute Percent Errors

Region and Subdivision	Difference between Estimates and Census 2000					MAPE	
	1990 Census Base		Undercount Adjusted Base		% due to Census Undercount	1990 Census Base	Undercount Adjusted Base
	Number	Percent	Number	Percent			
All States	-6,813,550	-2.42	-2,874,106	-1.02	57.82	2.57	1.47
Northeast	-1,460,378	-2.72	-1,051,712	-1.96	27.98	2.61	1.98
Midwest	-797,850	-1.24	-384,335	-0.60	51.83	1.44	0.95
South	-3,013,859	-3.01	-1,178,952	-1.18	60.88	2.89	1.35
West	-1,541,463	-2.44	-259,107	-0.41	83.19	3.16	1.76

Notes: The estimates are derived from the components of change between 4/1/90 and 4/1/2000 addting to the 1990 Census count. The 1990 census base estimates are based on the 1990 census official count as enumerated. The undercount adjusted base estimates are based on the undercount adjusted 1990 census counts. The components of change include births, deaths, net domestic migration, net international migration, federal-civilian movement and residual adjustments.

The reason for such dramatic shifts in explaining the errors in projections is that the state population estimates are not only used as the starting population base to launch

projections, but also are used as the population controls to developing fertility, mortality, and migration rates. This can be seen from the correlation between percent errors in projected births and percent errors in state estimates (0.59), and the correlation between errors in projected deaths and state estimates (0.39).

3. State population estimates and 2000 census count

Since the population estimates affect the accuracy of the projections considerably, we need to evaluate the accuracy of the state population estimates. As shown in Table 7, the estimates based on the official 1990 census count under-estimated the U.S. population by 2.4 percent or a total of 6.8 million people. Almost all states had the estimated population lower than the census count except West Virginia. The West had the highest percent error (3.2 MAPE), followed by the South, and the Northeast region. The Midwest had the lowest percent errors (1.4%).

If we use the undercount adjusted 1990 population as the base to derive the state estimates, we can see a dramatic reduction of estimation errors. The amount of under-estimation for the U.S. as a whole decreases from 6.8 million to 2.9 million, a 57.8. percent reduction. The percent error of the estimates for the entire U.S. decreases from 2.4 to 1.0 percent. The mean absolute percentage error (MAPE) for all states dropped from 2.6 percent to 1.5 percent. The reduction of percent errors in state estimates based on the 1990 census undercount adjusted count is so overwhelming that all regions have a reduction of estimation errors, especially in the West. The reduction of errors by the use of undercount adjusted 1990 census population ranges from 28.0 percent in the Northeast to 83.2 percent in the West. The overall effect of the 1990 census undercount accounts for more than one half of the errors.

The strong relationship between the census undercount and errors in population estimates indicates that the 1990 census undercount not only has a direct impact on the

accuracy of the projections, but also has a very strong indirect impact channeled through its impact on the accuracy of population estimates.

VI. Conclusions and Implications

The accuracy of population projections depends on many factors. It has been shown that the level of accuracy or magnitude of errors depends on whether the census counts or the population estimates are used for evaluation. If the population estimates are used to evaluate the accuracy, the Census Bureau's projections for the nation and states are very accurate. The net census undercount rates in the 1990 census are significantly higher than the undercount rates in the 2000 Census. Consequently, the major factor contributing to the discrepancy between the projections and 2000 census counts is the undercount of the 1990 census. For the national projections, the 1990 census undercount accounts for 2/3 of the discrepancy, far more than the errors contributed by the international migration and other source combined.

The multiple regression analysis of factors affecting the variation of accuracy of the projections among 50 states and the District of Columbia shows that errors in the state estimates is the most important variable contributing to the state projection errors, followed by the 1990 census undercount. The errors in the projected components - births, deaths, domestic migration and international migration should have contributed a significant amount of errors in projections. But, when the state estimates and the 1990 census undercount are taken into account, the impact of errors in the components becomes insignificant.

The 1990 census undercount accounts for more than one-half of errors in the state population estimates, which in turn, contributes most of the errors in the state projections. Since the census undercount contributed 2/3 of the discrepancy between that national projections and the census count and the national projections are used to control the state projections, the ultimate source of the errors for the latest projections is the 1990 Census undercount. This further indicates the importance of the accuracy of base year population in producing accurate projections.

These results suggest that if we want to improve the projection, we need to pay attention to the accuracy of base year population and the accuracy of population estimates. However, whether we use the undercount adjusted base population for projections is subject to Census Bureau's policy. Therefore, it is necessary first to

ensure the accuracy of projected births because it explains the largest proportion of projection errors among the components. It will be more cost-effective to do so because any improvement in projecting births can have a noticeable effect on projection accuracy. On the contrary, it may take more effort to make improvement in the domestic migration component for projections because its direct impact is mixed - it can go in either direction depending on other errors. This does not mean we should not pay attention to this important component in projections. We should know that no matter what we do to improve this component we may not expect to get the expected results. In other words, we do not need a complicated model to project the migration. What we need is a simple, reasonable, and understandable model to explain to the user what we do. Demographers repeatedly indicate that complex techniques did not produce more accurate forecasts or projections (Smith and Sincich, 1992),

Since the 2000 census is more accurate than the 1990 census in terms of smaller net undercount rates, we should have a better population base for projections. However, we do not know how the next census will do. The uncertainty of the net undercount rates for the next census will affect how the accuracy will be measured. The uncertainty of the base year population and the target year population for evaluation will certainly complicate the determination of the performance of the projections.

References

Campbell, Paul R. 1996a, "How Accurate were the Census Bureau's State Population Projections for the early 1990's?" Paper presented at the Federal Forecasters Conference, Washington, D.C. May 2.

Campbell, Paul R. 1996b, "Population Projections for States by Age, Sex, Race, and Hispanic Origin: 1995 to 2025," PPL-47, U.S. Census Bureau, Population Division, October.

Campbell, Paul R. 1997, "An Evaluation of the Census Bureau's 1995 to 2025 State Population Projections - One Year Later." U. S. Census Bureau, Paper presented at the Population Association of America Meeting, Washington D.C. (March)

Hollmann, Frederick, Tammany Mulder, and Jeffrey Kallan, 1999, "Methodology and Assumptions for the Population Projections of the United States: 1999 to 2100." U.S. Census Bureau, Population Division Working Paper, No. 38, December, 1999.

Hogan, Howard and Gregory Robinson, 1993. "What the Census Bureau Coverage Evaluation Programs Tell Us About Differential Undercount," Proceedings of 1993 Research Conference on Undercounted Ethnic Populations, U.S. Census Bureau

Robinson, J. Gregory, 2001a, "Accuracy and Coverage Evaluation: Demographic Analysis Results," U.S. Census Bureau, DSSD Census 2000 Procedure and Operations Memorandum Series B-4* (March 12, 2001)

Robinson J. Gregory, 2001b, "ESCAP II: Demographic Analysis Results," U.S. Census Bureau, Executive Steering Committee for ACE Policy II, Report No. 1 (October 13, 2001).

Smith, Stanley K. and Terry Sincich, 1990, " The Relationship Between the Length of the Base Period and Population Forecast Errors," Journal of the American Statistical Association, Vol 85, No. 410, 1367-1375.

Smith, Stanley K. and Terry Sincich,1992, "Evaluating the Forecast Accuracy and Bias of Alterative Population Projections for States." International Journal of Forecasting, 8, 495-508.

Swanson A. David, Jeff Tayman, and Charles F. Barr, 2000, "A Note on the Measurement of Accuracy for Subnational Demographic Estimates." Demography, Vol.37, No. 2 May 2000:193-201.

Tayman, Jeff and David A. Swanson, 1999, "On the Validity of MAPE as a Measure of Population Forecast Accuracy," Population Research and Policy Review: 18: 299-322.

Tayman, Jeff, David A. Swanson, and Charles F Barr, 1999, "In Search of the Ideal Measure of Accuracy for Subnational Demographic Forecasts." Population Research and Policy Review, 18: 387-409.

U.S. Census Bureau, http://www.census.gov/population /www/projections/st_yr95to00.html
U.S. Census Bureau, http://www.census.gov/population /www/projections/st_comp-chg.html

U.S. Department of Commerce, Press Release, 2001, "Statement of Acting Census Bureau Director William Barron Regarding the Adjustment Decision." CB01-CS.08, October 17, 2001.

Wang, Ching-li, 2002. "Evaluation of Census Bureau's 1995-2025 State Population Projections" U.S. Census Bureau, Population Division, Working Paper No. 69 (forthcoming)

Wetrogan, Signe I. And Paul R. Campbell, 1990, "Evaluation of State Population Projections." Paper presented at the Population Association of America, Toronto, Canada, May 3-5.

FORECASTING UNCERTAINTY IN UPCOMING CENSUS BUREAU POPULATION PROJECTIONS

Frederick W. Hollmann, U.S. Census Bureau

In formulating its first round of national population projections based on Census 2000, the Population Projections Branch at the U.S. Census Bureau is seeking to incorporate systematic estimates of forecast uncertainty. In previous editions of Census Bureau population projections, the issue of uncertainty was addressed through simultaneous release of "highest", "medium", and "lowest" series. The highest and lowest series incorporated assumptions on fertility, mortality, and international migration that were, respectively, most favorable and least favorable to population growth. The criteria for selecting fertility, mortality, and international migration for the extreme scenarios was confined to the researchers' judgement, with no attempt to assess probabilities that any result would fall within the range between the two extremes. For each component, we allowed the range embraced by the extremes to expand over time, reasoning that uncertainty would increase with time from the base year to the reference year. However, we made no allowance for the effects of deviations from medial assumptions that would either cancel or reinforce each other over time. In fact, we have never specifically addressed the question of whether the extent to which an assumption was high or low was measured in a cumulative sense, or in a static sense. To consider the case of an assumed total fertility rate in 2050, for example, the cumulative sense relates to the number of babies born from the base date to 2050, and would best represent the impact of our uncertainty on the population in 2050, yet the static sense was most likely closer to our focus in proposing these scenarios. It is our aim therefore, to systematize the projection of error distributions of the central series, so as to represent the year-to-year fluctuations from trend that have governed historical time patterns in the components of population change. These can become a model for future uncertainty, which we would quantify as part of our product.

This attempt must occur in a significantly altered data environment, with respect to the reporting of race. All national population projections until now have assumed that a respondent could provide only a single race. Moreover, persons in the census base population who responded in a way that would place them outside of four major racial categories (to take the example of the 1990 census) were assigned a race within one of these categories. This could occur either through census editing, or for a special imputation for the purpose of establishing a base population for estimates and projections consistent with other administrative data. With Census 2000, there are two major changes. One of the 1990 race categories, Asian and Pacific Islander has been divided into two–"Asian" and "Native Hawaiian and other Pacific islander". The 1990 categories of White, Black, and American Indian Eskimo, and Aleut, are retained. However, a second change, far larger in its technical implications, is that respondents in Census 2000 could report as many races as applied, rather than a single race. As in 1990, the residual category ("some other race") is allocated to one of the major categories for purposes of estimates and projections only. Even with the remaining five major races, there were then 31 possible combinations of responses. Hispanic origin, defined as in 1990, can be collapsed into two major categories–non-Hispanic and Hispanic–whose cross-classification with race yields a possible 62 racial and ethnic categories to estimate or project.

The purpose of this paper is to suggest a general framework for projecting the national population in 2002, that incorporates systematic estimates of uncertainty, while representing a new distribution of race based on Census 2000 We start by indicating what we hope to achieve, with respect to product and method, then discuss the major issues we must encounter in reaching those objectives.

Generalized goals

We strive to achieve four major goals in this series of population forecasts that differentiate them from population projections produced by the Census Bureau in the past.

1) Projections will be probabilistic, in the sense that (at minimum) there will be prediction intervals identified around any central series of total population, and some detailed categories.

2) Probabilistic projections will be generated via a probabilistic treatment of the components of change. Stochastic procedures will be used to generate distributions of age-specific fertility and mortality schedules, as well as international migration, which will be

used to generate a large number of cohort-component projections of population, (called "realizations.") Distributions from these realizations will yield the necessary prediction intervals.

3) Race and ethnicity differentials must be incorporated in the assumptions, to allow for "composition effects" on population growth resulting from differences in vital rates. This means that assumptions regarding components of change will be imposed on *some categories* of race and Hispanic origin, and that the total population for each realization will be a sum of the resulting series. This practice has been incorporated in past series, but since the number of racial groups was much smaller, we could simply treat the racial and ethnic groups as a separate population summing to the whole, for purposes of carrying out cohort-component realizations.

4) The results of the projections will need to reflect the new Office of Management and Budget (OMB) guideline regarding the definition of race, with full cross-classification by Hispanic origin. Specifically, our ultimate product will consist, at minimum, of two matrices for each projected reference date.

      a) We must show age by sex by five "minimal" (sole reported) race categories, plus a sixth race category that is all multiple report categories combined–all for the non-Hispanic population as well as the two values of Hispanic origin combined. These categories will sum to the non-Hispanic and total populations, respectively, by age and sex. In addition, we show the Hispanic population by age and sex.

      b) We must show age by sex by five "maximal" race categories, consisting of all those who reported the race at least once. Again, this distribution is to be reported both across Hispanic origin and for the non-Hispanic population. These categories will sum to the total projected number of MARS race reports (people weighted by the number of groups reported) for the total and non-Hispanic populations, respectively, by age and sex.

It is *not* stipulated that stochastic procedures yielding prediction intervals will be applied to the entire detail described here.

We are accepting as a corollary of requirement 4 that it will be necessary, as an intermediate step in the process, to produce age-sex distributions for 62 cross-categories of race and Hispanic origin. The 31 racial categories will be determined as all combinations of 1, 2, 3, 4, or 5 categories,

to be cross-classified by two categories of Hispanic origin. This is not to say that projected population series for each of the 62 categories need to conform to the level of technical sophistication required for the larger aggregates described above.

As a corollary of requirements 3 and 4, I am assuming that racial groups–even for females only--can no longer be perceived as "closed" with respect to childbearing as they have been in the past. In particular, single-race parents of different race will produce some multiple-race children, so "mother rule", and "father rule" are no longer helpful by themselves. (In our last projections, we used "mother rule", rightly or wrongly.)

Forecasting components of change via a stochastic renewal process

Various approaches to forecasting uncertainty in the demographic components of change have been proposed in the literature on population forecasts. Generally, they tend to follow two major
approaches. One approach is to solicit expert opinion on future demographic rates "ex ante", both as to their level and also as to possible ranges of uncertainty. On this basis, it is possible to develop a distribution of future scenarios for each of the components, that can be randomized to produce a large number of projected population series, each representing a scenario. The scenarios thus generated would incorporate trends in fertility, mortality, and migration that would fall, with a certain pre-specified probability (e.g., 90%) within the specified ranges (Lutz, Sanderson, and Scherbov, 1998). While expert opinion produces future values of the components of change, the experts would undoubtedly take account of historical data series in making their assessments.

A second approach bases uncertainty systematically on the observation of extant data series, either by observing "ex post" the capacity of forecasting methods to reproduce existing data, or by observing trends and fluctuations in existing series directly, and generating models of trend and error variance using established time series methods. This approach is applied in the work of Ronald Lee, Juha Alho, and others (for example, Alho, 1990, Lee and Tuljapurkar, 1994, Lee, 1998). Lee, in particular, argues that the "random scenarios" approach described in the previous paragraph is unrealistic, and even contains statistical biases. Because all projected time trends in any of the three components are linear multiples of each other, there is no possibility of fluctuations in the

trends, so that the error distributions are likely to be understated. Moreover, Lee is skeptical of the ability of experts to assign ranges within probabilities, even if they are able to evaluate current information to hypothesize a future central trend (Lee, 1998.)

The proposed alternative (e.g., Lee, 1998) is an autoregressive stochastic renewal process. Parameters that may change over time are determined for each of the components of population change. Time series methods, such as autoregressive integrated moving averages (ARIMA) are used to develop stochastic models with the parameters as random variables, that may be correlated with each other. Through the selection of random numbers, realizations of the models for the parameters are generated, and the resulting parameters are used to calculate the schedules of age-specific fertility and mortality rates, and international migration allowances. These can then be applied to produce population projections via the cohort component method. This process is repeated a large number of times, to form a data base of realizations of components of change and population, from which any population indicator may be calculated, with all the elements of a probability distribution, including measures of central tendency and dispersion, and probabilistic ranges of outcomes.

Parameters of fertility, mortality, and migration

Because fertility, mortality, and international migration are not single values, but matrices of single year of age by sex by race by Hispanic origin, it is necessary to summarize distributions in the form of a smaller number of time-dependent parameters. Time series models can treat vectors as independent variables, but they become impractical if the vectors have a large number of elements. Various methods have been proposed to model age schedules of fertility and mortality to minimize the number of parameters that are time dependent, and whose time trend captures most of the impact of the component on projected population. Race and Hispanic origin cannot be treated in this way, as their cross-categories represent a non-metric variable with multiple categories.

In the case of mortality, the prominent method in the recent literature is the one developed by Ronald Lee and Lawrence Carter (Lee and Carter, 1992). This method isolates a single time-dependent scalar parameter from an age schedule of mortality that is defined as the age-independent multiple of a time-independent schedule of age-specific changes in the log of the age-specific mortality rates. This method allows the estimation of a rather simple stochastic time-series model for the scalar

parameter. For any realization of the model, the entire schedule of mortality rates can be regenerated from the forecast value of the time-dependent scalar, and the time-independent age data. Whether a single-parameter model will ultimately yield an appropriate model for our purposes remains to be determined, however, we anticipate that the process of modeling mortality will follow this general paradigm.

A number of methods have been proposed to summarize the age pattern of fertility. Ronald Lee and Shripad Tuljapurkar provide one example, wherein the total fertility rate is accepted as a single time-dependent parameter (Lee and Tuljapurkar, 1994). While this may not be optimal for a study of fertility by age of mother, Lee and Tuljapurkar make the case that shifts in the age pattern of fertility in the long term have limited impact on population change. Alternatively, functions of age fitted to the pattern of age-specific fertility allow multiple parameters that describe the fertility schedule, and submit to time-series modeling. An example of this is the gamma curve, application of which is discussed by Nico Keilman and Pham Dinh Quang (Keilman and Quang, 1998). However, neither of these models facilitates the consideration of trends in age-parity data, that provide information on the effects of postponement of childbearing on completed fertility. We expect to develop fertility parameters that reflect level of fertility, as well as age-parity information that can be incorporated in an autoregressive time series model.

Considerably less research has been addressed to the stochastic treatment of international migration in projections. Much of the work that has been done in the area of probabilistic projections has focused on fertility and mortality, while allowing international migration to be deterministic. Because of the nature of demand for Census Bureau forecasts, it is our aim that the probabilistic projection of population should be exhaustive, in the sense that we must be able to account for prospective error in all components of change. We anticipate there will be at least two time-dependent parameters, allowing for trends in the age-sex pattern of net migration, as well as the magnitude.

Finally, it can be argued that we need to incorporate uncertainty in the base population, since studies of decennial census data clearly indicate errors stemming from coverage. We do not plan to address this issue in our population forecasts, primarily because Census 2000 results will define the population universe for our projections. Under this definition, all population results, including prediction intervals, must be interpreted under whatever constraints this definition imposes.

It was observed at the outset that our product requirements dictate that we produce projections (at least in the production stages) of a vector of 62 cross-categories of race and Hispanic origin, for each value of age and sex. Many of these categories are very small (especially those involving some multiple-race combinations), and the methodology described above would not be practical for such categories. In addition to this, since the present definition of race is new with Census 2000, there is no historical basis for estimating trends in the components of population change for these categories.

We are proposing to establish a "first stage" of the population projection process include only the following three categories: 1) Hispanic origin, of any race, 2) non-Hispanic, single-reporting Black, and 3) all others. The process of projecting these groups, with projections of uncertainty, will determine the total population by age and sex and its forecast uncertainty. This three-category distribution has the following desirable properties, from the standpoint of population forecasting.

1) The categories are large, so there should be no danger of projecting single-year age distributions too "thin" to produce good results.

2) Differential fertility and mortality among the three categories can be documented from historical data, although there is some potential for bias in the comparisons. Hispanic origin was not directly affected by the change in race definition, so fertility and mortality rates can actually be computed directly for the first group. The non-Hispanic Black population was affected by the change in race definition, but relatively marginally, so that assumptions could be made to produce reasonable estimates of fertility and mortality.

3) Differences among the categories with respect to demographic rates are substantial. The age pattern of childbearing and excess mortality (notably for infants) for the non-Hispanic Black population are distinctive. Fertility continues to be somewhat higher for the Hispanic population than for most non-Hispanic racial categories. (We may assume convergence over time, but the convergence should not be immediate.)

We would claim that there are no other categories of race and Hispanic origin that enjoy these properties. One could make a weak case for separating the non-Hispanic aggregate of Asians and native Hawaiians and other Pacific islanders (NHOPI). There is some evidence of exceptionally low mortality rates for this group, although the extent of the differential is subject to bias caused by differential reporting in numerators and denominators. Fertility, while unexceptional in level follows a later age pattern. The category would arguably be large enough to generate stable forecasts. However, it fails very badly on point 2), since a substantial portion of the "old" Asian and Pacific islander category for which rates have been measured has become multiple-race under the "new" definition.

In the case of non-Hispanic American Indian, Eskimo and Aleut, the bias in vital rates caused by differential reporting of "old" race in vital registration and census data is both large and unknown. Furthermore, the "new" race category based on single responses is most likely a small subset of the old one.

In order to implement the autoregressive stochastic renewal process for the three components, we must treat the dependent variable of all autoregressive time series models as a vector with a variance-covariance matrix that can be estimated. Thus, for example, the trend in the fertility of the Hispanic and the non-Hispanic Black population in a particular realization of the model can be assumed to be positively correlated with that of the non-Hispanic white population, without assuming that the three are in "lock step", (i.e., correlations of unity). In this way, the compositional effects of the changing racial and ethnic distribution of the population can be reflected in the stochastic population projections.

Absence of closure of racial categories with respect to childbearing

In the last series of long-term population projections, we assumed that children were born into the race of their mothers, and retained their mothers' race categories throughout their lives. While we recognized that racially and ethnically exogamous unions could produce offspring of a race other than that of the mother, we assumed that the effect of father's race on the way a child's race would be reported in a census would net to zero. Otherwise stated, the overstatement of a given race of child for mothers of that race, would be matched by the understatement of the same race for mothers of other races. We made this assumption even in the presence of some unpublished research that indicated, based on census data on children in households, that this is not always the case. We were most likely understating the number of Black children, since children of one Black parent are more likely to be reported Black than the race of the other parent. We were most likely overstating the

number of Asian children somewhat, because exogamy is far more frequent among Asian women than among Asian men, and parental couples in which the mother is Asian are more likely than not to report the father's race for the race of children.

With the new race definition used in Census 2000, assuming closure for racial groups of women with respect to reproduction, besides producing biased projections, would be illogical. Clearly, multiracial categories are composed largely of people with parents of different race, whether single or multiracial. Hence, single-race mothers will produce children of multiple race–a likely event if the father is of a different race. On the other hand, mothers of multiple race are less likely to produce children of a single race.

To make matters more complicated, there is no deterministic logic that correctly specifies race of child by the race of both parents. There is evidence that parents of different race frequently select the race of one parent rather than include the race of both parents in a child's race report. Hence, an algorithm that would generate race of child simply as the union of the races of the child's parents would overstate multiracial children.

The method of allocating child's race from parents' race must be based on data from the census on children within family households, as it is currently the only major data source. Because cohort-component projections are likely to assume away any "drift" in race reporting, it is advisable to base race allocation on the age of children of various ages, as opposed to age "under 1". This can serve as a foil against racial drift. Ultimately, it would be appropriate to consult birth registration data for this purpose, but it is altogether possible that no comprehensive registration data providing new-definition race of both parents from birth certificates will be available until late in the present decade.

Even the three large racial and ethnic categories selected as candidates for stochastic forecasting are not "closed" with respect to childbearing. Most importantly, Hispanic and single-race Black mothers will have children who belong to the residual category. It would still be necessary to allocate births by race and origin, based on mother's and presumably father's race and origin. Assuming the first-stage process involves three race/ethnic categories, we have 9 combinations of mates producing children, and three possible outcomes with probabilities summing to unity for each combination. We may assume that fertility of women of a given category is independent of her mate's race, but the distribution of race

of child for each parental combination must still be estimated, and will evolve throughout the projection series with the evolution in the distribution of available mates

Forecasting the entire race distribution

Having produced stochastic forecasts of three major racial categories by age and sex, it remains to produce information for the entire distribution of 62 categories. Within Hispanic origin, there are a total of 31 racial groups. Single-reporting non-Hispanic Blacks comprise a single group. The third, residual category includes the remaining 30.

We are proposing to treat this expansion of the matrix deterministically (without estimating uncertainty) in the following way.

1) Based on census data for the entire racial distribution, we determine a set of sub-categories of the three large ones that can effectively be projected via the cohort component method. This is a qualitative decision that would require scrutinizing the census 2000 base population. Simply stated, categories need to be large enough that age-to-age fluctuations in single-year age groups are not large compared to the age groups themselves–except where this is expected (e.g., cohort of 1946, and cross-racial age heaping.) We would hope that the non-Hispanic, single-report categories of all the major OMB races would qualify.

2) We decide which categories are large enough or different enough to warrant separate estimates of fertility and mortality. It is presently assumed that most will not. For those that do, projections must be generated. The most likely candidates for this treatment would be the aggregate of the Asian and the Native Hawaiian and Other Pacific Islander categories, for which we might impute fertility and mortality schedules from the "old" Asian and Pacific Islander category, to reflect the lower mortality and later fertility schedules characteristic of this group.

3) Based on census 2000 data on households, a matrix needs to be developed that shows the probability distribution of race/origin of child for each race/origin combination of parents. This matrix will be very large, as there are 3,844 parental combinations with 62 categories of race and origin. If we assume that no child can have a race combination with elements not found in the response of either parent, the possible outcomes for the 3,844 parental combinations combine is in excess of 90,000.

4) Cohort-component projections must then be carried out

for each of the race-origin cross-categories identified in step 1. The results for components of change can be proportionately adjusted to a realization of the median values of the projection results for the large categories.

5) Populations for any race categories not yet projected will be determined by constant, census-based ratios of the population matrix produced in the previous step.

Stochastic forecasting in the presence of unprecedented detail

The task before us is, to the best of our knowledge, without precedent, because we are projecting characteristic detail for a variable, the composite of race and Hispanic origin, that is not "closed" with respect to reproduction. Hence, it would appear necessary–at minimum–to model random variation in the distribution of father's race for each race of mother, as it is a palpable factor in the determining child's race in a multiple-race environment. Yet, we have no historical basis for estimating this variation, since it would require a historical time series of data on fertility under the new race definition.

Primarily for this reason, we are proposing that the "first step" described above, projecting three race-ethnic groups by age and sex, represents the extent of detail for which we can predict uncertainty in the race distribution. At the same time, it appears important to be able to predict uncertainty by race at this level, since one cornerstone of the overall procedure is that racial differentials in fertility should be a factor driving the projection of total population.

Our suggestion that we cannot provide a stochastic treatment of the full matrix of 62 race-origin groups is grounded in a modest consideration of feasability. First, producing sufficient realizations to generate distributions of this much detail would most likely be prohibitive from the standpoint of computer resources. Second, there is no historical basis on which to carry out time-series analyses of racial groups that are significantly different between the pre-1999 and post-1999 OMB classifications. Third, even if such series existed, some of the categories are most likely too small to submit to this kind of analysis. Fourth, even if it were possible to produce results, they would not be sufficiently comprehensible to submit to review, much less interpretation.

One likely consequence of this problem is that we will not be able to show uncertainty in projections of the population in racial categories. Showing uncertainty for

three large categories, as it would be computed, would inevitably invite criticism of statistical favoritism.

A further point of future disappointment that we need to consider is that the projection of uncertainty for fertility rates from historical data is extremely sensitive to the choice of historical period on which the uncertainty is based. Stated simply, including the post-war baby boom in any analysis tends to produce distressingly high levels of uncertainty. This may be mitigated somewhat by basing the analysis on cohorts of women, even if some of the cohorts are incomplete.

Discussion

In generating and promulgating population forecasts with projected uncertainty as a public product, we are facing a challenge beyond what is inherent in an illustrative projection in a methodological presentation. Projections issued by the U.S. Census Bureau are considered to be an official product, and are widely adopted as such by both the government and the private sector–often without sufficient evaluation. We fully expect that many users will appropriate the central values of the forecast, with little consideration of the prediction intervals, as the basis for planning and decision-making, much as the middle-level series has been adopted in the past. More to the point, we expect that a different class of users, more "in tune" with the science of producing forecasts, will over-interpret our ability to forecast uncertainty, assuming that prediction intervals are absolutely definitive of the uncertainty of our projections. Like the central forecasts themselves, prediction intervals are based on historical observation, which, like the basis for the central forecasts, can be misleading. Moreover, every aspect of such a forecast–including its uncertainty--is based on a myriad of decisions made by researchers involved in producing the forecast, all of which are fallible.

What we hope to gain from this endeavor is a better appreciation by our users that forecasts (or projections, read as forecasts) are not definitive of a future reality. Rather, they are a reflection of a cumulative set of observations, interpreted by demographers and statisticians, that lend a great deal of uncertainty to the interpretation.

References

Alho, Juha, 1990, "Stochastic Methods in Population Forecasting", *International Journal of Forecasting,* vol. 6., pp. 521-530.

Keilman, Nico, and Pham Dinh Quang, *Predictive Intervals for Age-Specific Fertility,* Statistics Norway, Documents, No. 98/19, November 1998.

Lee, Ronald D., and Lawrence R. Carter, 1992, "Modeling and Forecasting U.S. Mortality", *Journal of the American Statistical Association,* vol. 87, no. 419, pp. 659-675.

Lee, Ronald D., and Shripad Tuljapurkar, 1994, "Stochastic Population Forecasts for the United States: Beyond High, Medium, and Low", *Journal of the American Statistical Association,* vol. 89, no. 428, pp. 1175-1189.

Lee, Ronald D., 1998, "Probabilistic Approaches to Population Forecasting", in Wolfgang Lutz, James W. Vaupel, and Dennis A. Ahlburg (eds.) *Frontiers of Population Forecasting,* supplement to vol. 24 of *Population Development Review,* pp. 156-190.

Lutz, Wolfgang, Warren C. Sanderson, and Sergei Scherbov, "Expert-Based Probabilistic Population Projections", in *Frontiers of Population Forecasting,* supplement to vol. 24 of *Population and Development Review,* pp. 139-155.

This page intentionally left blank.

OPTIMIZATION OF POPULATION PROJECTIONS USING LOSS FUNCTIONS WHEN THE BASE POPULATIONS ARE SUBJECT TO UNCERTAINTY

Charles D. Coleman, U.S. Census Bureau, 4700 Silver Hill Rd., Stop 8800, Washington, DC 20233-8800
Email: ccoleman@census.gov

1. Introduction

Coleman (2000) discussed the use of loss functions for the evaluation and optimization of population projections. Implicit in that article was the assumption that base populations are not subject to revision. The present article uses Coleman's (2000) framework and adds uncertainty to the base populations upon which projections are based. The projected populations are modeled as the product of the base populations and the projected growth ratios.[1] The projector, in effect, chooses growth ratios to minimize expected total loss: the expected value of the sum of the individual areas' losses. The growth ratios and base populations are subject to both risk in Knight's (1921) sense: quantifiable probability distributions; and to Knightian uncertainty (Knight, 1921): the existence of events whose probabilities are not quantifiable. Knightian uncertainty can also refer to residual uncertainty in the subjection probability distributions: the projector may not be quite sure of the forms of these distributions. In all cases, the future population, against which the projection is compared, is assumed to be revealed with certainty. This is a very strong assumption, as it assumes that uncertainty in knowledge of population ends at a known point. This equivalent, say, to assuming that census errors cease at a known point in time.[2]

Section 2 briefly introduces the use of loss functions to measure the accuracy of cross-sectional projections. This Section begins by assuming the presence of an impartial decision-maker who has preferences over outcomes. Since this decision-maker is unlikely to exist, Webster's rule is proposed, as it possesses several desirable properties. (Coleman, 2002)

Section 3 holds the base populations constant and selects optimal projected growth ratios. Both Knightian risk and uncertainty in the growth ratios are considered. Section 4 likewise holds the growth ratio constant and considers the effects of uncertain base populations on the selection of the optimal growth ratios. This Section finds that base population uncertainty is equivalent to growth ratio uncertainty in that they both produce the same optimization problem. An important upshot is that the optimal projected growth ratio generally differs from the a priori known growth ratio as a result of solving an optimization problem. Section 5 allows both growth ratios and base populations

to be stochastic. Both Sections 3 and 4 are useful as reference points from which to study the interaction of uncertainties in Section 5. Section 5's main conclusion is that this interaction generally complicates the solutions for the optimal projected growth ratios.

Section 6 concludes this paper.

2. Loss Functions[3]

A loss function measures the "badness" of the departure of a projection from its actual value. The total loss function for a set of projections is

$$\mathcal{L} = \sum_{i=1}^{n} L(P_i; A_i) \equiv \sum_{i=1}^{n} \ell(\varepsilon_i, A_i) \qquad (1)$$

where i indexes the n areas projected, P_i and A_i are the projected and actual populations for area i, $\varepsilon_i = |P_i - A_i|$ is the absolute value of the projection error, and L and ℓ are the individual loss functions. Also, let $P_i = B_i g_i$, where B_i and g_i are area i's base population and projected growth ratio, respectively. In all cases, A_i and B_i are assumed positive, while g_i is assumed nonnegative. $\mathcal{L}$ is taken to be additive in order to satisfy the von Neumann-Morgenstern (1944) expected utility axioms. (Coleman, 2001a and 2001b) A total loss function which satisfies the von Neumann-Morgenstern axioms has the useful, if clumsily stated in this context, property that the loss associated with a gamble is equal to the probability-weighted sum of the losses.[4]

The individual loss functions are built by assuming an impartial decision-maker who has preferences over outcomes. The assumptions needed to create these functions are summarized below. For a fuller explanation, see Coleman (2000). Subscripts are dropped in this Section, as they are not needed. The development is also stated in terms of P, as only the projected values are needed for computing loss, and not the means by which they were calculated.

Assumption 1 (symmetry): $L(A + \varepsilon; A) = L(A - \varepsilon; A)$ for all $A > 0$.

Assumption 2 (monotonicity in error): $\partial \ell / \partial \varepsilon > 0$ for all

---

[3]This exposition follows Coleman (2000, Section 2).

[4]See von Neumann and Morgenstern (1944) for a statement of the axioms and the proof of this statement in terms of expected utility. Markowitz (1959, chap. 10) has an amended version of the von Neumann-Morgenstern axioms.

---

[1]Growth ratios are used for simplicity instead of rates.

[2]I am indebted to Dave Word for pointing this out.

$\varepsilon > 0.$

Assumption 3 (monotonicity in actual value):
$\partial\ell/\partial A < 0$ for all $A > 0$.

Assumption 1 is very strong, as it implies that the decision-maker is indifferent between positive and negative errors. Assumption 2 simply states that smaller errors are preferred to larger ones. Assumption 3 states that an error of a given magnitude in a small area is worse than the same error in a large area. This can be best understood using an example. Suppose the error is 500. This is a serious error when the true value is 1,000, but almost a rounding error when the true value is 1,000,000.

The simplest loss functions that satisfy Assumptions 1–3 and admit Property 1 below are:

$$L(P, A) = |P - A|^p \, A^q \qquad (2a)$$

and

$$\ell(\varepsilon, A) = \varepsilon^p A^q \qquad (2b)$$

where $\varepsilon, p > 0$ and $q < 0$.

Finally, several mathematical and statistical reasons exist to explain why absolute percentage errors decrease in the size of the area. To handle this, we assume Property 1:

Property 1: The loss function defined by equations (2a) and (2b) increases in A for any given absolute percentage error. This is assured whenever $q > -p$, or, equivalently, $p + q > 0$.

Coleman (2000, Subsection 2.1) has an example of evaluating population projections using loss functions.

3. Growth Ratio Uncertainty Only[5]

This situation was discussed by Coleman (2000). Initially, assume that the joint subjective probability (Savage, 1954) distribution of the actual values is given by the Lebesgue-Stieltjes-measurable probability density function $dF(A_1,...,A_n)$. That is, the subjective probabilities associated with the actual values obey the customary laws of probability. Thus, we are dealing with Knightian risk. The probabilities are subjective in that they exist only in the mind of the projector. The future is unknowable, but the projector can make an estimate of dF. This estimate itself is based on a von Neumann-Morgenstern utility function on lotteries on all real n-tuples $(A_1,...,A_n)$ (Anscombe and Aumann, 1963). The subjective expected total loss associated with a point forecast is the Lebesgue-Stieltjes integral

$$\mathbf{E}\mathcal{L} = \int_A \sum_{i=1}^n L(B_i g_i, A_i)\,dF \qquad (3)$$

where $\mathbf{A}$ is the set of all real n-tuples $(A_1,...,A_n)$.[6]

The objective of projection optimization with no base population uncertainty is to choose a set of point growth ratios $\mathbf{g}^* = \left(g_1^*, \ldots, g_n^*\right)$ to minimize $\mathbf{E}\mathcal{L}$, given dF.[7] Coleman (2000, Section 3) solves a one area example using the Webster's Rule loss function, that is, with $p = 2$ and $q = -1$ (Coleman, 2002). A is assumed to have a triangular distribution with minimum, mode and maximum, $\underline{A}, A_m$ and $\overline{A}$, respectively. This solution, stated for g^*, the one area growth ratio, is given for reference:[8]

$$g^* = \frac{(\overline{A} - \underline{A})(\overline{A} - A_m)(A_m - \underline{A})}{2B\left[\begin{array}{l}(A_m - \underline{A})\overline{A}\log\overline{A} \\ +(\underline{A} - \overline{A})A_m\log A_m \\ +(\overline{A} - A_m)\underline{A}\log\underline{A}\end{array}\right]}. \qquad (4)$$

Letting $\underline{A} = B\underline{g}$, $A_m = Bg_m$ and $\overline{A} = B\overline{g}$ obtains the equivalent expression

$$g^* = \frac{(\overline{g} - \underline{g})(\overline{g} - g_m)(g_m - \underline{g})}{2\left[\begin{array}{l}(g_m - \underline{g})\overline{g}\log\overline{g} \\ +(\underline{g} - \overline{g})g_m\log g_m \\ +(\overline{g} - g_m)\underline{g}\log\underline{g}\end{array}\right]}. \qquad (4a)$$

Now, assume that there exists Knightian uncertainty over the growth ratios. Several methods exist for handling Knightian uncertainty, of varying usefulness for different applications (Walley, 1999). The method used in this paper is Choquet capacities, which give rise to the Choquet integral (Choquet, 1953). At the heart of this method is the concept of nonadditive probability. That is, given two events X and Y,

$$\Pr(X) + \Pr(Y) \le \Pr(X \cup Y) + \Pr(X \cap Y). \qquad (5)$$

This is in contrast to the usual concept of Lebesgue-Stieltjes-measurable probability, in which the inequality in (5) is replaced by an equality. It should be noted that the probability of the entire event space remains 1. For any given event X and probability density function dF,

---

[6]For all infeasible g, $dF = 0$. These include all vectors with at least one impermissible value, such as a negative.

[7]Minimizing expected loss is equivalent to maximizing expected utility. (Coleman, 2001b) See Coleman (2000, p. 29, fn. 8) for a more detailed explanation.

[8]Equation (4) is a correction of Coleman (2000, eq. 6). Care is required in its computation if the differences are small relative to the input variables.

---

uncertainty aversion can be defined by

$$c(dF, X) = 1 - \Pr(X) - \Pr(X^c) \qquad (6)$$

where X^c is the complement of X in the event space. "This number measures the amount of probability 'lost' by the presence of uncertainty aversion."[9] The "lost" probability reflects both the projector's ignorance over future events and his aversion to bearing uncertainty.[10]

The simplest assumption is constant uncertainty aversion.[11] Letting c be the uncertainty aversion, the corresponding Choquet capacity is $dF_c = (1 - c)\, dF$. Using the Choquet integral, Dow and Werlang (1992, p. 202) show that $\mathbf{E}_c\mathcal{L}$, the expected total loss which incorporates uncertainty aversion c, is given by[12]

$$\mathbf{E}_c\mathcal{L} = c \sup_g \mathcal{L} + (1 - c)\mathbf{E}\mathcal{L}. \qquad (7)$$

The case $c = 0$ corresponds to complete certainty over dF and reduces $\mathbf{E}_c\mathcal{L}$ to $\mathbf{E}\mathcal{L}$. When $c = 1$, the projector has complete uncertainty aversion and sets his expected loss to be the maximum possible. In essence, his expected loss is his worst-case scenario. This scenario will be on the boundary of **g**. He will choose a point estimate which minimizes his maximum total loss. That is, he will exhibit maximin behavior.[13] This point is further explored in Subsection 3.1. Intermediate values reflect the projector's possession of incomplete information about the future. In this case, $\mathbf{E}_c\mathcal{L}$ is a weighted combination of $\mathbf{E}\mathcal{L}$ and the worst-case loss. Thus, the loss-minimizing projection is intermediate between the two polar cases.[14]

3.1 Maximin Behavior

This is best exemplified by a one area problem. Using the notation of Section 3, when $c = 1$, the choice problem becomes to choose g^* to minimize

$$\max_{[\underline{A}, \overline{A}]} \left[L(Bg^*, \underline{A}), L(Bg^*, \overline{A}) \right]. \qquad (8)$$

Given a loss function which obeys Assumption 1, Bg^* solves

---

[9]Dow and Werlang (1992, p.200).

[10]See Schmeidler (1989, p. 582) for a formal definition of uncertainty aversion.

[11]Constant uncertainty aversion is a convenient assumption, but is not necessarily satisfied in reality.

[12]In terms of Dow and Werlang (1995), this is really $-\mathbf{E}_c(-\mathcal{L})$. The difference is that Dow and Werlang, (1995) Example 4.7, is concerned with maximization, while this problem is one of minimization.

[13]Strictly speaking, the projector exhibits minimax behavior with regard to expected total loss. See Coleman (2000, p. 29, fn. 8) for a fuller explanation.

[14]These cases are explored in more depth in Coleman (2000, Subsections 4.1 and 4.2).

$$L(Bg^*, \underline{A}) = L(Bg^*, \overline{A}). \qquad (9)$$

This equation results because both $\underline{A}$ and $\overline{A}$ are worst-case scenarios. Divergence from equality increases the loss with regard to one of $\underline{A}$, $\overline{A}$; thereby increasing the maximum loss. For Webster's Rule, equation (9) is solved by the geometric mean of $\underline{A}$ and $\overline{A}$. That is,

$$Bg^* = \sqrt{\underline{A}\,\overline{A}} \qquad (10)$$

obtaining

$$g^* = \sqrt{\underline{A}\,\overline{A}} \Big/ B. \qquad (11)$$

This can be generalized to n areas.

4. Base Populations Uncertainty Only

This case is akin to doing a short-term population projection in that, in the short-run, the ratio of change of the population is approximately constant. Let dB' be the joint subjective probability distribution of the revised base populations B_i'. Then, equation (3) is replaced by

$$\mathbf{E}\mathcal{L} = \int_{\mathbf{B}'} \sum_{i=1}^{n} L(B_i g_i, B_i' g_i)\, dB' \qquad (12)$$

where **B**′ is the space of vectors of revised base populations. Again, **g*** is chosen to minimize $\mathbf{E}\mathcal{L}$. It is important to note that it is not necessarily true that **g*** is the (known) fixed growth ratio. This results from obtaining **g*** from an optimization problem. Thus, the presence of uncertainty in the base populations may cause the projector to use projected growth ratios other than the known ones. The effect of this uncertainty need not be great. In a one area example, let $B = 1$ and B' have a symmetric triangular distribution with mode (and expected value) 1 and minimum and maximum of 0.9 and 1.1. That is, the projector considers B to be an unbiased estimated of B' with a range of ± 10 per cent, a very wide range indeed. Then, from equation (4), g^* is less than g by 0.17 per cent. For example, if $g = 1.01$, then $g^* = 1.0083$, to four places. Similar results obtain for biased expectations. Other loss functions may produce more dramatic changes, but are not considered herein.

Another implication of equation (12) is that growth ratio and base population uncertainty are equivalent. This can be seen by substituting A_i for $B_i' g_i$ in equation (12) and noting that its distribution is dF. The result is equation (3).

The analysis of Knightian uncertainty is similar to Section 3. Given an uncertainty aversion c', not necessarily equal to c of Section 3,

$$\mathbf{E}_{c'}\mathcal{L} = c' \sup_{\mathbf{B}'} \mathcal{L} + (1 - c')\mathbf{E}\mathcal{L}. \qquad (13)$$

In the one area case, the maximin solution, in the first term in the addition in equation (13), is obtained by choosing g^* to solve

$$L(Bg^*, \underline{B}g) = L(Bg^*, \overline{B}g) \qquad (14)$$

where $\underline{B}$ and $\overline{B}$ are the lower and upper bounds of the revised base population. Equation (14) has the same logic as equation (9): divergence from equality increases the maximum loss. Letting $\underline{A} = \underline{B}g^*$ and $\overline{A} = \overline{B}g^*$, equation (14) becomes identical to equation (9) and, therefore, has the same solution. Thus, Knightian uncertainty of growth ratios and base population produces equivalent effects.

Using Webster's Rule and substituting for $\underline{A}$ and $\overline{A}$ obtains

$$g^* = g\sqrt{\underline{B}\,\overline{B}}\,\big/\,B\,. \qquad (15)$$

The effect of Knightian uncertainty in the base population is clear: the optimal growth ratio under complete uncertainty aversion is a positive multiple of the known growth ratio. The two are equal only if $B = \sqrt{\underline{B}\,\overline{B}}$.

5. Uncertainty in Both Growth Ratios and Base Populations

The risk-only case is represented by a double integral equivalent to equation (3):

$$\mathbf{E}\mathscr{L} = \int_{\mathbf{g}'} \int_{\mathbf{B}'} \sum_{i=1}^{n} L(B_i g_i, B'_i g'_i)\, d\mathbf{B}'\, d\mathbf{g}'\,. \qquad (16)$$

where $\mathbf{g}' = (g'_1, \dots g'_n)$ is vector of true growth ratios and $d\mathbf{g}'$ is their Lebesgue-Stieltjes-measurable subjective probability distribution, relative to the revised base populations. The order of integration may be reversed. Again, a $\mathbf{g}^*$ is chosen to minimize $\mathbf{E}\mathscr{L}$. Note that the B_i are constant, as they are the unrevised base populations.

Because the one-area solution to equation (16) for triangularly distributed growth ratios and revised base populations is so complicated, an example is done using uniform distributions. Let g' be uniformly distributed on $[\underline{g}, \overline{g}]$. Then, equation (4) has solution

$$g^* = \frac{B(\overline{g} - \underline{g})}{\log B\overline{g} - \log B\underline{g}}\,. \qquad (17)$$

Now, letting B be uniformly distributed on $[\underline{B}, \overline{B}]$ obtains a solution to equation (16):

$$g^* = \frac{(\overline{B} - \underline{B})(\overline{g} - \underline{g})}{B(\log \overline{B} - \log \underline{B})(\log \overline{g} - \log \underline{g})}\,. \qquad (18)$$

Equation (18) shows, unsurprisingly, that the double integration in equation (16) makes the solution for g^* more complicated.

This author does not know whether Knightian uncertainty can be incorporated via the Choquet integral. The problem lies in the double integral in equation (16). Choquet (1953) developed his integral in a univariate setting.

6. Conclusion

This paper has considered the problem of creating point projections of population in order to minimize their expected total loss when the base population is subject to revision. This exercise requires the strong assumption that the future populations are known with certainty. An important result is that the mere possibility of revision, as revealed in a nondegenerate, additive subjective probability distribution function, can cause the projector to use growth ratios other than ones known with certainty. The equation to be solved is the same as that for uncertain growth ratios with no possibility of base population revisions. However, an example shows that this effect may be minuscule. The combination of base population revisions and uncertain growth ratios leads to complicated expressions for the optimal growth ratio projection. A simple case using uniform distributions provides an illustration. The problem in which both the growth ratios and base populations have Knightian uncertainty appears intractable, due to the limitations of the Choquet integral.

While this paper has concerned itself with population projections, it applies to any kind of forecast of positive cross-sectional data, when the base data are subject to revision. Moreover, the methodology can be extended to nonpositive data, along the lines of Coleman and Bryan (2002).

7. Acknowledgements and Disclaimer

References

Anscombe, F.J. and Aumann, R.J., (1963) "A Definition of Subjective Probability," *The Annals of Mathematical Statistics*, **34**, 199-205.

Choquet, Gustave, (1953) "Theory of Capacities," *Annales de l'Institut Fourier*, **5**, 131-295.

Coleman, Charles D., (2000) "Evaluating and Optimizing Population Projections Using Loss Functions," *Federal Forecasters Conference 2000: Papers and Proceedings*, Washington: U.S Department of Education, Office of Educational Research and Improvement, 27-32.

Coleman, Charles D., (2001a) "Metrics for Assessing the Accuracy of Cross-Sectional Estimates and Forecasts," manuscript, U.S. Census Bureau. Previous version presented to the 1999 Meetings of the Southern Demographic Association, San Antonio, TX, October 1999.

Coleman, Charles D., (2001b) "Total Loss Functions for Assessing the Accuracy of Cross-Sectional Estimates and Forecasts," submitted to *SIAM Review.*

Coleman, Charles D., (2002) "Webster's Rule for Assessing the Accuracy of Cross-Sectional Estimates and Forecasts," manuscript, U.S. Census Bureau.

Coleman, Charles D. and Thomas Bryan, (2002) "Loss Functions for Detecting Outliers in Panel Data when the Data May Change Sign," manuscript, U.S. Census Bureau.

Dow, James and Werlang, Sérgio Ribeiro da Costa, (1992) "Uncertainty Aversion, Risk Aversion, and the Optimal Choice of Portfolio," *Econometrica*, **60**, 197-204.

Knight, Frank, (1921) *Risk, Uncertainty and Profit.* Boston: Houghton-Mifflin.

Markowitz, Harry M., (1959) *Portfolio Selection: Efficient Diversification of Investments*, New York: Wiley. Second edition, 1970: New Haven, CT: Yale Univesity Press. Third edition, 1991: Cambridge, MA: Blackwell.

von Neumann, John and Morgenstern, Oskar, (1944) *Theory of Games and Behavior*, Princeton University Press, Princeton. Second edition, 1947; third edition 1953: Princeton University Press, Princeton. Reprinted 1963 and 1967: John Wiley and Sons, New York.

Schmeidler, David, (1989) "Subjective Probability and Expected Utility without Additivity," *Econometrica*, **57**, 571-587.

Walley, Peter, (1999) "Towards a Unified Theory of Imprecise Probability," *Proceedings of the 1st International Symposium on Imprecise Probabilities and Their Applications*, Ghent, Belgium, June 29 - July 2, http://decsai.ugr.es/~smc/isipta99/proc/055.html.

This page intentionally left blank.

Forecasting the Number of Veterans and Veterans Health Care Services

Chair: Kathleen Sorensen, U.S. Department of Veterans Affairs

The Veterans Actuarial Model (VAM2001)

Peter J. Ahn and M. Floyd Watson, U.S. Department of Veterans Affairs

The Veterans Actuarial Model is the veteran population projection model developed by the Office of the Actuary. It provides information that will assist VA staff to estimate future program needs. It has been expanded to estimate and project the number of veterans by age, gender, period of service, disability status, degree of disability, officer/enlisted, living/deceased, branch of service , dependents number, and marital status. Including these variables will enhance the effectiveness, accuracy, and efficiency of the VAM2001 to estimate future benefit costs, workload, and utilization of veterans' benefits. Examples using Excel PivotTable capabilities will be presented along with descriptions of veteran mortality and disability.

Forecasting Veterans' Disability Workload Received and Timeliness Performance

J. Reyes-Maggio, Veterans Benefits Administration, U.S. Department of Veterans Affairs

Forecasting the volume of veterans' disability claims received each year evolved considerably in the '90s. Likewise, VBA also re-evaluated its approach to assessing the impact of workload received as well as other legislative and process changes on its performance forecasting, i.e, how timely are veterans' disability claims processed. Variables such as the number of military personnel separating each year, the age and gender of these separatees, and the claim rate for these veterans are now used in projecting workload received. Forecasting timeliness of claims considers variables such as attrition of the workforce, trainee effectiveness, and hours employees actually work per day. VBA disability claims experts, with the assistance of consultants familiar with higher level applied mathematics, have captured the dynamic interaction of these variables and integrated them into simulation and worksheet tools that provide reliable estimates.

When Curiosity Killed the Statistical Trend, And Maybe The Cat, But Preserved the Military Veteran

Steve Pody, National Cemetery Administration, U.S. Department of Veterans Affairs

From the whirlwind of death and sacrifice at the hands of international terrorists on September 11, 2001, there came to be documented an unexpectedly upbeat statistical anomaly; a sort of cataclysm dividend. This incredible event of horrific destruction seemingly initiated a major deviation from statistical norm, subsequently causing, in grand irony, the prolongation of life for some one thousand U.S. citizens. The statistical data prompting this unusual analytical conclusion will be discussed in this presentation.

Continued

Long Term Care Model

Dan Culver, Veterans Health Administration, U.S. Department of Veterans Affairs

This paper describes the VHA's Long Term Care (LTC) Model. Used for program planning, strategic planning, and budget planning, the LTC Model projects demand for nursing home (NH) and home health (HH) care in terms of NH average daily census and annual HH care patients for a projected enrolled population. Projections may be generated for any VA facility, network, or nation for any target year between 2000 and 2010 based on the age (21-64, 65-74, 75-84, 85+) and disability level of the VA enrollees associated with the facility by enrollment priority using national (non-VA) surveys as the underlying utilization standard incorporating policy variables to "scale" the results to fit within budget constraints.

Uninsured Veterans and the Veterans Health Administration Enrollment System

Donald Stockford, Mary E. (Beth) Martindale, and Gregg A. Pane
Veterans Health Administration, U.S. Department of Veterans Affairs

The VHA Enrollment System is part of the ongoing restructuring of VA health care, and over 6 million veterans are now VHA-enrolled but, as with veterans in the veteran population, many enrolled veterans are uninsured. Historical trends in the uninsured may not predict the future very well for many reasons. Although there is economic uncertainty in the U.S., there is also new re-visioning of the very definition of the term "uninsured," reflecting a high-level policy change. This paper examines historical and current data on the uninsured and looks ahead to 2003 to see that VA will remain a "safety net" provider for veterans with health insurance coverage problems.

The Department of Veterans Affairs Health Care Enrollment Projections

Mary E. (Beth) Martindale, Randall J. Remmel, and Gregg A. Pane
Veterans Health Administration, U.S. Department of Veterans Affairs

The passage of the Veterans Health Care Reform Act of 1996 changed the way the Department of Veterans Affairs delivers health care to veterans. To receive care under this law, most veterans must be enrolled. This paper describes the Department's projections of health care enrollment, utilization, and expenditures that are utilized each year for important policy decisions such as who the Department can continue to serve under its appropriation and other resources. An example of the impact of external and internal changes in policies and other factors upon projections of one type of VA service, prescription drugs, is explored.

When Curiosity Killed The Statistical Trend, And Maybe The Cat; But Preserved The Military Veteran: A Brief Survey of Side-Effect Phenomenon Occurring During A Time of National Trauma, And The Subsequent Impact of Statistical Projection

Steven Pody, Department of Veterans Affairs

The National Cemetery Administration and, of course, the entire Department of Veterans Affairs, have always been subject to statistical fluctuation of its customer base, the projections of which are vulnerable to the irregular tides of peace and war. Major conflict or military call-ups may occur wherein millions of veterans are introduced into a system which a few years prior had no inkling that so many "customers" would be in existence. Think of the statistician projecting 30-year figures during the first administration of President Woodrow Wilson (1912-1916). Within that 30-year period were launched two unprecedented world wars, generating for the United States of America over 340,000 battle deaths, and almost 21-million veterans - plus attendant spouses and dependents. Who could know?

But of course not every conflict is major, and not every violent happenstance produces any greater a military force then that standard normally maintained by the nation. Many variations of action and response are possible as history compiles its daily record. And so it has come to pass, through recent circumstances, that historical evolution has formulated yet another variation. An incident of major non-military violence has unexpectedly produced an impact upon a military-oriented organization, shaking up its large body of experience and precedent-rich statistical prognostications. Through a great cataclysm, mass destruction has combined with mass media to produce a new and unpredicted effect, which has acted upon an organization normally quite used to events creating high yields of fatal carnage.

Specifically, the event alluded to was the terrorist assault on America on September 11, 2001. The organization affected in such a secondary, but pronounced manner by this event was the National Cemetery Administration. However… the primary subject within this famous context is not that of death. Ironically, within this dark and bitter moment has arisen surprising evidence of a seeming statistical tribute to some fundamental, stubborn, and possibly noble facets of human nature. Herein lies a paean to life, and to a dedication to purpose inherent within the living. It is a subject for which no statistic may account, nor be employed to foresee.

In September of 2001, the National Cemetery Administration, or NCA, was projected to host 7,000 burials within its system of national cemeteries. The average interment rate, the number of burials per month, had been 7,165 for the prior 11 months within the fiscal year.

But September 2001 was not to be an average month, or even a reasonably predictable month of any kind. This was the month of the infamous terrorist attacks in New York City and in Washington, D.C. While daily records are not available to highlight intra-month activity, specifically September 1st to 10th versus the period of September 11th to 30th, the month as a whole fell short of projection by 996: A 14.2% plunge (to 6,004) from statistical expectation. This one anomalous month alone, the last month of the fiscal year, not only concluded far from estimate, but threw the entire year's projection off target by a considerable margin.

Was this trough, this shortfall statistical hiccup, a mutual coincidence within a tragic month, or was it truly related to events of the month? With a remaining World War II population of some 5 million veterans at a median age of 78, along with 3.5 million Korean Conflict veterans only a few years younger, was there really the possibility that even a tremendous news event could impede the tide of nature and natural mortality for the better part of a month?

To give an idea of the magnitude of this subject deviation, and defining this month as more than just the occasional anomaly, several factors need to be considered. First of all, September, statistically, is already a month of low interment expectation. Only three other months are expected annually to have a lower yield of interments. Therefore, the projection, the planned figure, was reasonably set at a low number (relative to the rest of its companion year) by virtue of long prior-year precedent. In such months, in this age when interments have been consistently increasing each year since the early 1980's, the occasional surprise is generally in achieving even greater figures than a normal progression would call for. And yet, September 2001 yielded for the NCA considerably less "business" than planned.

In other words, not only were interments under projection, calculated as they were for normal conditions, but were uncommonly low in the face of an incident of uncommonly high death. National mortality was suddenly augmented by 3,000+ deaths, many of whom may be expected to have been military veterans. In fact, open national cemeteries lay within close proximity of each target city, and were ready to make emergency accommodations.

Historically, within regular NCA projections, acceptable-range interment-rate anomalies do occur - with actuals occasionally missing projection by four, five or even six hundred in a month. However, such activity is always in the unstable winter months (December to March) wherein weather, interacting with old and weary bones, can become a very capricious, powerful and deadly variable. The final third of the Federal fiscal year, June to September, is a time of generally friendly late spring and summer weather, reliably unfolding each year in a very stable way, statistically speaking. With little variance from month to month within this seasonal period, there is again no precedent for any wild departure from the norm.

So… How much less was this particular September figure from the norm that should make it so remarkable? As previously stated, each year - through almost two decades, has shown continually increasing interments. In fiscal year 2001, this trend produced record high months in October, November, January, February, March, April, June, July and August. Did September merely look less significant in a relative sense when compared to this large collection of record months? An already statistically unremarkable month coupled, coincidentally, with a statistical trough could highlight that month as unusually low. So what factors exist to qualify as advancing beyond the unusual and into the realm of the remarkable?

As it turned out, September of 2001 recorded the lowest September interment rate since 1997 - a four year old figure which, incidentally, 2001 only managed to just surpass by 29 interments. Special note should be taken that the NCA world of 1997 was more than just a time a few years removed from the present. There have been many changes, including considerable construction in the intervening years. In 1997, there were five fewer national cemeteries in the National Cemetery Administration. Since that 1997 figure was recorded, national cemeteries have been built to serve the major metropolitan areas of Albany, Chicago, Cleveland, Dallas/Ft. Worth, and Seattle/Tacoma. In NCA facility terms, those 5 new national cemeteries represented an 8% increase in the category of fully open (all grave-space types available) national cemeteries. The new cemeteries tabulated huge population advances in service benefit, adding millions of veterans to NCA-creditable service areas since 1997.

Furthermore, columbaria - memorial walls in which cremated remains may be placed, increased in number from seven to seventeen within the NCA, widening the availability of this option for large segments of the ever-increasing segment of the population who favor this burial preference.

In sum, all of the tens of millions of dollars of construction effort and the newly acquired millions in service-area population gains apparently ceased to exist - at least in gross interment numbers, for this extraordinary month of September 2001.

September 2001: The month when terrorists ignited the earth and sky, and murdered some 3,000+ people. Has anyone forwarded the insane postulation that such events were so horrible and so gut-wrenchingly dynamic and so perversely interesting that people refused to die just so they could see what would happen? Curiosity is indeed a powerful motivating force.

And nobler elements within human nature are also worthy of serious consideration. Is it not possible that this particular class of statistical sampling - that we in the NCA call military veterans (and their spouses and dependents), felt once again, strongly and patriotically, the call of their country? Did their country once again need them – their stability, their experience? A nation in alarm, in danger: September 2001 exhibited mass examples of patriotism, with the American flag raising and waving everywhere. It is not known whether such observations were highlighted at the time, but possibly the same effect occurred in December of 1941, when a similarly costly and shocking national challenge took place. And perhaps last September the old call to service was once again sounding, and nature, an inexorable pull of gravity and time, was resisted that that call might yet again be answered.

Upon consideration, it seems unlikely during such a moment of very much heightened national patriotism that the projected people populating NCA's missing thousand truly did expire - but did so only after alternately choosing to be buried at a place other than a national cemetery.

Be such as it may, through the range of bold philosophies and red-white-and-blue suppositions, a fine set of annual and September-month calculations found themselves in the trash …a solid trend lost in the inscrutable motives of those 1,000 shortfall numbers. These statistics are displayed graphically below. The initial graph presented shows that touted June-to-September time of relative stability, as represented by the last ten years. The progressive tracks, when followed individually below, exhibit fairly small variance as they make their way through the final third of their respective fiscal years. September, 2001, however, fairly plunges through the graph, atypical of its month category, seasonal pattern, and ultimately, year-long summary.
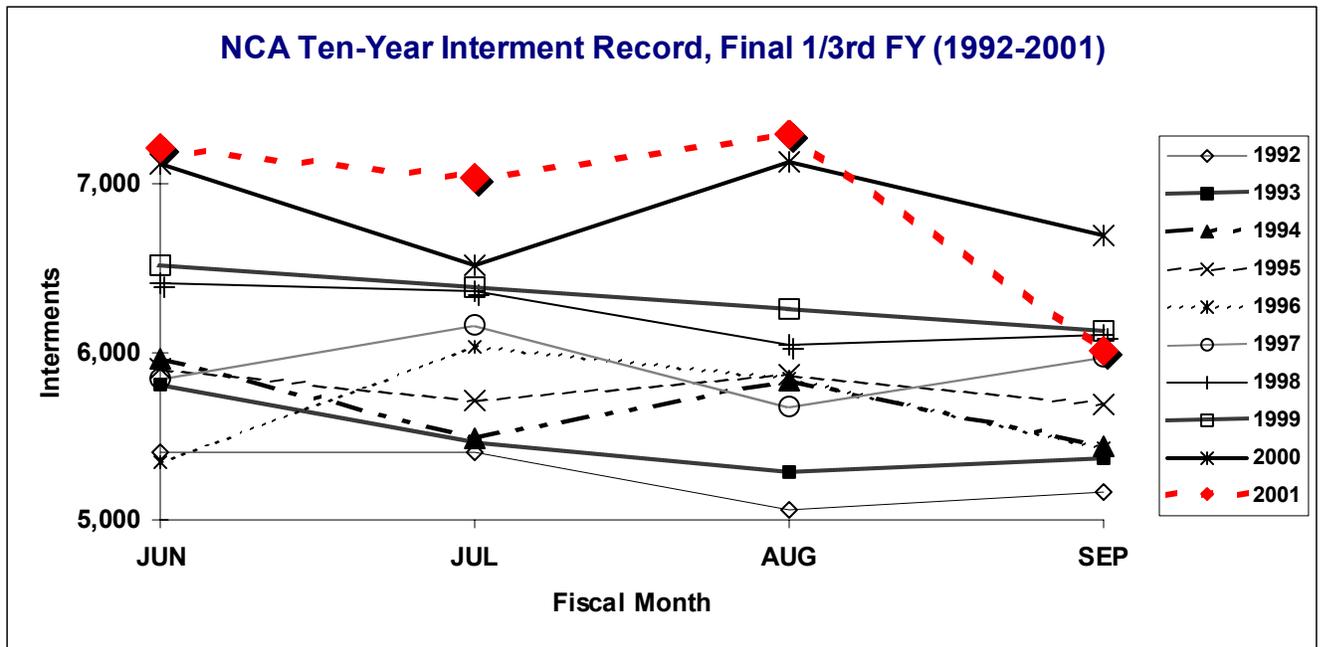
Visual clarity may be enhanced through the following 13-month chart of NCA interment activity. Vouching for a wide body of NCA statistics, the September trough, as displayed in Graph Two is indeed a one-of-a-kind statistical occurrence. Again, the strength of statements concerning trend validity rest both within the 10-year sample, and in light of organizational evolution which has seen the addition of millions of veterans to the service-benefit system.

NCA Interments, Calendar Year 2001
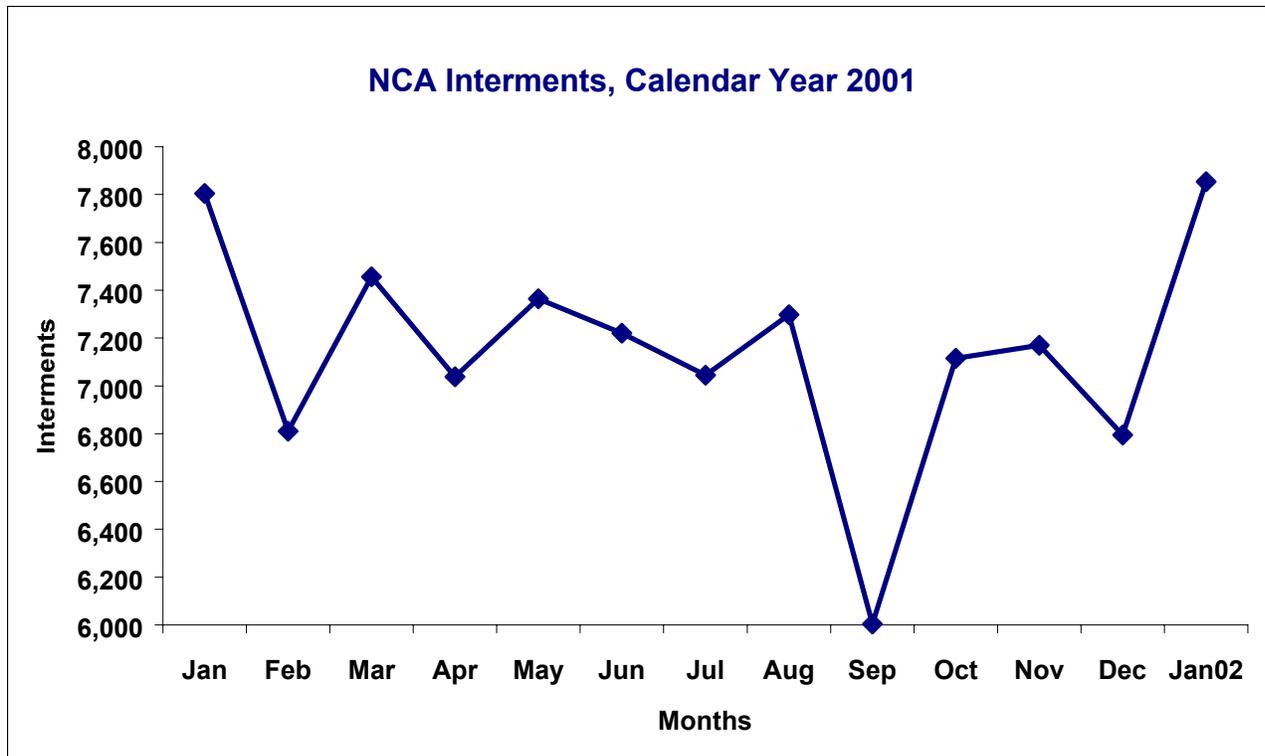
In this second graph, calendar-year interments for the NCA in 2001 are shown, each month relative to the whole. Two-thirds of the months occupy a narrow range within 400 interments of each other. As may be seen, if September had followed the pattern, registering traditionally within that narrow-window range, the stability normally inherent in the late spring to early autumn would be evident. After September, numbers generally resumed their precedent-established patterns. October, November, and January of the new year 2002, all continued to support long-running trends of not only high interment numbers, but record high interments for their months: continuing that reliable 18-year progressive death-rate pattern.

Reviewing Graph Two, an inventory of graph troughs from January through July would reveal nothing significant. February is always a trough, having two or three less production days available than regular months. The April and July troughs are mild enough to be attributable to favorable weather …or even that gently capricious variation in the life-clocks of a very large population of veterans (plus spouses and dependents), which is possible at any time.

However, December of 2001 has some further interesting qualities, which may possibly relate to the terrorist attack of the previous September. This particular December (2001) was the first December in eleven years (when data began to be published within monthly reports), which registered less interments than its prior companion month of November. Generally, national weather becomes more severe and challenging to health as late autumn progresses into early winter. Historically, as far as the statistical record runs for an organization containing numerous northern installations experiencing true winter, December registers more interments than November, and January registers more interments than December.

The December graph trough is far less severe than September's, but it may be seen to concurrently exist across the various cemetery-source graphs presented within this paper. However many non-NCA installations are cited though, this December phenomenon can only be accredited as a documented and unique event for the NCA alone. Long-running and corroborating data for the other source groups is not available. Nonetheless, a brief attempt will be made to advance and explain possible causes for this December interment trough as experienced in the NCA, and in possible conjunction with other examined industry partners.

Specifically, the month contained several possible motivating reasons, the impact of which could have translated as a lesser interment rate from an expected norm. In December of 2001, after a month of fact-gathering anticipation through September and half of October, plus a month-and-a-half of spectacular military retribution through mid-October and November, the conflict in Afghanistan came to a decisive first-phase conclusion. The terrorist and terror-supporting organizations in control of Afghanistan, the originators of the attack on New York City and Washington were, dramatically and publicly, ousted from national power. Further, within the month was hosted the sentimental and widely publicized 90-day commemoration of the September 11th attack. This was a major milestone of national reflection and healing.

Additionally, the parallel anthrax scare from mid-October settled from wild fear of mass annihilation to a condition of caution and watchfulness. Capping major events, and offering as great a measure of closure on this issue as America is likely to get, December was a watershed for those observing mass death and those who sacrificed in the whirlwind of mid-September. December saw a close of the retribution that followed, and also of the threat of deadly disease hanging over the nation. This month was not the beginning of the end; war would continue …but perhaps this was a closure of the beginning. December registered upon our NCA tabulations in the fashion of an aftershock of an earthquake, a final sigh before resumption of regular patterns of normal reality.

As mentioned, other organizations contributed data to this study, and their graphs may be seen below. These organizations include Arlington National Cemetery, a State veterans cemetery in New Jersey, and a collection of five Catholic cemeteries reporting as one group in the St. Paul/Minneapolis, Minnesota region. The NCA, with 120 cemeteries, forms a large sampling unto itself, but other input is advanced to broaden the base of this paper's contention. Multi-cemeterial organizations, which both compile and analyze statistical data, are not common. But, if possible, it was felt that a couple of non-NCA military cemeteries might lend a wider perspective, and a civilian-oriented multi-cemetery group was also found to expand the sampling field.

It may not, in the long run, be possible to corroborate our set of results with other segments of the cemetery industry. The NCA is a large organization, perhaps unique within its field for analytical potential. The NCA oversees over a hundred cemeteries, all operating under the same information-gathering scenario and serving a select segment of the national population.

For the large, disparate category of State veteran cemeteries there is little data uniformity. There exist annual, but no monthly tabulations summed amongst the group of State veteran cemeteries collaborating with the NCA. Likewise, there are no known umbrella organizations within the non-military cemetery community tracking burials on a monthly basis. It is unknown if national trends within the private cemetery business recorded similar patterns, except for the one sampling given in this report. Perhaps, or perhaps not... (Therefore, in the uncertainty of outlying data it might be presumptuous to assert that the military veteran alone constitutes a distinctly reacting sub-population to September 11. And yet, the only broad-sample, long-term documentation of a population subset reacting to events rests squarely with this group.)

Each graph is a valid representation relative and relevant within its own 13-month statistical rendition but, as will be seen, scale cannot be retained throughout all samples. Single installations and larger organizations process differing magnitudes of numbers. Interestingly, though a wide variety of single-year patterns are evident within these additional organizations, the general trend of a large September drop in interments, and a minor one in December (relative to the preceding October/November) may be observed to some degree in each case.

Arlington National Cemetery, a national cemetery belonging to the Department of the Army rather than the NCA within the Department of Veterans Affairs, displays a roughly similar interment pattern to the NCA model. Statistics for this military-oriented national shrine show a major shortfall in pattern for September, and December slumps after the month of November.

Arlington National Cemetery, Calendar Year 2001
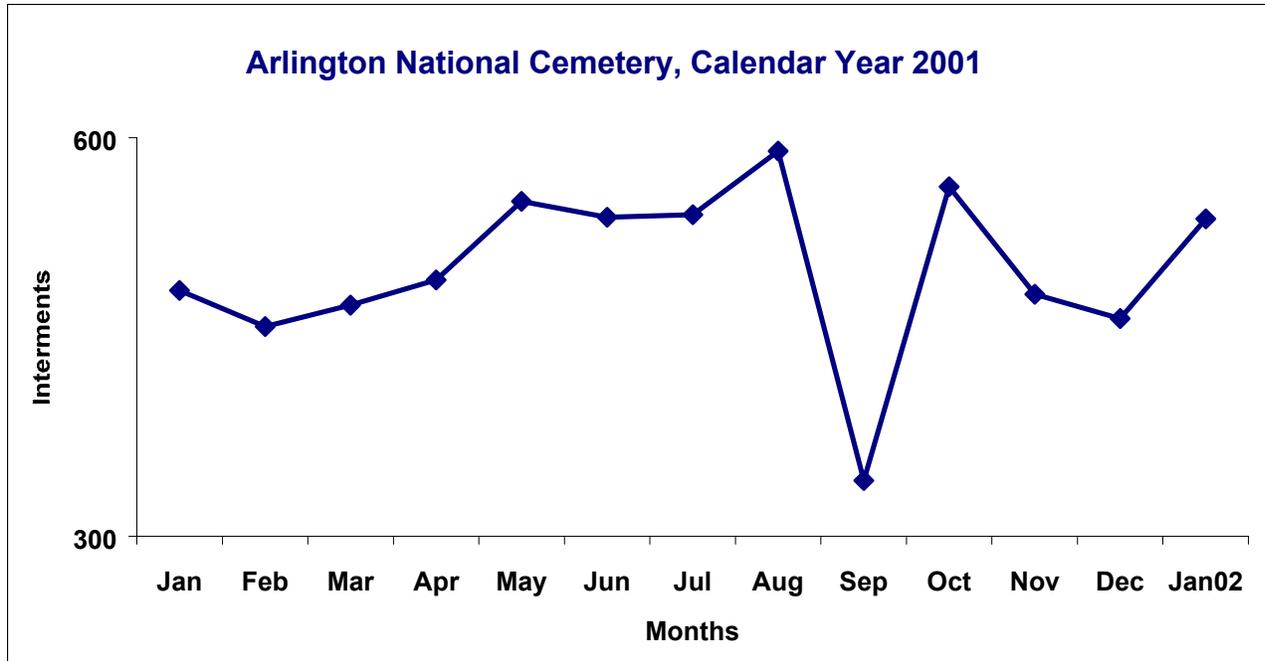
Single installations, like Arlington National Cemetery, do differ from national organizations in that such sites reflect more limited local factors. The geographically widespread system of the NCA tends to average-out the fluctuating impact of seasonal variation. For Arlington, if the local Virginia winter wasn't too harsh, winter interment numbers might display a weak showing, relatively, within the annual run of months.

Brigadier General William C. Doyle Veterans Memorial Cemetery, a State veterans cemetery in New Jersey, varies slightly in the presented scheme. September is again represented by a sharp plunge in figures, relative to its neighbor months. November, however, registered one less interment than December, though both months, like Arlington, form a collective statistical dip. If not coincidental, the reason for both November and December to form this mutual trough may lie in the proximity of these two cemeteries to the terrorist strike points near their respective locales – Brig Gen Doyle to New York City, and Arlington to Washington, D.C.

Again, these are single installations, and have no averaging affect as might be expected in a national organization. What registers on the graph happened locally. For Brig Gen Doyle and Arlington, September was the big trough, the month of shock. October became the catch-up month as nature reasserted itself. But this reassertion masked the continuing hold earthly events still held upon the mortal populace; so November/December registered a dual repeat of depressed interment figures as world events continued to dramatically sort themselves out. Populations, which would likely patronize these two cemeteries, would be the same as those living and working in the shadow of the attacks, with major commitment to unfolding retribution.

**Brigadier General William C. Doyle State Veterans Cemetery
Calendar Year 2001**
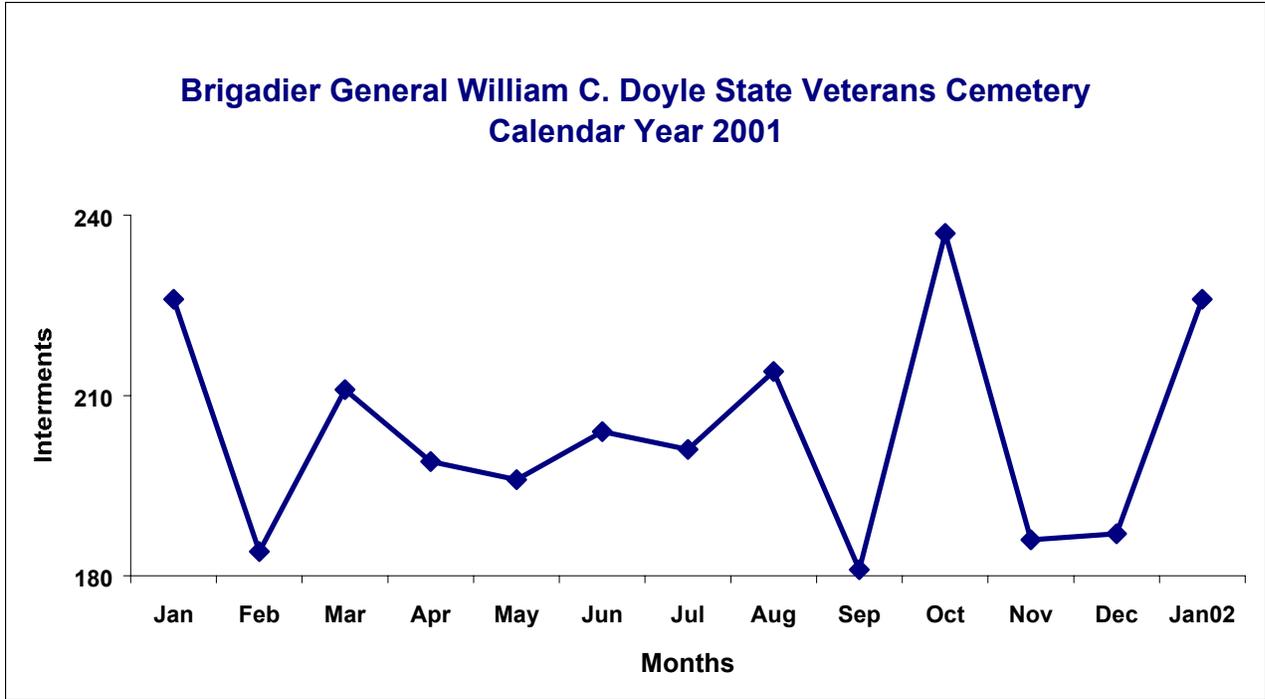
Reporting from the NCA, the author can most credibly give a veteran-interment slant on potential trends: Hows and Whys, with long years of numerical documentation and trend precedent. Nonetheless, in an attempt to be more inclusive, a query was made to a sampling cluster of private cemeteries. The five cemeteries that make up the Catholic cemetery system of the Archdiocese of St. Paul and Minneapolis kindly submitted their collective interment data for inclusion within this study. Minnesota has severe winters. Not only do low temperatures generally prevail, but conditions of deep snow exist. Under such circumstances, making holes in the ground is difficult, and a heavy burden is imposed upon a likely frail audience. Therefore burials, per the graph (and like many northern burials), are often delayed until spring.

Archdiocese of St. Paul and Minneapolis, Calendar Year 2001
(5 cemeteries)
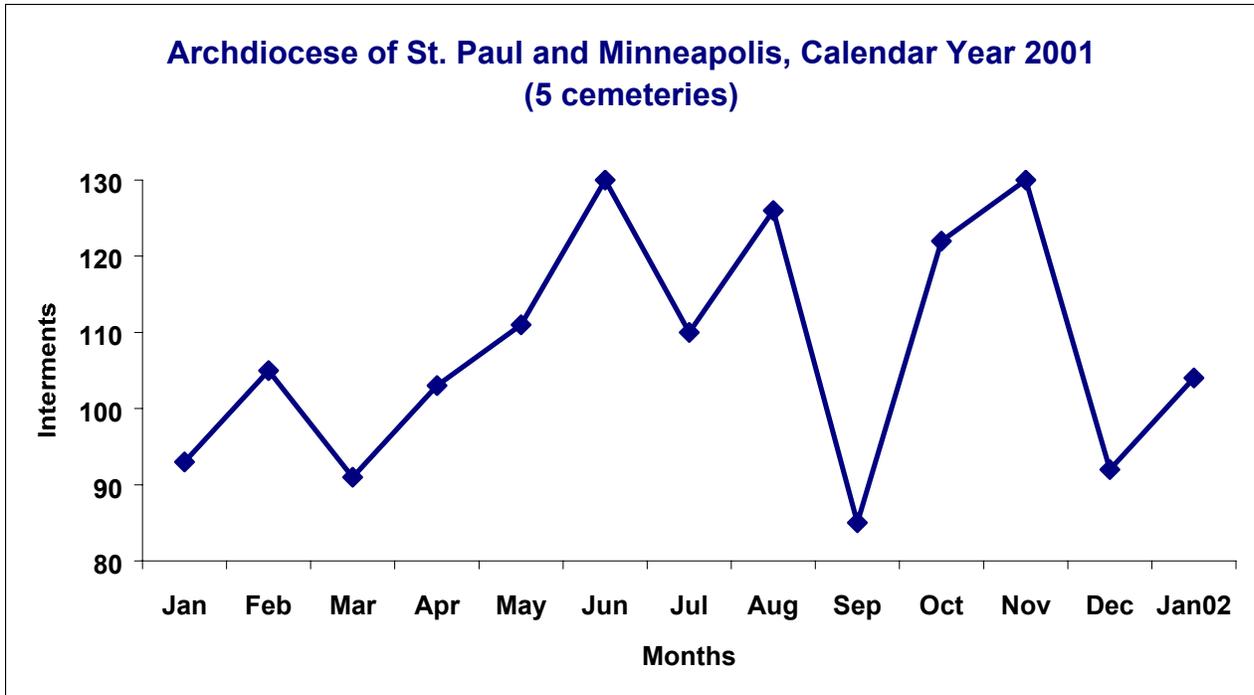
In all honesty it cannot be given as definite assurance that the pattern seen for the Archdiocese, or even the two military cemeteries, is, or is not, a normal and regularly recurring annual trend. Report years, where reports are compiled, vary along calendar and multi-choice fiscal year lines. Long-compiled monthly records are often not available, and what can and is displayed for this study is one unified, snapshot 13-month record. For non-NCA cemeteries given in example herein, September interments graphically dip, it is true. But did the pattern dip in the year 2000? Do numbers for these installations always trough in September? Unlikely, but not known for certain. The pattern gives every indication of widespread uniformity, but… When all is summed, it can only be said for certain that, for the NCA, the September 2001, statistical trough was remarkably contrary to long-prior NCA trend.

And yet, if this present study were to accept that the above displayed non-NCA-military and civilian graph patterns were a true event deviation, and genuinely representational of at least part of their respective categories, then the entire subject may be advanced to a conclusion of credible probability. Encompassing the myriad American cemeteries, there exists the potential that, if true for some, then true for others. Therefore, extrapolating out to the nation, many, many thousands of deaths may have been delayed by that single major September disruption.

Additionally, if valid - and the likelihood is great, not only veterans would have heard the call of their nation and steeled themselves, even against death, to offer support. …And not only veterans would have been curious enough to remain within the fragile veil of life to find out how this tremendously significant threat to their nation might turn out.

So, in sum, what is seen on these graphs, like a heart electrocardiogram, is the pulse, the shock of a nation. The import of such a collective graph record could not fail to show people being people, reacting as a nation together, whether individually once wearing a military uniform or not. Initially, the pattern was perhaps most easily seen in military cemeteries where collective records are kept and are convenient for large-scale analysis. What the pattern revealed was a genuine reaction to a tremendous event, within an unlikely subject area - burials of the dead. But sampling of other systems would indicate that the September pattern, while unique on a nationwide basis, was not necessarily unique to any given population sub-group within that nation.

Let us freeze and examine for a moment that September juxtaposition between a national tragedy and the National Cemetery Administration. Ultimately, it may be cause to wonder that if there were 3,000 victims on that terrible day of September 11th, cannot a full 25% be sub-tracted out of the total, strictly on a statistical basis, for those veterans

and their spouses and their dependents - those 1,000 who continued to live when statistics pointed otherwise? Might some psychological credit be gained to help serve as a partial balm to horrific events? Within the context of the same enormous tragedy, we are given the paradox wherein 3,000 people died who should not have died, and yet one thousand lived who should otherwise have died.

Thus, it appears, was presented to national and to NCA history a moment which stayed the natural termination of one thousand nameless numbers of a statistical September of a statistical NCA fiscal year. This event became statistical discontinuity personified. Here was a twenty-September-day-long moment when, because of great tragedy, some personal grief was ironically stayed; when thousands who might have wept did not weep, and when statistical projection in our NCA world of interments and final bugle calls took a major shift - and a curiously joyous detour, into the land of the X-factor.

Acknowledgements

For information selflessly shared, and pleasantly and professionally communicated, thanks are given to Mr. John C. Metzler, Jr., Director of Arlington National Cemetery, VA, Mr. Stephen Abel, Director of Brigadier General William C. Doyle Memorial Veterans Cemetery, NJ, and Mr. Paul Letourneau, Accounting Manager of the St. Paul/Minneapolis, MN, organization of Catholic cemeteries.

Disclaimer

While an employee of the National Cemetery Administration and the Department of Veterans Affairs, the author wishes to stress that this report is an official document of neither organization. Rather, it is a compilation and presentation of data, conclusions and potentialities formulated by the author through observation and while in the performance of NCA-related duties. The views and opinions expressed within this document are not necessarily those of the NCA or the VA.

This page intentionally left blank.

LONG TERM CARE MODEL

Dan Culver, Department of Veterans Affairs

The Long Term Care (LTC) Model projects demand for nursing home (NH) and home health (HH) care in terms of NH average daily census and annual HH care patients. It was developed by the Veterans Health Administration Office of Policy and Planning in consultation with subject matter experts from the Geriatrics and Extended Care Strategic Healthcare Group, the Agency for Health Care Policy and Research, and the University of Michigan.

1. **Attributes:** Projections may be generated
   - for any VA parent facility, Veterans Integrated Service Network (VISN), or nation
   - for any target year between 2002 – 2010
   - by Priority level (1a, 1b, 2, 3, 4, 5, 6, 7)
   - based on the projected enrollee population of the facility
   - based on the age (21-64, 65-74, 75-84, 85+) and disability level of VA enrollees
   - using national (non-VA) surveys as the underlying standard

2. **Underlying assumptions:**
   - Historical VA LTC data are not completely adequate for projecting future LTC requirements for reasons specified below.
   - NH and HH requirements depend on both the age and disability level of the population at risk.
   - The population at risk is the VA enrollee population.
   - Enrolled veterans will require NH and HH care at the same age- and disability-adjusted rates as do US males, as measured by national surveys.
   - The 1996 Medical Expenditure Panel Survey of NH is representative of national NH use.
   - The 1998 National Home and Hospice Care Survey is representative of national HH care use.
   - VA will be unable to provide LTC to all enrollees who request it, consequently, an affordable market share (MS) is required.

3. **Why non-VA surveys are used:**
   Since VA long term care budgets for both nursing home and home health care have been historically constrained, any projections based purely on VA experience would merely perpetuate the current situation. Furthermore, not only would a VA NH model be required, so would separate models or approaches for community and state NH. Finally, such models would only be able to project future workload for geographic areas with existing NH programs.

   For home health care, VA experience is even less indicative of true demand, since VA has historically referred the preponderance of HH demand to Medicare.

4. **Overview of the Model:** The LTC Model essentially calculates the product of three variables.
   - LTC Use Rates for males
   - Enrollee Population projected for each facility
   - Market Share percentage VA will provide

   This calculation occurs for every possible combination of age, disability level, and Priority level of the Enrollee Population.

5. **Data Sources for the three variables in the LTC Model:**

   a. **LTC Use Rates** for males, by age and disability level are derived from national, non-VA surveys.
      - for NH from the 1996 Medical Expenditure Panel Survey.
      - for HH care from the 1998 National Home and Hospice Care Survey.

b. **Enrollee Populations** are based on actuarial estimates distributed into disability levels based on the results of a telephone survey of 27,000 enrollees conducted in February 1999.

c. **Market Shares** are based on a combination of historical precedent, legislative mandate, LTC program policy, resource availability, and budgetary policy.

6. Definitions of Terms and Variables:

a. **NH Use Rate** $=$ $\dfrac{\text{male NH residents (by age and disability) on January 1, 1996}}{\text{male residents of the US (by age and disability) in 1996}}$

b. **HH Use Rate** $=$ $\dfrac{\text{male HH care recipients* (by age and disability) during 1998}}{\text{male residents of the US (by age and disability) in 1998}}$

* receiving one or more visits from any health care provider (physician, visiting nurse, nurse aide, therapist, home health aide, homemaker, or social worker)

c. **Enrollee Population Attributes:**

Priority Levels are as defined in Eligibility Reform legislation, except as noted below.
    1a = 70%+ Service Connected (SC)
    1b = 50-69% SC
    2
    3
    4
    5
    6
    7

Age groups
- 21-64
- 65-74
- 75-84
- 85+

Disability levels for both NH and HH programs are defined in terms of Activities of Daily Living (ADL) deficiencies. For the HH program only, disability levels are also defined in terms of Instrumental Activities of Daily Living (IADL) deficiencies.

The ADL deficiencies include difficulties in
- eating
- bathing
- dressing
- getting in and out of bed or chairs
- using or getting to toilet
- walking across a room

The ADL deficiency score represents the total number of these activities a patient or enrollee receives help doing.

The IADL deficiencies include difficulties in
- using the telephone
- managing money
- shopping for personal items
- getting around the community
- preparing meals
- doing light housework

The IADL deficiency score is only considered in the HH program and then only if the ADL score equals zero.

d. Historical Market Shares:

The VA NH Market Share (MS) is defined as follows:

ADC treated in VA, Community, and State NH under VA auspices
Total NH ADC estimated for enrollee population by the LTC Model

where the numerator and denominator both pertain to the same time period. The MS represents the percentage of the anticipated total enrollee demand provided by VA.

The overall nursing home MS for FY00 is computed as follows:

MS NH = FY00 ADC / estimated FY00 enrollee NH demand based on LTC Model
   = 31,090 / 146,000
   = 21%

Although the MS over all Priority levels was 21% in FY00, the NH MS for Priority 1a enrollees (Service Connected rating of 70% or more) was 35%, and for Other enrollees was 16%.

The numerator of the home health MS is the patient load treated in either VA HH programs or in contract programs paid for by VA. The denominator is the LTC Model projection for the corresponding year. In FY00, the HH MS was 10%, with no special MS provision for Priority 1a enrollees. HH patients VA cannot treat due to resource constraints are referred to Medicare.

e. Future Market Shares: Future MS do not necessarily have to equal the historical MS. In general, they require policy decisions based on a combination of historical market share precedent, legislative mandate, program policy, resource availability, and budgetary policy. VHA corporate and VISN level targets and rates are subject to policy and budget decisions. The Model is intended to be flexible and to accommodate VHA policy.

Nursing Home Care

- For NH care, the Millennium Bill mandates that VA provide needed NH care to any Priority 1a enrollee who requests it. However, since not all Priority 1a enrollees needing NH care will request it from VA, the Model incorporates a planned MS for Priority 1a enrollees that increases from 35% to 85% by 2008.
- The future MS for Other Priorities is discretionary. The relevant policy decisions have not yet been made.

- The Federal Advisory Committee on Long Term Care recommended tripling, from 8% MS to 24% MS, the VA expenditure on home and community-based care over five years, with FY99 as a baseline. This MS can be adjusted per VHA policy decisions.
- VHA budgetary policy is currently under consideration.

7. **Output of the Model:** The LTC Model generates projections of
   - Nursing home Average Daily Census (ADC)
   - Home Health care annual patients

required by enrolled veterans by age, disability level, and Priority; for any parent facility, VISN, or the nation; for any year between 2002 – 2010.

a. The ADC for NH care is a combined total of VA, community and state ADC.

b. The annual patient load for HH care represents the patient workload in the following programs:
   - Home Based Primary Care
   - Contract Home Health Care
   - VA Adult Day Health Care
   - Contract Adult Day Health Care
   - Homemaker/Home Health Aide

To permit maximum flexibility, any of the variables in the LTC Model can be adjusted for analytical or planning purposes. In addition, any combination of facilities or any combination of VISNs can be grouped together and treated as a single facility or VISN.

Nursing Home Component of LTC Model

$$NH_{fac} = \sum_{year} \sum_{age} \sum_{ADL} \sum_{Priority} UR_{age, ADL} \times Pop_{facility, age, ADL, Priority, year} \times Mkt\ Share_{Priority}$$

NH im   =   projected NH ADC for **fac**ility i in **year** m

UR jk   = $\dfrac{\text{males nursing home residents by age j and ADL level k in 1996 MEPS}}{\text{male resident population of US by age j and ADL level k in 1996}}$

Pop ijklm   =   enrollee pop of facility i, age j, ADL k, Priority l in year m

Age j   =   21-64, 65-74, 75-84, 85+

ADL k   =   0, 1, 2, 3, 4, 5, 6

Priority l   =   1a, 1b, 2, 3, 4, 5, 6, 7

Mkt Share l   =   for Priority 1a, rises to 85% by 2008; under debate for Priorites 1b - 7

Year m   =   2002 - 2010

Home Health Care Component of LTC Model

$$HH_{fac} = \sum_{Year} \sum_{Age} \sum_{ADL\ Priority} UR_{age\ ADL} \times Pop_{facility\ age\ ADL\ Priority\ Year} \times Mkt\ Share_{Priority}$$

HH im	=	projected HH patients at **fac**ility i in **year** m
UR jk	=	males in age j and ADL level k receiving HH visits in 1998 NHHCS
		male resident population of US by age j and ADL level k in 1998
Pop ijklm	=	enrollee pop of facility i, age j, ADL k, Priority l in year m
Age j	=	21-64, 65-74, 75-84, 85+
ADL k	=	0, 1, 2, 3, 4, 5, 6, or IADL Only
Priority l	=	1a, 1b, 2, 3, 4, 5, 6, 7
Mkt Share l	=	16%
Year m	=	2002 - 2010

Market Share Primer

Item 6d of the paper defines market share in the context of the LTC Model, but perhaps the simplest way to explain the definition of market share is with a diagram.

In the diagram below,

1. Box 1 represents a population.
2. Box 2 represents members of the population who "need" LTC, according to the Model.
3. Box 3 represents members needing LTC who actually receive LTC from VA.
4. Box 4 represents members needing LTC who presumably receive LTC from other sources.

Market share, as used in the LTC Model, equals the fraction formed with a numerator consisting of Box 3; and a denominator consisting of Box 2, with "members" replaced by "average daily census." This is the classical definition of market share.

A simple, real-world analogy may be illustrative: market shares of cola bottlers. One definition of Coke's market share is [ounces of Coke sold / ounces of all colas sold]. Pepsi's market share would have the same denominator, but a numerator equal to ounces of Pepsi sold. Other cola bottler market shares would be computed analogously. The sum of all numerators equals the denominator, accounting for 100% of the total market demand.

The LTC Model uses precisely the same methodology for market shares. In the context of cola market share, the VA market share numerator (NH average daily census "sold" by VA) corresponds to ounces of Coke sold. Similarly, the market share denominator (NH average daily census expected to be "sold" by all sources to the enrollee population) corresponds to total ounces of cola sold.
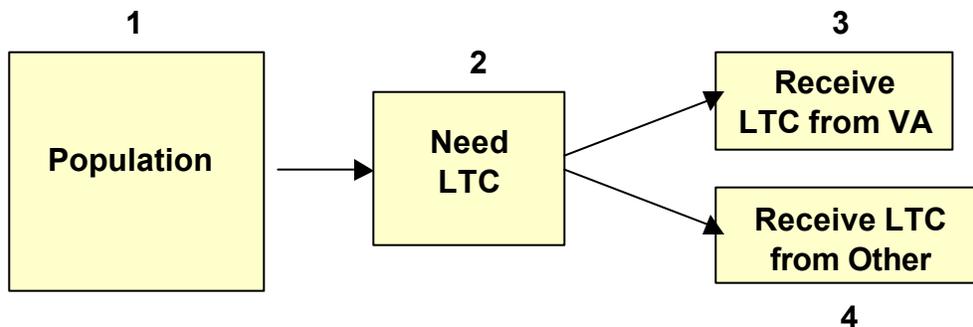
$$\text{VA Market Share} = \frac{\text{Box 3}}{\text{Box 2}}$$

This page intentionally left blank.

Uninsured Veterans and the Veterans Health Administration Enrollment System 2003

Donald Stockford, M.A., Mary E (Beth) Martindale, Dr.P.H.

Gregg A. Pane, M.D., M.P.A.

Department of Veterans Affairs

Introduction

The VHA Enrollment System is part of the ongoing restructuring of VA health care, and over 6 million veterans are now VHA-enrolled. However, as with veterans in the veteran population, many enrolled veterans are uninsured. Over the long term, the number of uninsured veterans in the veteran population has been decreasing and the number of uninsured enrollees in the VHA Enrollment System has been increasing. In this paper, linear regression techniques are used to look ahead to 2003 to try to gauge the impact of these competing trends on the historical role of VA as a "safety net" health care provider for veterans in need.

Part I: Some 38.7 million Americans (14.0% of all non-institutionalized persons in the U.S.), including 1.6 million veterans (6.3% of all non-institutionalized veterans), were without any health insurance coverage at all during all of calendar year (CY) 2000.

These latest data represent decreases from the equivalent uninsurance levels of the previous year, CY 1999, when some 39.3 million Americans (14.3% of all U.S. non-institutionalized) and 1.7 million veterans (6.6%

of all U.S. non-institutionalized) were uninsured during the entire year (**Chart 1**, **Chart 2**).

Chart 1. U.S. Population, Veteran Population Percent Uninsured (Adjusted) 1999 - 2000
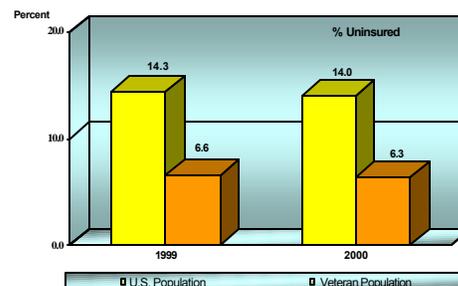
Sources: March 2000 and March 2001 Current Population Surveys; Veteran Population 2000 Model

These full-year uninsurance data (**Chart 1**, **Chart 2**) are the latest Bureau of the Census measures on uninsurance in the United States and are **adjusted** to reflect recent changes in the health insurance coverage probes[1, 2] in the Current Population Survey, which is Census' primary health insurance data collection instrument. Here is some clarification as to what is meant by **adjusted** and **unadjusted** rates: in March 2000 and March 2001, and in addition to the ordinary March CPS probes for health insurance coverage, Census included health

insurance verification probes in the CPS questionnaire, to re-probe people who said they had no coverage at all during the previous calendar year. Thus, uninsurance rates which are computed irrespective of the verification probes are called **unadjusted** rates, and uninsurance rates which are computed based on the regular CPS health insurance probes in combination with the verification probes are called **adjusted** rates. Also, since the verification probes are new to CPS as of March 2000 (CY 1999 coverage), the CY 1999 and CY 2000 health insurance coverage data reported here, begin **a new trend series of adjusted rates and numbers** that will be followed over time.

Note: Various health insurance data sources are used or referred to in this paper. The primary data source is the Current Population Survey (CPS), a monthly survey of civilian non-institutionalized persons in the U.S., conducted by Census for the Bureau of Labor Statistics. An additional data source is the 1999 Veterans Health Administration (VHA) Survey of Veterans Health and Reliance Upon VA. Another VHA survey, namely the 1999 VHA Office of Quality and Performance Large Survey of Enrollees was done about the same time as the smaller VHA OPP survey. Since results of the VHA OPP and VHA OQP surveys are largely in agreement and there is no dataset for the VHA OQP survey that is freely available for analysis, only the VHA OPP survey data are used this paper[3],[4]. The VHA OPP survey is discussed in more detail below.

Chart 2. U.S. Population, Veteran Population
Number Uninsured (Adjusted)
1999 - 2000
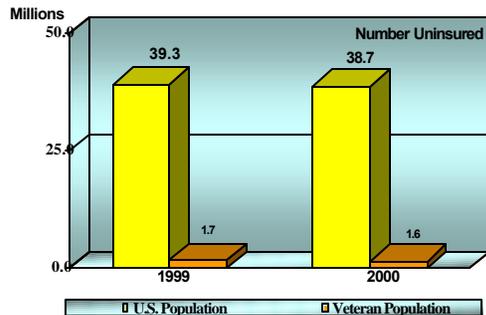
Sources: March 2000 and March 2001 Current Population Surveys;

Part II: U.S. population uninsurance rates have reversed a multi-year trend of increasing rates; veteran population uninsurance rates

continue a long-term decreasing trend; and both of these trends are expected to continue.

Adjusted data (**Chart 1**, **Chart 2**), reflecting a full-year of lack of health insurance coverage, as verified through the CPS verification probes, were used above to underscore the seriousness of health insurance coverage problems in the U.S., as well as for veterans and for VA. However, equivalent adjusted data for years earlier than CY 2000 do not exist, and so, for analyses of long term trends, it is necessary to examine historical **unadjusted** data (**Chart 3**).

Chart 3. U.S. Population, Veteran Population: Long-Term Trends in the Percent Uninsured (Unadjusted) CY 1987- 2000
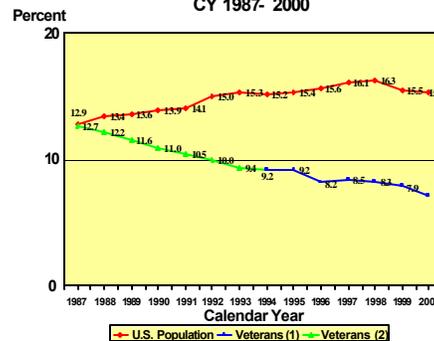
Sources: US. Population (March 1988-2001 CPS data); Veteran Population (March 1995-2001 CPS data, 1993 National Survey of Veterans (SOV-IV), 1987 National Survey of Veterans (SOV-III), and trending for CYs88-92). (Note: CPS data are CY; other data are point-in-time estimates and may be higher than equivalent full-year rates.)

According to the Current Population Survey, the **unadjusted** total number and percent of Americans who were without any health insurance coverage at all during the entire year dropped from 42.6 million (15.5%) in CY 1999 to 42.3 million (15.3%) in CY 2000. This represents the continuation of a reversal in trend that began one year earlier; from CY 1998 through CY 2000, uninsurance rates have been dropping, whereas, from CY 1987 through CY 1998, they increased.

On the other hand, the **unadjusted** total number of veterans who were without any health insurance coverage during the entire year continued a long term decline, from 2.0 million (7.9%) in CY 1999 to 1.8 million (7.1%) in CY 2000.

National full-year uninsurance trends and rates are lower for veterans than for the U.S. population for many reasons. In particular, the reported U.S. population rates include women

and children, while veteran rates reflect a population that is predominantly (95%) male. The overall U.S. rates, in particular, reflect: improvements in employer-based coverage during the expanding economy of the 1990's that were offset by declines in State-subsidized insurance such as Medicaid that had been occurring prior to 1996 Welfare Reform; further declines in Medicaid after 1996 Welfare Reform offset improvements in employer-based coverage even more; population growth also contributed; but the moderating of declines in State-subsidized coverage and the introduction of the State Health Insurance Program (SCHIP) in the late 1990's resulted in more children being covered. On the other hand, male veterans are about 16 years older[5, 6] on average than their adult male non-veteran counterparts. Veterans are also more likely to be in their peak earning years, have job training, vocational training, and a variety of other opportunities and resources, and compare favorably on health insurance and other socioeconomic measures[5, 6, 7] to their non-veteran counterparts. However, aging is the most profound factor affecting veteran trends and, as they age, more and more veterans, particularly "near elderly" veterans (i.e., veterans age 50 - 64 who may retire early and lose private coverage until age eligible for Medicare, etc.), are retiring and/or obtaining Medicare coverage, even if they were uninsured before.

The recent decline in U.S. population uninsurance rates and the long-term decline in veteran uninsurance rates shown in **Chart 3** are expected to continue,[8, 9, 10] **although U.S. population growth may cause the actual numbers of uninsured Americans to increase even as overall U.S. uninsurance rates decrease. The rest of this paper will focus on the veteran data and trends.**

Part III: The Veterans Health Administration (VHA) Enrollment System was mandated by Congress to help VA stay within its budget while it implemented major eligibility reforms to afford veterans a comprehensive package of services; enrollment participation by veterans is high and continues to grow.

The VHA Enrollment System was mandated by Congress (*Veterans Health Care Eligibility Reform Act of 1996*, P.L. 101-262) as a tool to manage the major eligibility reforms.

Enrollment priorities were conceived as a method to help VHA stay within its appropriation and other resources, as VA care is not an entitlement like Medicare. As a result of the Act, (most) veterans must be enrolled in order to obtain VA health care. They are assigned to one of seven distinct enrollment priority groups and subsequently enrolled (Chart 4). They have access to a comprehensive range of benefits and services (VHA's "Medical Benefits Package" [11]). Some of the veterans who do not have to enroll include veterans who: (i) have a service-connected compensation rating of 50% or greater, (ii) have been discharged in the past year for a compensable disability that VA has not yet rated, or (iii) want care for a service-connected disability.

Chart 4
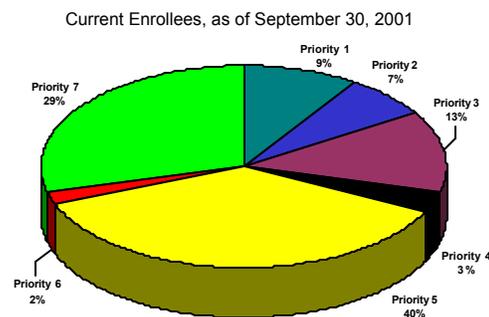
Current Enrollees, as of September 30, 2001

Source: Veterans Health Administration Data
Note: Total Current Enrollees = 5,848,067; excludes "ineligible", "declined", and "deceased".

Annually, VA assesses whether it will have the resources to meet the demand for care by veterans in all priorities. If, based on the Secretary's annual enrollment decision,[12, 13] it cannot, then VA may not continue to enroll veterans in the lowest level of priorities. However, for the last four years, VA has been able to open the VA health care system to all veterans, even higher income veterans, if they are willing to make co-payments. Other potential management efficiencies that might be achieved are also considered in the annual enrollment decision.

As of September 30, 2001, there were some 24,911,226 living veterans in the U.S. and P.R. and as of September 30, 2001, some 5,848,067 veterans (about 23% of all veterans living in the U.S. and P.R.) were enrolled in the VHA Health Care System. The existing Priority 7 includes

"higher income" non-service-connected veterans, who account for about 29% of all September 30, 2001 VHA enrollees. Also, the veteran population is declining over the next 10 years, but older age groups are increasing, trends that will have tremendous impact on VA.

Since the inception of VHA Enrollment, the number of Priority 7 veterans has shown the largest increase, both in absolute numbers and percent. They are, however, the lowest cost enrollees since they have other eligibilities and insurance and rely to a lesser degree on VA than enrollees in other priorities. They may be coming to VA to bridge gaps in their insurance coverage or to reduce their out-of-pocket costs. Based on the enrollment projections, developed for the Secretary's annual enrollment decision,[12, 13] enrollee demand shows no sign of decreasing, with a 31% increase in the number of enrollees from 6.1 million in 2002 to 8.0 million in 2010. Most of the increase is due to increases in Priority Category 5 and 7 enrollees. The latest VHA enrollment projections (as of September 2001) show VHA enrollment will continue to increase through 2010 and expenditures will also continue to rise, if no constraints are implemented and if resources (supply) can meet the projected demand.

Part IV: Veterans Health Administration enrollees and patients are very highly likely to be uninsured.

As we saw earlier, some 1.7 million veterans, or about 6.6% (adjusted) of the total veteran population were uninsured during all of calendar year 1999. The 1999 VHA Office of Policy and Planning (OPP) Survey of Veterans' Health and Reliance upon VA[3, 4] **(also, see "NOTE" above)** provides us related information on VHA enrollees and patients. The 1999 VHA OPP telephone survey had n=19,686 total respondents. **Chart 5 c**ompares uninsurance data on enrollees and patients from the 1999 VHA OPP enrollee survey with (adjusted) CY 1999 data on the veteran population from the March 2000 CPS. Point-in-time estimates (approximately, as of February 1999) from the 1999 VHA enrollee survey show point-in-time uninsured among enrollees to be about 28% of all enrollees and about 31% of all patients.

Chart 5: U.S. Population, VHA Enrollees, VHA Patients, and Total Veterans Percent Uninsured - 1999
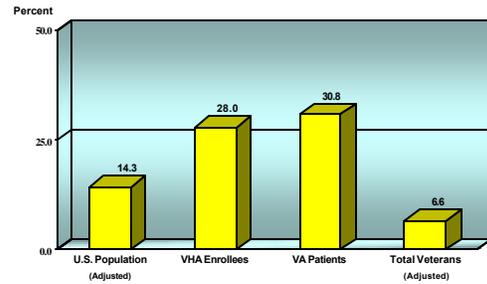
Sources: March 2000 Current Population Survey; 1999 VHA OP&P Survey of Enrollees.
Note: CPS data are adjusted full- year data for CY 1999; VHA OPP Survey data are as of March 1999; U.S. population data are for all ages; all other dat a are for persons age 18 or over.

Part V: As of February 1999, some 1.0 million VHA enrollees were uninsured and the (adjusted) upper limit in VA market share of uninsured veterans is about 59%, underscoring the role of VA as a safety net provider for many at-risk veterans

Since the VHA OP&P survey mentioned above was a survey of some 3,621,000 enrollees (as of February 1999), the observed 28.0% point-in-time uninsurance rate translates into a figure of about 1.0 million uninsured veteran enrollees. Since point-in-time VHA uninsured enrollees (numerator of market share) may be higher than corresponding full-year equivalent all veteran uninsured (denominator of market share), these data suggest that an upper bound for the (adjusted) VA market share of uninsured veterans is 59.0%. In light of the current health insurance coverage environment in the U.S., the fact that VHA is the largest health care system in the U.S., and the fact that VA has such a high market share of uninsured veterans, VA must be placed squarely at the center of national debates concerning the future of health care and health insurance coverage for Americans, particularly veterans. Access to as well as quality and equity of VA health care for veterans are related critical issues.[14, 15]

Part VI: Projecting Adjusted and Unadjusted Full Year Uninsurance Rates Of Veterans Using Simple Linear Regression Techniques

In this section, it is necessary to clarify what is meant by **adjusted** and **unadjusted** uninsurance rates, and, therefore, text from pages

1 and 2 above is repeated here: in March 2000 and March 2001, and in addition to the ordinary March CPS probes for health insurance coverage, Census included health insurance verification probes in the CPS questionnaire, to re-probe people who said they had no coverage at all during the previous calendar year. Thus, uninsurance rates which are computed irrespective of the verification probes are called **unadjusted** rates, and uninsurance rates which are computed based on the regular CPS health insurance probes in combination with the verification probes are called **adjusted** rates. Also, since the verification probes are new to CPS as of March 2000 (CY 1999 coverage), the CY 1999 and CY 2000 health insurance coverage data reported here, begin **a new trend series of adjusted rates and numbers** that will be followed over time.

The data of **Chart 3** show the major long-term trends in the **unadjusted** rate of full-year uninsured for the U.S. population and for veterans overall. It is evident from the chart that these divergent trends have strong linear components, and this is particularly so for the veteran data where aging is the greatest single factor in uninsurance rates. With this in mind, we can focus on the veteran data and seek a way to extrapolate the plotted veteran rates of **Chart 3** forward, to get an idea what the uninsurance rates for veterans might look like beyond CY 2000 (March 2001 CPS). At this writing, we are now in CY 2002, and we take cognizance of the fact that error in longer-term projections increases with time, so we will restrict our interest to trending to CY 2003.

Furthermore, a linear trend in the veteran data between 1987 and 1993 has already been hypothesized and included in **Chart 3**, so we further restrict analysis of the veteran uninsurance rates of **Chart 3** to the period of CYs 1994 - 2000 (and the March 1995 - 2001 CPS) for actuals, and CYs 2001-2003 for the projected numbers.

The basic method of analysis is to project the **unadjusted** veteran rates for CYs 1994 – 2000 forward to CY 2003 and to compute the **adjusted** rate for CY 2000 from the March 2001 CPS. We then project the **adjusted** data both <u>forward</u> to 2003 and <u>backward</u> to 1999

(coincident with the time of the 1999 Survey of Enrollees).

The method of analysis makes use of the **Euclidean Parallel Postulate**, one of the basic axioms of Euclidean Geometry: i.e., **given a line and a point not on the line , there is one and only one line through the given point that is parallel to the given line**. Also inherent in our method is the assumption that trends in **unadjusted** veterans rates, when extrapolated in a linear fashion, will define a line whose slope is the same as the slope of a similar line extrapolated from the equivalent **adjusted** veteran rates. This is a reasonable assumption, since we expect that there is an estimable difference between unadjusted and adjusted uninsurance rates that is approximately the same from year to year; also, there has been some standardization of the CPS health insurance verification questions and processes.

Baseline Regression: Regressing Unadjusted Veteran Uninsurance Rates Forward

Chart 6 shows the **unadjusted** veteran uninsurance rates for CYs 1994 – 2000, the **adjusted** veteran uninsurance rate for CY 2000, and the result of regressing the **unadjusted** rates of CYs 1994 – 2000 forward to 2001 - 2003. The method makes use of the **weighted least squares linear regression** add-in to Microsoft.

The data of **Chart 7** show the results of our baseline regression for projecting **unadjusted** veteran uninsurance rates forward. It is helpful at this point to mention that the linear regression process results in an estimated equation, **L1***, for a straight line, **L1**, that predicts **unadjusted** veteran uninsurance rates:

i.e., in terms of formulas, we estimate the trend in **unadjusted** veteran uninsurance rates,

L1: Y1 = a1 + b1*X

with the equation,

L1*: Y1* = (a1*) + (b1*)X,

where **a1* = 635.9714** and **b1* = -0.31429,**

from the Excel regression results in **Chart 7 below**.

Chart 6. Trend in Unadjusted Veteran Uninsurance Rates, CYs 1994 - 2003, and Actual Adjusted Veteran Uninsurance Rate for CY 2000
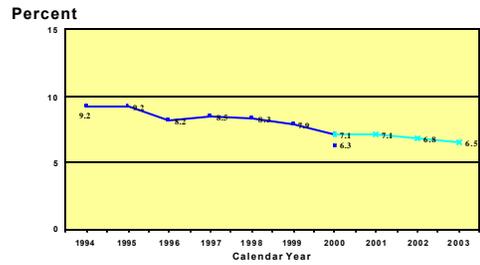
Sources: US. Population (March 1988-2000 CPS data); Veteran Population (March 1995-2000 CPS data, 1993 National Survey of Veterans (SOV-IV), 1987 National Survey of Veterans (SOV III), and trending for CYs88- 92). (Note: CPS data are CY; other data are point-in-time estimates and may be higher than equivalent full-year rates.)

Chart 7. Regression Results

SUMMARY OUTPUT

Regression Statistics

Multiple R	0.921479
R Square	0.849123
Adjusted R Square	0.818947
Standard Error	0.313506
Observations	7

ANOVA

	df	SS	MS	F	Significance F
Regression	1	2.765714	2.765714	28.13953	0.00318
Residual	5	0.491429	0.098286		
Total	6	3.257143			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	635.9714	118.3163	5.375182	0.003002	331.8303	940.1126	331.8303	940.1126
X Variable 1	-0.31429	0.059247	-5.30467	0.00318	-0.46658	-0.16199	-0.46658	-0.16199

The next step in our analysis is to use the results of the **L1/L1*** regression process to estimate an equation, **L0***, for a straight line, **L0**, that predicts **adjusted** veteran uninsurance rates.

Regressing Adjusted Veteran Uninsurance Rates Both Forward and Backward

We now use the results of the simple linear regression techniques of **Chart 7** to construct a linear trend line for the **adjusted** veteran uninsurance rates. To do this,

Consider:

L1: Y1 = (a1) + (b1)X

and

L0: Y0 = (a0) + (b0)X

where   **L1** and **L0**   are the simple linear regression equations for the trended **unadjusted** and **adjusted** veteran uninsurance rates, respectively.

We observe from **Chart 6** that (2000, 6.3) is a good estimate for a point on **L0**. This point on **L0**, recall, begins a new trend series of **adjusted** rates, as a consequence of inclusion by Census of health insurance verification questions in the CPS. (**NOTE**: Although Census included the new health insurance verification questions beginning in March 2000 (CY 1999), the March 2000 (CY 1999) CPS public use file does not have an indicator that allows computation of the **adjusted** rates, so our new trend series of **adjusted** rates for veterans really begins with March 2001 (CY 2000), even though the trend of **adjusted** rates for the U.S. population actually begins one year earlier with rates published by Census in official Census reports).

Since we have assumed **L0** is parallel to **L1** and that (2000, 6.3) is a point on the trend line for adjusted uninsurance rates, we have

b0* = b1* = -.314286 in

Y0* = (a0*) + (b0*)X ,

so that

6.3= (a0*) + (b0*)X$_{2000}$

and

6.3= (a0*) + (-.314286)(2000)

and, therefore, **a0* = 634.872**

That is, we have an estimate of the intercept term in our equation, **L0***. Thus, our estimated regression equation for **adjusted** veteran uninsurance rates is

L0*: Y0* = 634.872 + (-.314286)X

We can now use this estimated equation **L0*** to obtain some estimates of **adjusted** veteran uninsurance rates.   In particular, by **forward regressing**,

for X = 2001, we obtain Y0* = 6.0
and for X = 2002, we obtain Y0* = 5.7
and for X = 2003, we obtain Y0* = 5.4

However, we can also **backward regress** with the same estimated equation, **L0***, in order to estimate the percent of adjusted full-year uninsured veterans for CY 1999 with,

for X = 1999 in L0*, we obtain Y0* = 6.6

and, furthermore, we can plot all of these derived and trended rates (**Chart 8**).

Chart 8. Trends in Unadjusted Veteran Uninsurance Rates, CYs 1994 – 2003, and in Adusted Veteran Uninsurance Rates, CYs 2000 - 2003
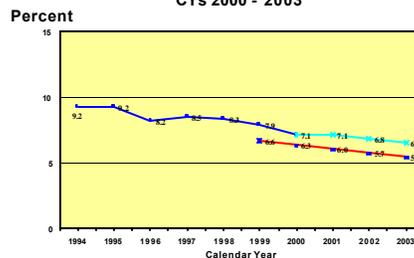
Chart 9 shows how the trended rates, i.e., **unadjusted** as well as **adjusted full-year uninsurance rates,** translate into projected numbers of uninsured veterans. Results of the **forward regression** of **unadjusted** rates and of the **forward and backward regressions** of **adjusted** rates are shown in the table.

Full-Year Uninsured (and Non-Institutionalized) Veterans, 2003

The data of Chart 9 provide us a snapshot of uninsurance for veterans in 2003, and approximately as of mid-year, i.e., June 30, 2003. In particular, the Chart shows that there will be some 1,299,705 full-year uninsureds among veterans in 2003.

These data require a few caveats. For example, the estimated numbers of uninsureds in **Chart 9** reflect non-institutionalized veterans only, since CPS is a survey of the civilian non-institutionalized population. Also, the veteran population figures upon which these estimates were calculated reflect the VA Veteran Population 2000 (Vet Pop 2000) model, and are contingent upon the assumptions inherent therein. An additional assumption is that the March CPS data on veterans (from the CPS veteran probe) each year should reflect veteran population model data estimates of the count of civilian non-institutionalized veterans in the 50 states and D.C.

Most importantly, full-year uninsurance rates are lower than corresponding point-in-time uninsurance rates, and it behooves us to examine the relationship between full-year and point-in-time rates to better understand the veteran uninsurance issue.

Chart 9
CPS Full -Year Uninsurance:
Unadjusted and Adjusted Rates and Numbers

Veteran Uninsurance

	Actual CY 1999 (Mar 2000 CPS)	Actual CY 2000 (Mar 2001 CPS)	Projected CY 2001	Projected CY 2002	Projected CY 2003
Unadjusted Rate	7.9%	7.1%	7.1%	6.8%	6.5%
6/30 Vet Pop #	25,757,063	25,453,121	24,999,449	24,536,924	24,068,609
Unadjusted Uninsured #	2,034,808	1,807,172	1,774,961	1,668,511	1,564,460
	Projected CY 1999 (Mar 2000 CPS)	Actual CY 2000 (Mar 2001 CPS)	Projected CY 2001	Projected CY 2002	Projected CY 2003
Adjusted Rate	6.6%	6.3%	6.0%	5.7%	5.4%
6/30 Vet Pop #	25,757,063	25,453,121	24,999,449	24,536,924	24,068,609
Adjusted Uninsured #	1,699,966	1,603,547	1,499,967	1,398,605	1,299,705

Part VII: Full-Year Uninsured vs. Point-in-Time Uninsured (Calendar Year Uninsured vs. Current Uninsured) and the Impact Upon Total Veteran and VHA Enrollee Uninsurance Rates

Full-year uninsurance and point-in-time uninsurance data have both been presented in this paper. It is essential to note that point-in-time uninsurance rates will generally be higher than full-year rates because anyone uninsured for the full calendar year will have been uninsured at every point-in-time during the calendar year. The question arises as to what the difference between the full-year and point-in-time rates might be.

There are little data available, but in CY 1992, the Survey of Income and Program Participation (SIPP) [16, 17, 18] showed that about 20% of the U.S. civilian non-intstitutionalized population had a coverage lapse during 1992, while CPS data (March 1993 CPS) for CY 1992 showed that about 15% of the U.S. population was uninsured the full-year.[1] Thus, any point-in-time rate during CY 1992 could be construed to have been as much as one-third (33%) higher than the corresponding full-year rate for CY 1992.

These data are not particularly recent, are **unadjusted**, and are only rough estimates. Nevertheless, the number of uninsured persons at any point in time, such as the point-in-time uninsured VHA enrollees of the 1999 VHA enrollee survey, might be considerably higher than any full-year rate alone might otherwise suggest. This means that there could be far more uninsured VHA enrollees than the CPS data of **Chart 9** alone might suggest (**Chart 10**, **Chart 11**). That is, point-in-time uninsured enrollees might easily outnumber full-year uninsured veterans overall.

Even with declining veteran uninsurance rates, VHA enrollment has increased dramatically over the last four years, and we would expect the total number of uninsured VHA enrollees in 2003 to be higher than at any other time in history, numbering well over one million.

Part VIII: Uninsurance of VHA Enrollees by Priority

Charts 10 and 11 show data on the uninsurance status of VHA enrollees, from the 1999 VHA Survey of Veterans' Health and Reliance Upon VA.[3] At this writing, a 2002 update and improvement on the 1999 enrollee survey is being fielded, and the 2002 data which are the equivalent to the 1999 data on uninsurance are not yet available. The data from the 1999 enrollee survey and the ongoing 2002 enrollee survey are approximately of the same time period, March of the survey year.

As Chart 10 shows, Priority 5 veterans are generally "lower income" and are the least likely to be covered. Priority 7 veterans are generally "higher income" and the most likely to be covered. However, even veterans with health insurance coverage or other eligibilities are coming to VA. As enrollees of all priorities age and increasingly depend on Medicare insurance, veterans continue to seek VA care for gaps in their insurance coverage such as pharmacy, long-term care, or to reduce their out-of-pocket costs for an expensive insured benefit. Also shown in Chart 11 are data from the latest VHA enrollment projections[13]; the enrollment projections are fiscal year based but are here interpolated to March of 2003 to be consistent with the time period of the enrollee surveys.

Chart 10

VHA Enrollees
% Uninsured by Priority Status

Percent

Age 18 or Over, % Uninsured

	Total	P1	P2	P3	P4	P5	P6	P7
	28.0	32.2	28.8	26.8	24.1	32.7	23.8	14.9
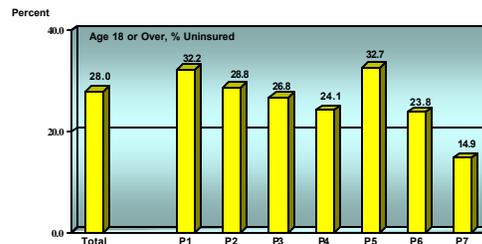
Source: 1999 VHA OP&P Survey.
Note: Data are as of March 1999.

SUMMARY

Even with declining veteran uninsurance rates, VHA enrollment has increased dramatically over the last four years, and it is expected that the total number of uninsured VHA enrollees in 2003 will be higher than at any other time in history, numbering well over one million. There will be data soon from the 2002 VHA Survey of Veterans' Health and Reliance Upon VA which, with past data from the 1999 enrollee survey and projections of future (2003) enrollment, will provide us more information on uninsured veterans who are enrolled in the VHA Enrollment System.

Chart 11. Point-in-Time Uninsured VHA Enrollees, and Projected Total VHA Enrollees

	Mar-99 _1/ % Uninsured	# Uninsured	Projected Mar-03_2/ Projected Enrollees
Total	28.0	1,013,046	6,598,880
P1	32.2	131,771	517,012
P2	28.8	82,182	399,424
P3	26.8	138,847	835,488
P4	24.1	23,140	168,286
P5	32.7	527,914	2,372,204
P6	23.8	13,044	147,581
P7	14.9	96,148	2,158,887

_1/ From the March 1999 VHA OPP Survey of Enrollees

_2/ Based on official VHA Projections; interpolated from EO Sept Projections

References

1. R. J. Mills, "Health Insurance Coverage, 1999: Consumer Income", *Current Population Report P60-211*, U.S. Census Bureau, September 2000.

2. R. J. Mills, "Health Insurance Coverage, 2000: Consumer Income", *Current Population Report P60-215*, U.S. Census Bureau, September 2001.

3. "Survey of Veteran Enrollees' Health and Reliance Upon VA", Veterans Health Administration Office of Policy and Planning, U.S. Department of Veterans Affairs, March 1999 and 2000.

4. Y. Shen, A. Hendricks, S. Zhang, L. Kazis, et al; "Health Insurance and Use of Services by Veterans (March 1999)", Veterans Health Administration Office of Quality and Performance, and Center for Health Quality, Outcomes, and Economic Research, U.S. Department of Veterans Affairs, Summer 2001.

5. R. E. Klein and D. Stockford, 'Data on the Socioeconomic Status of Veterans and on VA Program Usage", VA Office of the Actuary, Office of the Assistant Secretary for Policy and Planning, U.S. Department of Veterans Affairs, May 2001.

6. R. E. Klein, "The Changing Veteran Population, 1990- 2020", VA Office of the Actuary, Office of the Assistant Secretary for Policy and Planning, U.S. Department of Veterans Affairs, April 2001.

7. "1993 National Survey of Veterans", National Center for Veteran Analysis and Statistics, Office of the Assistant Secretary for Policy and Planning, U.S. Department of Veterans Affairs, April 1995.

8. "Employment-Based Health Benefits: Trends and Outlook", *EBRI Issue Brief, No. 233*, Employee Benefit Research Institute, May 2001.

9. News Release (of 09/06/01), The Kaiser Family Foundation (Health Research and Educational Trust, HRET), data from the 2001 Annual Employer Benefits Survey, September 6, 2001.

10. "State Health Insurance Mid-Year Reports", Health Insurance Association of America, June 2000.

11. Department of Veterans Affairs, 38 CFR Part 17, RIN2900-AJ18, Enrollment – Provisions of Hospital and Outpatient Care to Veterans, Final Rule, Section 17.38, Medical Benefits Package, p. 54217.

12. "Enrollment Level Decision Analysis, FY 2002", Veterans health Administration Office of Policy and Planning, Department of Veterans Affairs, September 2002.

13. "Enrollment, Utilization, and Expenditure Analyses, Fiscal Years 2002-2010, Task Order #1, Modification #9", by Condor Technology Solutions and Milliman USA, Inc. for Veterans Health Administration Office of Policy and Planning, Department of Veterans Affairs, September 2001.

14. Mary E. (Beth) Smith (Martindale), G. Sheldon, R. E. Klein, T. Feild, R. Feitz, D. Stockford, S. Krumhaus, C. Alpert, "Data and Information Requirements for Determining Veterans' Access to Health Care", *Medical Care* 1996; 34:3; MS 45-54.

15. Maria L. Fonseca, Mary E (Beth) Smith (Martindale), R. E. Klein, G. Sheldon, "The Department of Veterans Affairs Medical Care System and the People It Serves", *Medical Care* 1996; 34:3; MS 9-20.

16. R. L. Bennefield, "Comparative Analysis of Health Insurance Coverage Estimates: Data From CPS and SIPP", Bureau of the Census, paper presented at the 1996 American Statistical Association Meetings, August 1996.

17. " Health Insurance Coverage – Who Had a Lapse Between 1991 and 1993?", *Bureau of the Census Statistical Brief*, SB/95-21, August 1995.

18. C. T. Nelson, R. J. Mills, "The March CPS Health Insurance Verification Question and Its Effect on Estimates of the Uninsured", Bureau of the Census, paper presented at the 2001 American Statistical Association Meetings, August 2001.

Contact: Donald Stockford, VHA Office of Policy and Planning (105D), Dept. of Veterans Affairs, 810 Vermont Ave, NW, Washington, DC 20420, Tel. (202) 273-5112, or E-Mail to Donald.Stockford@mail.va.gov., or FAX to (202) 273-9030.

This page intentionally left blank.

THE DEPARTMENT OF VETERANS AFFAIRS HEALTH CARE ENROLLMENT PROJECTIONS

Gregg A. Pane, MD, MPA, Chief Policy and Planning Officer, Mary E. (Beth) Martindale, Dr PH, Randall J. Remmel, PhD, MBA, Don Stockford, MA, Health Policy Analysis Group, Office of Policy and Planning, Veterans Health Administration, U.S. Department of Veterans Affairs

Abstract

The passage of the "Veterans' Health Care Eligibility Reform Act of 1996" (P.L. 101-262) changed the way the Department of Veteran Affairs (VA) delivers health care to veterans. To receive care under this law, most veterans must be enrolled. This paper describes the Department's projections of health care enrollment, utilization, and expenditures that are utilized each year for important policy decisions such as who the Department can continue to serve under its appropriation and other resources. An example of what the impact of internal or external changes in policies and other factors might be upon projections of one type of VA service, prescription drugs, is illustrated.

Background

For over 60 years, VA has been providing quality healthcare to America's veterans through the Veterans Health Administration (VHA), the nation's largest integrated health care delivery system. However, it is supported by a congressionally appropriated budget and is not an entitlement. Thus, the available resources have always determined to a large degree who the system will serve, what it will provide (or supply) to whom, how the delivery system is structured, and continues to do so today. The system began as a hospital system with very little outpatient care provided initially. Complex eligibility rules determined which veterans could be treated, where and how they could be treated, and for what condition. These developed over time into a patchwork maze of eligibilities, difficult for both clinicians and veterans to understand and navigate through to access needed care. In October 1996, Congress passed Public Law 104-262, the "Veterans' Health Care Eligibility Reform Act of 1996". For the first time in its history, once enrolled into the VA health care system, VA could provide the care to an enrollee that is needed to promote, preserve, or restore the health of the individual through a very comprehensive medical benefits package—the right care, at the right time, in the right place. The Law emphasized preventive medicine and primary care, as well as the specialty care, for service-connected disabled and special populations. It included most inpatient and outpatient care in accord with generally accepted standards of medical practice. Because of resource constraints, the law also mandated a system of enrollment as a tool to help VHA balance the demand for care with the resources available. After the passage of the "Veterans' Health Care Reform Act of 1996", VHA's Office of Policy and Planning (OPP) developed an actuarial health care services demand projection model through a contract with Condor Technology Solutions, Inc., and Milliman USA, Inc., an actuarial firm. This is the fifth year this model has been used to make enrollment-related projections and analyses.[1]

General Approach

This model projects enrollees, utilization and expenditures, and patients for the next Fiscal Year (FY) and future years based upon the accrual of actual health care enrollment before the annual projection model update is begun each year. Actual enrollment experience is tracked and reported monthly, with an enrollment-related database that is created and disseminated by OPP to all individuals and offices for their own business functions. OPP receives monthly updates of enrollment from VHA's Health Eligibility Center, which is responsible for the business operation of veteran enrollment. OPP merges the enrollment data with various measures of enrollee utilization and costs that are provided by VHA's Office of Finance and the Veterans Integrated Service Networks (VISN) Support Services Center. For the projection model, a master file of every enrollee and all the events about the veteran's enrollment and health care utilization has been created and is updated at least annually before the projection model is run.

---

[1] Department of Veterans Affairs, Enrollment, Utilization, and Expenditure Analyses, Fiscal Years 2002 – 2010, Contract #GS – 23F – 8025H, Task Order #1, Modification #9, September, 2001, Condor Technology Solutions, Inc., Milliman USA, Inc., Kathi S. Patterson, FSA, MAAA, Merideth A. Randles, John P. Cookson, FSA, MAAA, Michael J. Dekker, ASA, John W. Leo, Ph.D., Gary W. Massingill, FSA, MAAA, Stanley A. Roberts, FSA, MAAA.

The actuary applies the private sector's current experience of providing the services included in the VA Medical Benefits Package (MBP)[2] to the projected enrollee population. Private sector utilization norms are adjusted to the VA enrollee population by age, gender, morbidity, and reliance upon VA. This utilization is also adjusted by the degree of management within the VA system compared to the community private sector's degree of management. Projected enrollee expenditures are calculated by multiplying VA unit costs by the adjusted private sector utilization norms for VA enrollees. Unique patients are also projected based upon the enrollee and utilization projections.

The enrollee, workload, expenditure, and patient projections have been projected at the national, VISN, and the parent preferred facility level. Enrollees, utilization, and expenditures have been projected by county of residence and zip codes for some applications. In addition, enrollment-related projections have been made through the current year, and for future years, *e.g.*, through FY 2022. Applications include:

- The Secretary's Enrollment Level Decisions
- The Capital Asset Realignment for Enhanced Services (CARES)
- Enrollee Cost Sharing Analyses
- Budget Formulations
- Market and Unmet Demand Analyses
- Planning Model for VISNs
- Scenario Testing
- Policy Decision Analyses
- Private Sector Contracting

Policy Issues and Implications
VA Senior Management considers the following:

Demand--Whom do we serve?
Financial--Where are we going?
Services/Supply—What services do we provide and how?

Whom do we serve?

VA has traditionally served veterans with service-connected disabilities, VA pensioners, populations with special rehabilitation needs and specialty care, the low income veterans as a safety net, and other veterans as resources permit. With the "Veterans' Health Care Eligibility Reform Act of 1996", Congress established priorities for enrollment. The priorities listed below are being updated and proposed for regulatory action this year. Because of budgetary constraints and continuing increases in demand, VA may have to consider policies that might limit whom we serve.

---

[2] Each year VHA verifies with VA's Office of the General Counsel whether any changes have been made in the interpretation of what benefits are covered or not covered in the Medical Benefits Package as described in the current enrollment regulations: Department of Veterans Affairs, 38 CFR Part 17, RIN 2900-AJ18, Enrollment-Provision of Hospital and Outpatient Care to Veterans, Final Rule, Federal Register/Vol.64, No.193/Wednesday, October 6, 1999/ Rules and Regulations, 54207-54218.

Figure 1. VA Enrollment Priorities, FY 2001

Priority	Description
1	Veterans with service-connected conditions rated 50 percent or more disabling.
2	Veterans with service-connected conditions rated 30 to 49 percent disabling.
3	Veterans who are former POWs Veterans with service-connected conditions rated 10 to 29 percent disabling. Veterans discharged from active duty for a disability incurred or aggravated in the line of duty. Veterans awarded special eligibility classification under 38 U.S.C., Section 1151. Purple Heart Veterans.
4	Veterans who are receiving aid and attendance or housebound benefits. Veterans who have been determined by VA to be catastrophically disabled.
5	Non-service-connected veterans and service-connected veterans rated 0 – 9 percent disabled, whose income and net worth are below the established dollar thresholds.
6	All other eligible veterans who are not required to make co-payments for their care, including: World War I and Mexican Border War veterans. Veterans solely seeking care for disorder associated with exposure to a toxic substance, radiation, or for disorders associated with service in the Persian Gulf. Compensable zero percent service-connected veterans.
7	The following Priority Level 7 subgroups have been considered: Priority Level 7a Zero percent non-compensable service-connected veterans enrolling prior to a specified date with income above the statutory threshold; who agree to pay specified co-payments. Priority Level 7b Zero percent non-compensable service-connected veterans enrolling after a specified date with income above the statutory threshold; who agree to pay specified co-payments. Priority Level 7c Non-service-connected veterans enrolling prior to a specified date with income above the statutory threshold, who agree to pay specified co-payments. Priority Level 7d Non-service-connected veterans enrolling after a specified date with income above the statutory threshold; who agree to pay specified co-payments.

For the Secretary's enrollment level decision concerning whom we can serve in the coming fiscal year, OPP analyzes the expenditures needed to make the MBP available to all the next fiscal year's projected enrollees. For example, last Fall the projected demand for enrollment, utilization and expenditures resulted in the following information.

Table 1. FY 2002 Enrollment-Related Projections

Priority	Projected Live End-of-Year Enrollees	Projected Average Enrollment	Projected Total Unique Enrollees	Projected Unique Patients	Projected Medical Benefits Package Expenditures	Cumulative Projected MBP Expenditures
1	519,686	521,840	540,660	431,137	$3,680,025	$3,680,025
2	396,302	391,741	410,149	279,089	$1,217,133	$4,897,158
3	819,210	798,902	846,837	525,220	$1,992,862	$6,890,020
4	168,349	168,051	180,610	158,101	$2,600,802	$9,490,822
5	2,322,426	2,263,640	2,416,167	1,787,627	$7,735,314	$17,226,136
6	138,335	128,845	141,024	75,471	$160,049	$17,386,185
7a & 7b	88,636	85,535	91,691	49,603	$117,627	$17,503,812
7c	1,607,478	1,640,430	1,673,824	854,295	$1,732,099	$19,235,910
7d	320,544	143,039	326,955	156,584	$141,720	$19,377,630
Total	6,380,966	6,142,023	6,627,916	4,317,127	$19,377,630	

Since the projected MBP expenditures above ($19.378 billion) were greater than the estimated resources initially available for supporting the MBP reported below in Table 2 ($18.937 billion), VA assessed through what priority VA could continue to enroll veterans. After all resources and efficiencies were considered, there was a final MBP difference of $142 million (Tables 1 and 2).

Tables 1 and 2 revealed that VA could continue to enroll all priorities of veterans except those non-service-connected (NSC) veterans in Priority 7d, who were not already enrolled prior to December 1, 2001. This was the date determined to divide proposed subpriorities 7c and 7d (referred to as *7iii* and 7*iv* in the proposed subpriority regulation RIN 2900-AK50) if enrollment were stopped for Subpriority 7d.

Table 2. FY 2002 Estimated MBP Resources Available vs. Projected MBP Expenditures

Resources and Expenditures	(Billions $)
Estimated Medical Care Appropriation	$21.331
Collections for Copayments, Deductibles, Third-Party Reimbursements, Other Revenue, and Carry-Over Funds	1.606
Subtotal, Resources for All MBP and non-MBP Services	*22.937*
Less Resources for Non-MBP Services[3]	(4.001)
Subtotal, Resources for the MBP	*18.937*
Projected MBP Expenditures	19.378
Subtotal, Initial Difference (Projected MBP $ and Resources Available)	*(0.441)*
Less Projected Policy and Management Efficiencies[4]	0.299
Final Difference, Projected MBP $ and Adjusted Resources Available	**($0.142)**

[3] Certain types of services for specified veterans are not included in the MBP, but are provided under other authorities*, e.g.,* long-term care, domiciliary care, dental care, per diem payments for State Homes, emergency care, CHAMPVA, readjustment counseling, certain prosthetic services, and counseling treatment for sexual trauma.

[4] Management savings through improved standardization policies and compliance in the procurement of supplies, pharmaceuticals, equipment and other capital purchases; adherence to national criteria established to promote operational efficiencies in current and new Community-Based Outpatient Clinics (CBOCs); and improved guidance and control of centrally managed programs.

Last Fall, VHA's total FY 2002 medical care appropriation was estimated to be $21.3 billion. This is supplemented by additional funds from collections for copayments, third-party reimbursements for services, and other revenue, including the effect of new outpatient copayments. The sum of these resources is $22.9 billion. These resources include $4.0 billion for services provided that are not included in the medical benefits package.[5] This leaves $18.9 billion available for the medical benefits package. When these available resources for the medical benefits package are subtracted from the projected expenditures ($19.4 billion), there is a resulting shortage of over $441 million. VA believes that this difference within the medical-benefit package can be lessened, but not eliminated, by taking additional management actions that are estimated to be $299 million. VA expects the Office of Management and Budget (OMB) to approve a supplemental request to Congress for funding to allow for continuation of full enrollment. Thus, the Secretary announced his decision on November 30, 2001, to continue to enroll all priorities of veterans in FY 2002. However, if supplemental funding is not received, additional enrollment action may be necessary during FY 2002.

Longer-Range Projections

It is impossible to determine how world events will unfold. Those events that impact our economy and the use of our military may have a profound impact on VA's long-range enrollment and expenditure projections. Nevertheless, long-range projections are given below, with the caveat that actual results will differ from those projected here for many reasons. It is important that actual enrollment and expenditures be monitored and the projections updated regularly.

Politically, VA and the legislative and executive branches of government found it difficult at this time to restrict care to veterans when a war on terrorism was being waged. It is expected that Congress will appropriate the funds to cover the unexpected high demand for enrollment this year. But VA continues to wrestle with the tension between demand for services and resources available in its budget planning processes for FY 2003 and 2004. As a response to these pressures, the Secretary of VA has consulted its leaders and constituents in several focused policy discussions about whom VA should serve and how. The projected demand for enrollment and services suggest this will be a continuing problem.

---

[5] Op. Cit. (3).

Chart 1. Average Projected Enrollment

	2002	2003	2004	2005	2006	2007	2008	2009	2010
□ Priority 7c	1,783,469	2,083,331	2,339,415	2,557,693	2,742,938	2,899,221	3,030,480	3,139,621	3,229,199
■ Priority 7a	85,535	91,844	97,134	101,536	105,158	108,090	110,393	112,162	113,451
■ Priority 6	128,845	148,267	165,354	180,323	193,423	204,783	214,182	221,870	227,907
■ Priority 5	2,263,640	2,377,693	2,462,747	2,524,824	2,568,283	2,596,853	2,613,558	2,620,425	2,619,467
■ Priority 4	168,051	168,312	167,775	166,575	164,840	162,696	160,267	157,663	154,987
□ Priority 3	798,902	837,798	863,722	878,811	884,890	883,619	876,508	864,803	849,561
■ Priority 2	391,741	399,956	404,144	404,961	402,991	398,702	392,585	385,070	376,500
□ Priority 1	521,840	516,764	510,993	504,280	496,439	487,497	477,658	467,113	456,221

Fiscal Years

Chart 2. Projected Annual Expenditures

	2002	2003	2004	2005	2006	2007	2008	2009	2010
□ Priority 7c	1,873,819,204	2,343,259,371	2,818,259,434	3,303,036,196	3,793,921,776	4,301,975,022	4,810,162,417	5,311,978,264	5,807,604,783
■ Priority 7a	117,626,923	133,670,963	149,911,522	166,538,691	183,238,430	200,674,105	218,210,653	235,480,552	252,449,128
■ Priority 6	160,048,855	199,355,466	240,339,542	282,667,138	325,747,501	371,650,046	417,658,774	462,609,242	506,013,746
■ Priority 5	7,735,314,158	8,522,810,369	9,311,560,004	10,100,673,509	10,871,328,989	11,651,481,405	12,414,238,839	13,150,163,555	13,868,527,764
■ Priority 4	2,600,801,813	2,715,994,974	2,842,269,620	2,975,158,016	3,102,489,376	3,230,242,587	3,348,227,315	3,452,836,121	3,550,985,812
□ Priority 3	1,992,861,802	2,177,713,482	2,363,271,259	2,541,873,869	2,709,228,864	2,872,615,543	3,027,730,257	3,172,699,712	3,308,752,868
■ Priority 2	1,217,132,559	1,301,894,310	1,389,372,512	1,475,053,023	1,555,537,722	1,634,909,280	1,711,723,191	1,784,183,210	1,851,954,716
□ Priority 1	3,680,024,755	3,831,550,910	3,996,040,239	4,167,529,065	4,330,895,917	4,498,470,597	4,664,808,274	4,819,791,530	4,953,722,181

Fiscal Years

General Model Description

The following outline provides a general description of the methodology used to develop the Veteran Enrollment, Health Care Utilization And Expenditure Projection Models (the model). These models were created by Condor Technology Solutions, Inc., and Milliman USA to support VA's Enrollment Level Decision, and later enhanced to support CARES analyses, VHA's budget formulation, and other policy decisions.

Enrollment Projections

1. Obtain baseline actual enrollment by scrambled Social Security Number (SSN)
2. Develop enrollment rates using historical enrollment and historical veteran population projections (VETPOP)
3. Develop projections of new enrollees using the rates developed in Step 2, the baseline from Step 1 and VETPOP projections
4. Apply mortality rates to enrollment projections

Workload Projections

1. Summarize private sector health care utilization averages by geographic area
2. Adjust utilization to reflect Medical Benefit Package and Millennium Bill health care services
3. Adjust utilization to reflect age and gender characteristics of the projected veteran enrollee populations
4. Adjust utilization to reflect the morbidity of the projected veteran enrollee populations relative to the underlying private sector populations (VA patient diagnosis data used to assess relative morbidity levels)
5. Adjust utilization to reflect the estimated degree of health care management observed within the VA health care system relative to the loosely managed level observed in the local community (VA inpatient diagnosis and workload data used to assess degree of health care management)
6. Adjust utilization to reflect the estimated veteran enrollee reliance on VHA for their health care needs (Veteran enrollee survey data and CMS[6] match data used to assess reliance)
7. Adjust utilization to reflect the residual differences between modeled and actual historical VA workload (estimates of unmeasured morbidity, reliance and degree of health care management differences)

Unit Cost Projections

1. Obtain baseline Cost Distribution Report (CDR)-based VA unit cost data
2. Unit cost data adjusted for health care service mix inherent in data
3. Adjust VA-based unit costs to residual differences between modeled and actual historical VA expenditures

Expenditure Projections

1. Enrollment, Workload and Unit Cost Projections are combined to produce Expenditure Projections

---

[6] Centers for Medicare and Medicaid Services, formerly the Health Care Financing Administration (HCFA)

VA Enrollee Health Care Projection Model
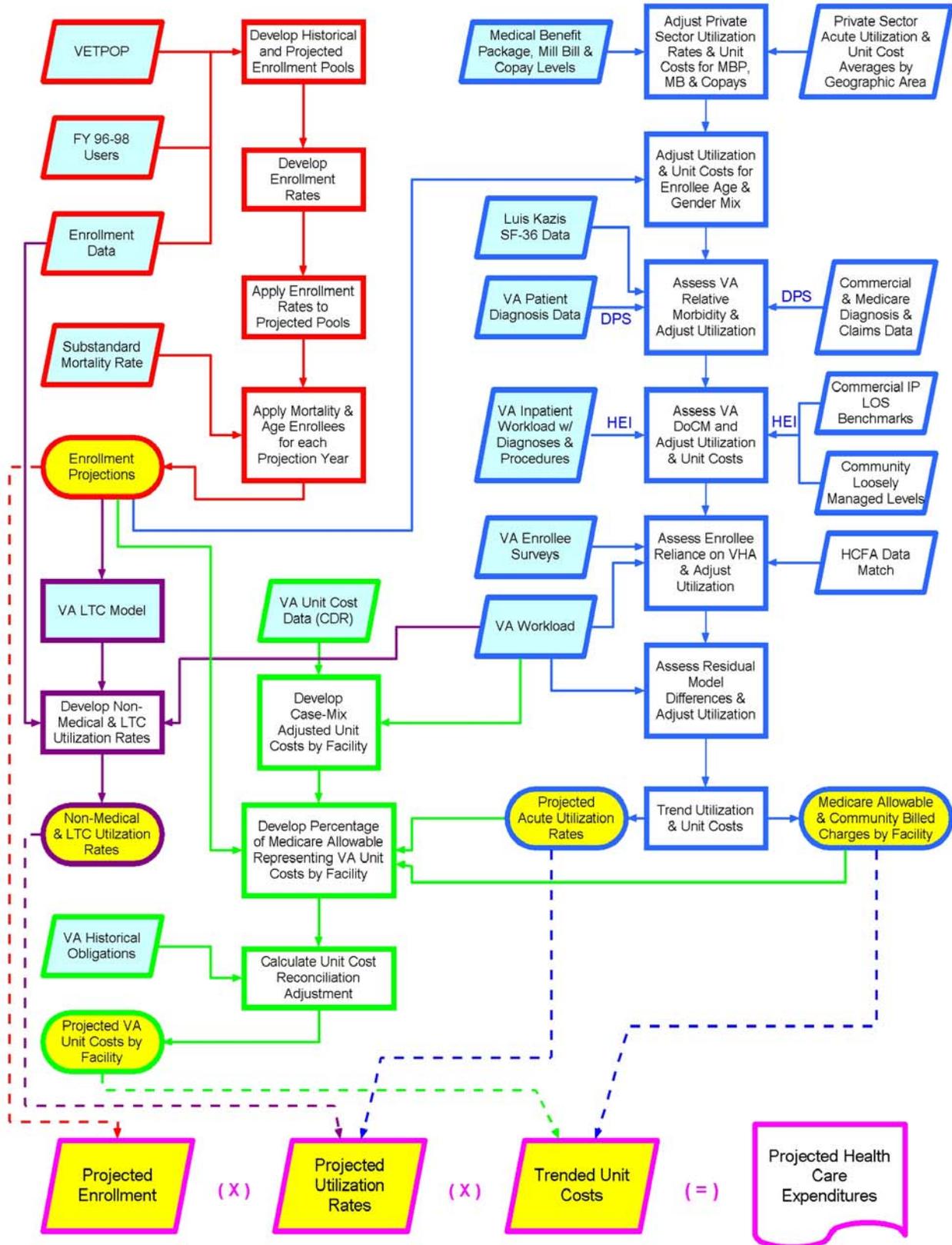
In most of the applications, enrollees are assigned to a preferred facility, where the veteran's care is managed by a VA health care provider. The cost models reflect the projected health care demands of the enrollees by preferred facility. It is not anticipated that all of the enrollees' VA demanded health care will necessarily be obtained from that preferred facility; consequently, most of the cost model applications are enrollee-based, not facility-based.

Costs were projected for providing the health care benefits defined in the Medical Benefits Package as well as other VA non-medical services to the Enrollees. Expected utilization by Facility, Enrollee Type (Enrollee Pre and Enrollee Post), Age Group (Under Age 45, Ages 45 to 64 and Ages 65 and Over), and Priority Level were developed using private sector utilization adjusted to reflect the veteran enrollment populations and an appropriate level of managed care for the VA. This health care utilization is detailed by several Inpatient and Ambulatory medical service categories. Estimated VA unit costs based on VA's Cost Distribution Report (CDR) and related data sources were applied to the expected utilization by medical service category. From the utilization and VA unit cost data, expected per member per month (PMPM) costs were calculated for each combination of Facility, Enrollee Type, Age Group, and Priority Level veteran Enrollees. The PMPM cost is the cost of providing health care to each member, in this case veteran Enrollees, for a one-month period of time.

Each cost model has been adjusted to reflect relative veteran morbidity and reliance on VA for obtaining health care services. These adjustments vary by VISN, Enrollee Type, Age Group, Priority Level, and service category.

The partial reliance adjustments reflect the fact that the majority of veterans (particularly those who qualify for Medicare) have another choice for health care services. Consequently, veterans can utilize health care from facilities both inside and outside VA concurrently. The partial veteran reliance in these models reflects estimated current veteran reliance on the VA health care system.

The relative morbidity adjustments reflect the relative health status of veteran Enrollees compared to the private sector populations underlying the utilization benchmarks. These adjustments are based on a diagnosis-based risk adjustment methodology which incorporates the responses to the OPP's 1999 Survey of Enrolled Veterans performed by Computer Hardware Maintenance Corporation (a division of Condor), and its 2000 Survey of Enrolled Veterans performed by Shugoll Research (through a contract with Condor) and the 1999 Health Survey of Veterans (Veterans SF-36 & Health Behaviors) supported and funded by VHA's Office of Quality and Performance.

Enrollment Trends
Short-and Long-Term Considerations

The total veteran population is declining over the next 10 years, but older age groups are increasing, *e.g.*, the age 85 and over more than doubles (113% increase) in actual numbers. If there are no interventions, enrollee demand shows no sign of decreasing, with a 31% increase in the number of enrollees from 6.1 million in 2002 to 8.0 million in 2010. Most of the increase is due to increases in both the mandatory low-income Priority 5 veterans and the higher income Priority 7 veterans. Since both are the largest priorities of veterans, with a relatively small current market share, there is considerable potential for an expanding demand for enrollment from these two population subgroups. VA already has a large market share of the veterans in the other service-connected and pensioner sub-populations who need specialized care, aid and attendance, or other complex care (Priorities 1-4).

A concomitant increase in enrollee utilization results in an increase in MBP expenditures of 76% from $19.4 billion in FY 2002 to $34.1 billion in FY 2010, an average of $1.8 billion per year. Priorities 1-6 (mandatory populations) account for $10.7 billion or 72.8% of this increase, 60% due to Priority 5.

These same long-term trends are exacerbated if the total, not just the MBP, expenditures are considered. These total VHA health care expenditures increase $17.9 billion or 75% from $23.8 billion in 2002 to $41.7 billion in 2010, an average increase of $2.2 billion per year. Priorities 1-6 increases account for $13.2 billion or 74% of these increases from 2002 – 2010. The $17.9 billion increase over time can be contrasted with the fact that OMB has projected increases for the same time period of $3.8 billion that is primarily only inflation of $0.5 billion per year.

**Chart 3. Total VHA Projected Health Care Expenditures
(Medical Benefits+Long Term Care+Services Under Other Authorities)**
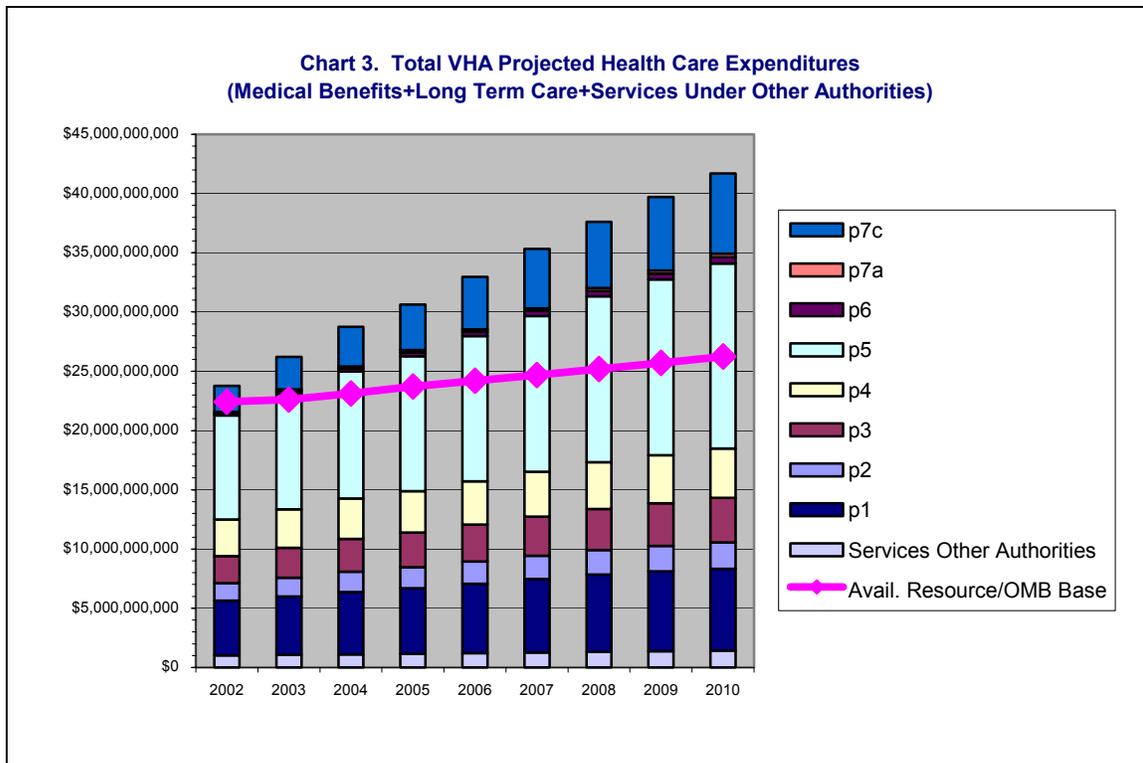
The "Veterans' Health Care Eligibility Reform Act of 1996" assured every enrollee of receiving a comprehensive package of high quality inpatient and outpatient care in a timely manner. Because of mounting financial pressures, VHA is considering the development of several different policies for modifying either whom we serve, what services are provided, or what out-of-pocket costs may be required. One of the more perplexing trends in service utilization is that for the pharmacy benefit. An example is given below of some of the actual and projected experience with this service and how such projections may shape major policy decision-making within this government agency. Many endogenous, as well as exogenous, factors have influenced our projections and the policies associated with the pharmacy benefit.

VHA Pharmacy Benefit

The demand for enrollment often reflects the personal economic decisions of veterans who seek to cover gaps in their insurance or other resources such as Medicare, or to reduce their out-of-pocket expenses for selected services. Many Medicare HMO's are dropping their coverage of the elderly and disabled in some geographic markets or are scaling back their coverage of drug costs, a service currently not mandated in the Medicare benefits package, but which was offered in some Medicare HMO's as an incentive to enroll with that type of Medicare provider. Medicare eligible veterans may be turning to the VA health care system at a time when its financial resources are also constrained.

A number of various analyses have been initiated within VA that are looking at VHA's actual historical experience to assess those who are using pharmacy services and to what degree. In addition, the VHA enrollee projection model generates long-range projections of pharmacy utilization and expenditures. Because of the influx of the Medicare eligible veterans who may be enrolling primarily for the pharmacy benefit, many of the analyses have focused on the Medicare eligible enrollee or the age 65 and over enrollee as a proxy for Medicare eligibility.

In FY1999, 52% of the veterans who had been VA patients in FY 1997, 1998, or 1999 (1,026,021), received care in both VA and Medicare systems. In FY 1999, the same percent (24%) used either VA only (not Medicare FFS) or Medicare only. In FY 1999, Medicare eligible veterans who had been VA patients in FY 1997, FY 1998, or FY 1999, received $18.8 billion in care from either VA or Medicare. Of this total FY 1999 expenditure in both systems, half ($9.5 billion) was borne by VA.

Chart 4. FY 99 Medicare Eligible Veteran Patients who used VA in FY 97, 98, or 99

Chart 5. FY 99 Expenditures Components for Dual VA and Medicare Users
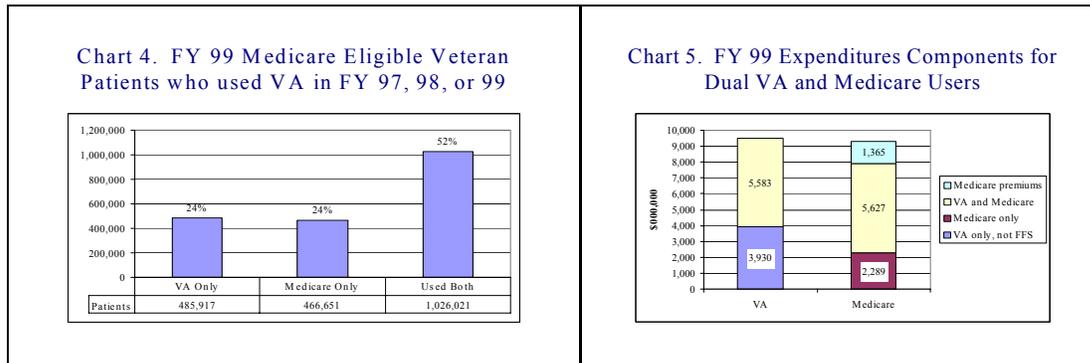
The more recent VA/Medicare dual user experience is not reflected in the older VA/Medicare matched data. In FY 1999 there were only 847,584 Priority 7 enrollees. By the end-of-year FY 2001 there were 1,747,591, a 106% increase. The increase in Priority 7 Medicare enrollees was even greater (138%), from 410,446 in FY 99 to 975,343 in FY 01. Total expenditures for enrollees age 65+ (a Medicare proxy) increased 121% from FY 99 to FY 01. This recent growth in total expenditures from FY 99 to FY 01 was almost entirely attributed to 'Post' enrollees who are new to the VA system after enrollment began (564% increase in total expenditures), *i.e.*, 'Post' enrollees are those enrollees who were not a VA user in FY 96, 97, or 98 prior to enrollment implementation. Priority 7 age 65+ enrollee total outpatient expenditures increased 170% from FY 99 to FY 01. Most of this growth occurred in the Post enrollees (714%) from FY 99 to FY 01.

Many dual VA/Medicare eligibles are coming to VA for services not covered by Medicare, *e.g.*, prescriptions. In just one VA network in Florida, VA's Inspector General found 43% of the P7 cases reviewed (949) indicated that since they had private sector primary and other specialty care, the sole purpose of their VA care was for prescriptions. These cases (949) represented an estimated $11.9 million annually in direct prescription costs.[8] In order for a veteran to currently receive prescriptions in VA, a VA clinician must exam the patient and prescribe the medication. Thus, duplication of services and poor coordination of care across systems and providers may occur with potential quality of care problems.

VA Pharmacy Projection Trends

It was projected last fall that VA pharmacy expenditures would increase 148% from $3.2 billion in FY 2002 to $7.8 billion in FY 2010. Pharmacy expenditures for the elderly were projected to increase 138% from $1.8 billion in FY 2002 to $4.3 billion in FY 2010. For the enrollees age 65 and over, Priority 7 pharmacy expenditures were expected to increase almost 257% from $386 million to $1.4 billion from FY 2002 to FY 2010. The projected pharmacy expenditure trend for enrollees age 65 and over by priority is illustrated in Chart 6. Chart 7 presents the projected total expenditure trend for enrollees age 65 and over by priority.

---

[7] Dual utilization of VA and Medicare Systems: Expenditure Trends from FY 98 and FY 99, Management Sciences Group, VHA's Office of Policy and Planning, VHA Budget and Policy Roundtable, September 17, 2001.

[8] Audit of VHA Pharmacy Co-Payment Levels and Restrictions on Filling Privately Written Prescriptions for Priority Group 7 Veterans, Report No. 99-00057-4, issues 12/20/0, VA Office of the Inspector General.

Chart 6. Projected Pharmacy Expenditures for Age 65+ Enrollees, FY 2002-2010
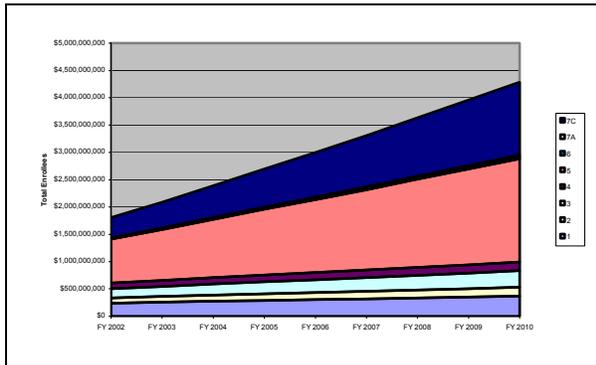
Chart 7. Total Projected Expenditures for Age 65+ Enrollees, FY 2002-2010
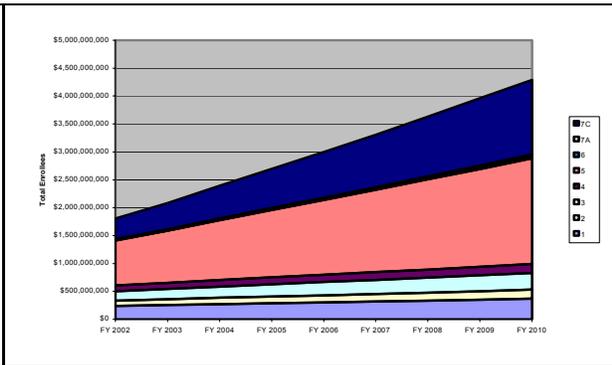
For the Medicare proxy (age 65 and over) enrollees, it was projected that non-service-connected Priority 7c enrollees would use more of its total health care expenditures for pharmacy services than any other priority. *i.e,* expenditures in Chart 6 divided by expenditures in Chart 7. In FY 2002, 33.3% of all Priority 7c expenditures were consumed by pharmacy (Chart 8), with an average percentage of pharmacy to total expenditures across all priorities of 18.3%. This Priority 7c percentage of pharmacy to total expenditures grew to 41.1% by FY 2010 (Chart 8). Higher priority enrollees who had always had access to prescriptions in VA, *e.g.,* Priorities 1-6, had the lowest percentages of pharmacy to total expenditures. The service-connected Priority 7a enrollees had the second highest percentages of pharmacy to total expenditures. The spread or difference in VA pharmacy use by the elderly Priority 7's versus those who are under age 65 and the remaining priorities is illustrated in Charts 8 and 9.

There were much smaller differences by priority in the percentage of pharmacy to total expenditures for the under age 65 (Chart 9). For the younger age enrollees, Priority 7c continued to have the largest VA pharmacy to total VA expenditure percentage, 17.9% in FY 2002, while the average of all priorities for the younger enrollees was 14.2%. This pharmacy percentage grew to 23% in FY 2010 for Priority 7c. The younger groups that are employed may also have improved insurance coverage with a pharmacy benefit, unlike most elderly enrollees.

Chart 8. Pharmacy as Percent of Total Expenditures (Age 65+)
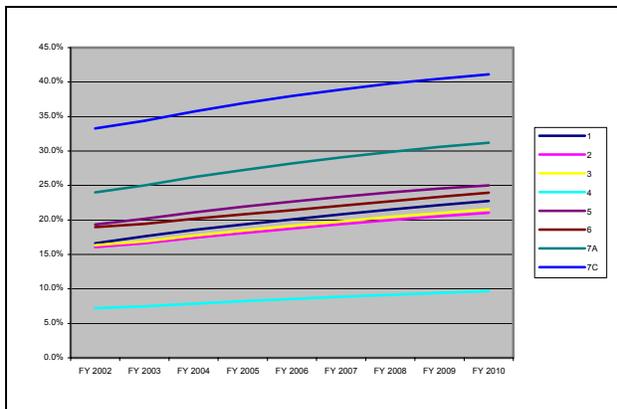
Chart 9. Pharmacy as Percent of Total Expenditures (Age <65)
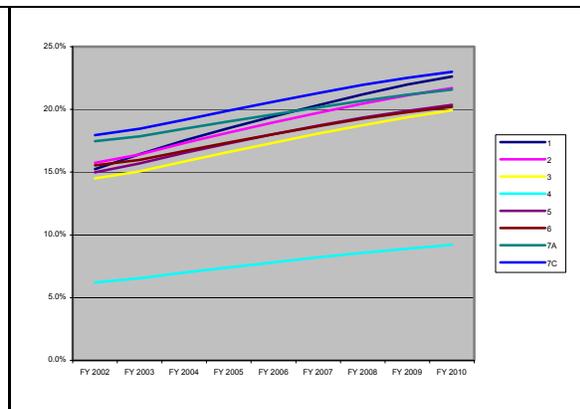
Implications

Until Congress enacts a pharmacy benefit for the Medicare population or VA alters its current policies on enrollment, VA will continue to experience increasing demand for VA health care that includes increased prescription utilization, especially by the elderly, higher income veterans who have not previously been able to access VA care.

Scenario Analysis

Chair: Karen S. Hamrick, Economic Research Service, U.S. Department of Agriculture

Scenario Analysis with a U.S. Computable General Equilibrium (CGE) Model

Kenneth Hanson, Economic Research Service, U.S. Department of Agriculture

A U.S. Computable General Equilibrium (CGE) model has been used at USDA-ERS, by Ken Hanson in cooperation with others, to assess economywide impacts of farm and food policies. Use of the CGE model in scenario analysis is an exercise in comparative static economic analysis. The model characterizes the state of the economy at a particular point in time. A policy change is analyzed by translating the policy change into a change in exogenous model parameters, imposing a change to these parameters in a simulation experiment with the model, and translating the model results in terms of how the economy would adjust to the new policy situation. We illustrate this exercise with two recent examples, a change in Food Stamp Program expenditures, and a movement of low-skilled workers from welfare to work due to welfare reform.

Agricultural Sector Scenario Analysis and the ERS Country-Commodity Linked System

Ralph Seeley
Paul Westcott
Economic Research Service, U.S. Department of Agriculture

The Country-Commodity Linked System (CCLS), a partial-equilibrium system with 24 commodity markets and 44 regions, is used for scenario analyses at the USDA's Economic Research Service. New scenarios may require increased capacities for the models, analysts, algebraic techniques, and overall system. New approaches were needed to model China's WTO accession with non-state trading enterprises and tariff-rate quotas, to model European Union enlargement and grain price policies, to model Western Hemisphere integration with bilateral tariff reduction, and to model Taiwan's pork industry collapse. The reasoning behind those scenarios, as well as initialization and calibration, may suggest solutions to other problems.

Scenario Analysis Using a Global Dynamic Applied Equilibrium Model

Agapi Somwaru, Economic Research Service, U.S. Department of Agriculture
Xinshen Diao, International Food Policy Research Institute (IFPRI)

A Global Applied Equilibrium Model, both static and dynamic in its specification, has been used to evaluate the effects of removing trade barriers, subsidies, and other trade distortions forms of support on the world economy. The model, based on a global database, utilizes trade and domestic support instruments to capture worldwide adjustments to policy changes. It simulates changes in policy as scenario analyses in counterfactual or "what-if" comparisons. Depending on the assumptions of factor mobility the model can assess medium- to long-run changes. The model results of complete trade agricultural reform increases world trade in agricultural commodities, but leaves the level of total agricultural production almost unchanged. In the medium term, some net agricultural importing countries suffer a welfare loss due to an adverse change in their terms of trade that reform causes. In the longer run, however, agricultural policy reform benefits almost all countries, and in particular, developing countries, due to the change reform induces in the developing countries' investment pattern, growth in capital stock, and growth in their total factor productivity.

This page intentionally left blank.

Scenario Analysis with a U.S. Computable General Equilibrium (CGE) Model

By
Kenneth Hanson
Economic Research Service, USDA
1800 M Street NW, S-2093
Washington DC 20036

(202) 694-5427
khanson@ers.usda.gov

Introduction

Scenario analysis with an economic simulation model provides a laboratory for gaining insight into the economic impacts from a policy change or an exogenous change in a model parameter characterizing some technological or behavioral feature of the modeled economy. The purpose of scenario analysis is "not strictly to predict the future but to facilitate a systematic exploration of … critical events within some explicit time frame" [Granger (1989, p. 224)]. A base set of assumptions is made, which is the "most probable" or "surprise-free" case. Plausible alternative scenarios are compared against the base case. Scenario analysis produces qualitative forecasts, not quantitative, point forecasts.

This paper provides a description of a Computable General Equilibrium (CGE) model for the United States, including the sources of data underlying the model. The paper also includes a brief review of how the model has been used in policy analysis, and how it compares with other types of CGE models. The first version of the model was used primarily for farm and trade policy analysis. The current model version, described here, has been developed to analyze food assistance issues as well as issues of farm policy. With a new policy focus, it has been necessary to reassess some components of model design.

Simulating a policy change in our CGE model is an exercise in comparative statics, a what-if comparison of two equilibrium states of the economy. The results of comparative static analysis are in terms of annual average changes in economic activity when the economy moves from the base equilibrium with the existing policies in place to a new equilibrium with the policy changes. The annual average changes occur over a time period required by the economy to adjust to a new equilibrium. We would generally expect equilibrium to be reached in 2 to 5 years. The length of that period depends on assumptions made about price-quantity responsiveness (elasticities) and whether aggregate capital stocks have time to adjust.

The new equilibrium is characterized by prices, which equate supply and demand in markets for goods and services, and satisfy the model closure rules. A CGE model describes and compares the old and new equilibrium but not the adjustment process. Though the paper does not provide any specific simulation analysis, it does reference various studies where the model has been used in policy analysis.

Use of the U.S. CGE Model in Scenario Analysis by the USDA-ERS

The use of CGE models in scenario analysis is a vast field of study with a long history, briefly explored in appendix A. At USDA-ERS, the focus of the U.S. CGE model has been on the analysis of farm and food assistance policy. An earlier model version documented in Robinson, Kilkenny, and Hanson (1990), was used primarily for farm and agricultural trade policy. Examples of farm program analysis are Kilkenny and Robinson (1990), and Bernat and Hanson (1995). An example of trade policy analysis is Robinson, Kilkenny, and Adelman (1989). Analysis of an export promotion program is Hanson, Vogel, and Robinson (1995).

A version of the model to examine food assistance policy is documented in Hanson, et al. (2002), and with greater household detail in Hanson and Hamrick (2002). These models have been used to assess the economywide and farm sector impacts from a change in Food Stamp Program expenditures and from converting the entitlement program's in-kind benefits to a fixed annual block grant to States with the State option to cashout the program. The impact of the Federal budget procedure on funding an increase in Food Stamp program expenditures has been explored in Hanson and Golan (2002). The version of the model with greater household detail has been used in Hanson and Hamrick (2002) to look at the labor market impacts of the reduction in the Food Stamp Program and TANF caseloads since 1996.

This static U.S. CGE model is programmed in the GAMS software [Brookee, Kendrick, and Meeraus (1992)]. The software allows the modeler to develop the program code for the model in general algebraic notation. The software serves as a front end to a selection of solvers for nonlinear optimization problems. An extension of the static CGE model is to generate a sequence of static equilibria consistent with baseline projections from a macroeconomic model [Robinson, Kilkenny and Adelman (1989), Hanson, Robinson and Tokarick (1993)]. The modeling exercise is tedious to be consistent with both macro and industry projections.

The USDA-ERS uses other types of CGE models for different types of issues. Regional and global trade agreements have been analyzed using several different CGE models summarized in Burfisher and Jones (1998). Articles in Burfisher and Jones (1998) include Gehlhar's use of the GTAP model for global trade analysis, Xinshen, Somwaru, and Raney's use of an inter-temporal multi-country model for analysis of Western Hemisphere trade policy integration, and Burfisher, Robinson, and Thierfelder's use of a static multi-country model for analysis of NAFTA. Another GTAP based multi-country model has been developed to focus on the impact of global climate change on United States and world agriculture [Darwin, et al. (1995)]. Two other models are a multi-region U. S. model and a single region, sub-nation model [Canning and Tsigas (2000), Vogel and Hanson (2001)].

A CGE Model for Scenario Analysis
Crucial issues of model design concern the specification of actors or economic entities and the rules governing their behavior and interaction in markets. The actors in a CGE model are specified to include all the major economic entities distinguished in the National Economic Accounts [Hanson and Robinson (1991)]. Consequently, the model includes all economic activity summarized in the National Income and Product Accounts [U.S. Department of Commerce, Bureau of Economic Analysis (1985), Seskin and Parker (1998)]. Five types of economic entities are households, producers, government, rest of world, and the capital account–the market for loanable funds where savings become business loans for new capital stocks, which serves as an institution for the allocation of savings to investment.

A CGE model is an economy-wide computer simulation model, which captures, in a stylized manner, the economic interactions among households, producers, and government. Each of these economic entities has multiple roles and all interact with each other. Households supply labor to producers, and consume goods and services using the income they earn. In addition, they receive income from the ownership of capital, receive government transfer payments, save, and pay taxes. Producers make goods and services for the market, use labor, capital stocks, and other goods as inputs into production. The government provides transfers and public services to producers and households, and receives tax revenue from producers and households.

Clearly, as households, producers, and government interact, a rather complicated economic process is taking place. A CGE model captures this economic process and provides a way to examine how shocks, such as changes in policy, affect the economy. A CGE model captures the linkages among economic entities and thus can trace the impacts from a shock through the economy.

One major contribution of a CGE model is its comprehensive look at the impact of policy change on the economy, as it works through the various linkages among the economic entities. In the case of welfare reform, the policy of interest is the shift from Assistance for Families with Dependent Children (AFDC) to Temporary Assistance for Needy Families (TANF) in 1996 and the impact it may have on the labor market. This policy change entails recipients shifting from welfare to work, whereby transfer payments decrease and labor market participation increases. The initial impact is a reduction in government spending targeted to low-income families, an increase in labor supply for low-skilled jobs, and a decrease in personal income taxes to offset the reduction in expenditures and maintain a budget neutral policy change. The CGE model can trace changes in household labor force participation through the labor market to industry demand for labor, and back to households through earnings. Other households are also affected as labor markets adjust to absorb the new labor supply. Each direct effect of a policy change creates its own set of ripple effects, captured by the CGE model.

Households and Producers
Each of the economic entities in the model may be aggregated at different levels of detail, refining the model's specification and making it relevant for specific policy issues. We aggregate producers by industry groups using the Input-Output Accounts [U.S. Department of Commerce, Bureau of Economic Analysis (1997)]. The industry groups are chosen to emphasize the role of agriculture and food processing in the U.S. Economy. Households are grouped to allow the scenario analysis with the model to focus

on changes in household work force status and on participation in government assistance programs, while recognizing that households with different family structures make choices under different circumstances.

We segment households into a number of social-economic categories using data from the Current Population Survey (CPS), March Supplement [U.S. Department of Commerce, Bureau of Census (1997a,b)]. The unit of analysis that we label "household" is our best approximation of a "consumption unit," and is not identical with the CPS household defined by a common address. For the household aggregation, we use four characteristics to segment households into distinct groups; family structure (5 types), income (3 levels), work force status of primary and secondary earners (3 categories), and participation in the Food Stamp Program (FSP) and Temporary Assistance for Needy Families (TANF). Not all combinations of household characteristics occur in the data, so our model includes 99 household groups.

Households receive income from three main sources: earnings from both wages and salaries and from self-employment; capital income from the ownership of assets–dividends, interest, and rent; and transfer income from government programs. Eleven government transfer programs are distinguished, some are for low-income households, while others are for the elderly. Data on household income by source are from the CPS, March Supplement [U.S. Bureau of Census (1997)]. Households use their income to consume goods and services, pay taxes, and save. Household expenditure shares are derived from the 1996 Consumer Expenditure Survey [U.S. Department of Labor, Bureau of Labor Statistics (1997)]. Savings and taxes are specified as fixed saving rates and tax rates specific to each household group. The tax rates are derived from the CPS March Supplement, while the savings rates are derived from the Federal Reserve, Survey of Consumer Finances [Bosworth, Burtless, and Sabelhaus (1991)].

Labor supply and demand are also treated in detail. The primary and secondary earner of each household type supplies labor in its own unique mix of occupations distinguished by skill (education and training) categories, using the CPS. Similarly, each industry demands labor in its own unique mix of occupations, using data from the U.S. Department of Labor, Occupational Employment Statistics. The work force status of unemployed and not-in-the-labor-force are also specified for both primary and secondary earners of each household type, where

appropriate. The greater detail in classifying household and labor occupations distinguishes this model from Hanson et al. (2002).

Households act as if they maximize utility, a measure of their well-being, through the purchase of an array of goods and services and the enjoyment of leisure, given a budget constraint, and a constraint on the total amount of time allocated to work or leisure. Firms maximize profits from the sale of goods and services to households, given their technology of production. Neither households nor firms are able to individually influence prices arising from the transactions taking place in these two markets. In equilibrium, the fabled invisible hand of the market determines the amount of goods and services produced by each firm, the prices they charge, how much they pay households for factor services, and how much of each good each household consumes.

Firms, grouped into industries by the type of commodity produced, purchase goods and services produced by firms in other industries. These inter-industry transactions are important because they link firms in different industries. As a result, a change in household consumption leading to a direct change in production in one set of industries and also leads to an indirect impact on production in another set of industries. The direct plus indirect impact on industry production from a change in final demand results in an impact on jobs by the education and skill categories.

While households decide how much labor they will supply to the market, industries are also determining how much labor they will demand, based on demand for their products. Both household labor supply and industry labor demand is by occupation, which capture differences in skill levels. Using a household-occupation matrix, we link the supply of labor and flow of labor income by skill level to specific household groups. Using an industry-occupation matrix, we link the demand for labor by occupation to industry production. This improves our ability to model labor market issues arising from welfare reform as low-income households move from welfare to work in low-skilled occupations.

Additional Institutions
We complete the specification of the CGE model by adding three other economic entities: government, Rest of World, and capital account. The government is divided into a Federal government and into an aggregated State and Local government. In light of the new federalism reflected in such legislation as the 1996 welfare reform, distinguishing the State and

local government from the Federal government is important. Federal block grants for welfare programs to State and local governments are an inter-governmental transfer.

These two levels of government have separate budgets, taxes, expenditures, and transfers. The State and Local government is not allowed to have a deficit, whereas the Federal government is able to finance budget deficits through borrowing on the capital market. Different types of taxes are distinguished for each level of government. Taxes include personal income tax, corporate profit tax on capital income, social security tax on labor income, a business tax on sales, and tariffs on imports. Expenditures for goods and services are distinguished by components of demand associated with different government activities for each level of government. Transfers from government to persons are aggregated into eleven programs and are distinguished by whether they are Federal and/or State and local programs.

The Rest of World (ROW) and the capital account are indirectly important for the analysis of welfare reform, through linkages from their demand to industry production and the occupational requirements by these industries. The treatment of the ROW as one actor rather than many countries is a standard simplification for single country models, and which we maintain in our CGE model. The ROW supplies imports to the United States and purchases exports from the United States. The model contains a balance-of-trade constraint in which the value of imports at world prices must equal the value of exports at world prices plus a number of capital income flows. These include net foreign investment, net foreign factor income payments, net foreign remittances, net foreign transfers by the U.S. government, and interest payments to foreigners on the U.S. government debt. These capital income flows are consistent with those distinguished in the National Economic Accounts. They are important in the analysis of domestic policy in that a change to them can have an impact on the price of U.S. exports and imports relative to domestic prices. A change in these prices has an impact on the ROW demand for U.S. exports and the U.S. demand for ROW imports. Changes in exports and imports can have an impact on labor demand by occupation, which may have a bearing on the jobs for low-income households.

Introducing the capital account as an actor into the model adds a market for loanable funds. In a CGE model, investment (demand for loanable funds) is driven by the availability of savings (supply of

loanable funds). Total savings are from households, businesses, government surplus or deficit, and net capital inflows from the ROW. Business savings are from depreciation of capital stocks and retained earnings. Investment is divided between changes in inventory, and the purchase of new capital stocks by industry (fixed investment). The new capital stocks are produced through the purchase of capital goods and construction services.

Factor Markets: Aggregate Factor Supply
Factors of production are land, labor by occupation, and capital. The treatment of aggregate supply for these factors is an important feature of a CGE model, how it responds to a policy shock, and thus the nature of the policy analysis. The aggregate supply of land for agriculture is fixed. Labor supply is either fixed or an endogenous outcome of household decisions. In the medium-run, the aggregate supply of capital is fixed, but stocks of capital may be reallocated among sectors of production. In the long-run, aggregate capital stocks adjust to maintain the return to capital at its original level. This treatment of capital stocks is for comparative statics, where investment is not linked to the change in capital stocks. In a dynamic model, this linkage between investment and capital stocks would be a part of the modeling framework.

Model Closure: Macroeconomic Constraints on Economic Activities
Model closure pertains to the three major macroeconomic balances or accounting identities which hold true for any macroeconomic or economywide model (Robinson, 1989; Arora and Dua, 1993). First, for the government account, the closure identity is that revenue less expenditure equals the surplus (deficit if negative). Second, for the trade account, which pertains to the relationship of the United States with the Rest-Of-World, there is the condition that exports less imports (balance of trade) equals the net value of capital income moving into and out of the United States. So, if imports are greater than exports as is the case for the United States at this time, then there will be a net capital income flow into the United States (net foreign investment). Third, for the capital account, savings equal investment.

Households and firms also have accounting identities that must hold true under all circumstances, which are similar to these aggregate balances for the government, trade, and capital accounts. These conditions are the household budget constraint and the accounting equation for firm net income or profit. They are not treated as part of model closure because

they are incorporated into the model's specification of how households and firms make economic choices. There is a qualitative difference in how the three aggregate balances are treated relative to the treatment of household and firm budget constraints. Closure for the trade balance summarizes the complex process by which the balance of trade, exchange rate, and foreign capital income flows adjust. Similarly, for the savings-investment balance, it summarizes a complex system of financial transactions whose net result is the allocation of financial savings into physical investment for residential and nonresidential structures and into equipment by industry and government. As for the government, the choice underlying closure is a political choice. In none of these cases is there an explicit set of behavioral equations attempting to describe the outcome of these complex market and political processes. Instead, it has proven expedient to introduce closure rules pertaining to these major aggregate balances.

Closure rules determine how the government, trade, and capital accounts adjust to maintain their accounting identity, in response to changes in economic activity. There are a number of possible closure rules, which influence the way a policy change works through the economy. The appropriate choice will depend on the type of policy change being considered. The following set of closure rules are used for the analysis of welfare reform under alternative macroeconomic scenarios: (i) endogenous federal government deficit and fixed personal income tax rates; (ii) exogenous real investment and endogenous household savings; (iii) endogenous trade balance and exogenous exchange rate.

Finally, a CGE model only solves price changes relative to a fixed price index, or numeraire. There is a choice among a number of possible price indices to fix. For policy applications we generally fix a domestic price index. All price changes are measured relative to the specified numeraire, so, if domestic prices fall for some goods they will rise for other goods. For instance, a policy change that reduces food stamps will reduce the demand for food, which will lead to a fall in food prices. If the government transfer funds are shifted into a tax rebate to households there is an increase in demand for nonfood goods and services. The greater demand will lead to higher prices for these goods and services. The fixed numeraire forces the weighted-average price change across commodities to balance out. Switching numeraires does influence how price changes are passed through the modeled economy,

still the final impact as measured by the change in household real income remains the same.

Model Parameters: Data and Calibration
The database underlying a CGE model consists of a Social Accounting Matrix (SAM), quantity measures for factors of production (labor, capital and land), and elasticity parameters. For this analysis, we are using a 1996 SAM developed and maintained at USDA-ERS. We chose 1996 as our base year for policy analysis because it was the last year prior to welfare reform and all data was available at the start of the project. Appendix B provides information on data sources for the SAM.

A SAM is a system of double-entry accounting that organizes data on the transactions and transfers among the economic entities making up the economy. The SAM entries consist of income flows and expenditures that occur throughout the economy among; households, firms, government, rest-of-world, and the capital account. The SAM depicts the structure of an economy for a particular year in a format that ensures consistency between income and expenditures, and is consistent with the National Economic Accounts, which were originally developed for macroeconomic models. For a CGE model, a SAM expands upon the National Economic Accounts using micro-survey data. For instance, households are segmented into household groups, while labor is segmented into occupations distinguished by skill.

The policy and share parameters for the equations in the CGE model are calibrated so that the equations characterizing household and firm behavior are consistent with economic theory, the SAM database, and elasticity parameters. The elasticity parameters are taken from the literature, and treated as "best judgment" values. An experiment is conducted by changing one or more parameters and resolving the system of equations.

References
Adams, Philip D. and Peter B. Dixon. "Generating Detailed Commodity Forecasts from a Computable General Equilibrium Model." *International Journal of Forecasting*, 13, 1997, 223-36.

Adelman, Irma, and Sherman Robinson. *Income Distribution Policies in Developing Countries.* Stanford Calif.: Stanford University Press, 1978.

Arora, Harjit K. And Pami Dua, "Budget Deficits, Domestic Investment, and Trade Deficits," *Contemporary Policy Issues*, 11, June 1993, 29-44.

Balistreri, Edward J., Christine A. McDaniel, and Eina V. Wong. "An Estimation of U.S. Industry-Level Capital-Labor Substitution Elasticities: Cobb-Douglas as a Reasonable Starting Point?" U.S. International Trade Commission, Office of Economics Working Paper No. 2001-12-A, December 2001.

Ballard, C. L., and J. H. Goddeeris. "Financing Universal Health Care in the United States: A General Equilibrium Analysis of Efficiency and Distributional Effects," *National Tax Journal,* 52, 1999, 31-51.

Ballard, Charles L., Don Fullerton, John B. Shoven, and John Whalley. *A General Equilibrium Model for Tax Policy Evaluation*. Chicago: University of Chicago Press, 1985.

Bernat, G. Andrew Jr., and Kenneth Hanson (1995). "Regional Impacts of Farm Programs A Top-Down CGE Analysis." *Review of Regional Studies*, 25(3), 1995, 331-50.

Blitzer, Charles R., Peter B. Clark, and Lance Taylor. *Economy-Wide Models and Development Planning*. New York: Oxford University Press, 1975.

Bosworth, Barry and Gary Burtless. "Effects of Tax Reform on Labor Supply, Investment and Saving." *Journal of Economic Perspectives*, 6(1), Winter 1992, 3-25.

Bosworth, Barry, Gary Burtless, and John Sabelhaus. "The Decline in Saving: Evidence from Household Surveys." *Brookings Papers on Economic Activity*, 1:1991, 183-256.

Brooke, Anthony, David Kendrick, and Alexander Meeraus. *GAMS: A User Guide*. San Francisco, CA: Scientific Press, 1992.

Burfisher, Mary, and Elizabeth A. Jones, eds., *Regional Trade Agreements and U.S. Agriculture*. U.S. Department of Agriculture, Economic Research Service, Agricultural Economic Report No. 771, November 1998.

Canning, Patrick, and Marinos Tsigas. "Regionalism, Federalism, and Taxation: A Food and Farm Perspective." U.S. Department of Agriculture, Economic Research Service, Technical Bulletin No. 1882, March 2000.

Darwin, Roy, Marinos Tsigas, Jan Lewandrowski, and Anton Raneses. "World Agriculture and Climate Change: Economic Adaptations." U.S. Department of Agriculture, Economic Research Service, Agricultural Economic Report No. 703, June 1995.

De Melo, Jaime and David Tarr. A General Equilibrium Analysis of U.S. Foreign Trade Policy. Cambridge MA: MIT Press, 1992.

Dervis, Kemal, Jaime De Melo, and Sherman Robinson. *General Equilibrium Models for Development Policy*. New York: Cambridge University Press, 1982.

Despotakis, Kostas A. and Anthony C. Fisher. "Energy in a Regional Economy: A Computable General Equilibrium Model for California." *Journal of Environmental Economics and Management,* 15, 1988, 313-30.

Despotakis, Kostas A. "Economic Performance of Flexible Functional Forms: Implications for Equilibrium Modeling." *European Economic Review* 30, 1986, 1107-43.

Devarajan, Shantayanan, and Delfin S. Go. "The Simplest Dynamic General-Equilibrium Model of an Open Economy." *Journal of Policy Modeling*, 20(6), 1998, 677-714.

Devarajan, Shantayanan, and Sherman Robinson. "The Impact of Computable General Equilibrium Models on Policy." International Food Policy Research Institute, TMD Discussion Paper No. xx, (forthcoming as of April 2002).

Dixon, Peter, Brian Parmenter, John Sutton, and David Vincent. *ORANI: A Multisectoral Model of the Australian Economy*. Amsterdam: North Holland, 1982.

Dixon, Peter B., Brian R. Parmenter, Alan A. Powell, and Peter J. Wilcoxen. *Notes and Problems in Applied General Equilibrium Economics*. New York: North Holland, 1992.

El-Said, Moataz, Hans Lofgren, and Sherman Robinson. "The Impact of Alternative Development Strategies on Growth and Distribution: Simulations with a Dynamic Model for Egypt." International Food Policy Research Institute, TMD Discussion Paper No. 78, September 2001.

Engen, Eric, Jane Gravelle, and Kent Smetters. "Dynamic Tax Models: Why They Do the Things They Do." *National Tax Journal*, 50(3), September 1997, 657-82.

Francois, Joseph F. And Kenneth A. Reinert, eds., *Applied Methods for Trade Policy Analysis: A Handbook*. New York: Cambridge University Press, 1997.

Gallaway, Michael P., Christine A. McDaniel, and Sandra A. Rivera. "Industry-Level Estimates of U.S. Armington Elasticities." U.S. International Trade Commission, Office of Economics Working Paper No. 2000-09-A, September 2000.

Ginsburgh, Victor A. and Michiel Keyzer. *The Structure of Applied General Equilibrium Models*. Cambridge: MIT Press, 1997.

Granger, C.W.J. *Forecasting in Business and Economics*, second edition. Boston: Academic Press, Inc., 1989.

Hamermesh, Daniel S. *Labor Demand*. Princeton NJ: Princeton University Press, 1993.

Hanson, Kenneth, Elise Golan, Stephen Vogel, and Jennifer Olmsted. "Tracing the Impacts of Food Assistance Programs on Agriculture and Consumers: A Computable General Equilibrium Model." Economic Research Service, U.S. Department of Agriculture, Food Assistance and Nutrition Research Report Number 18, April 2002. http://www.ers.usda.gov/publications/fanrr18

Hanson, Kenneth and Elise Golan. "Issues in Food Assistance: Economywide Effects from Changes in Food Stamp Expenditures." U.S. Department of Agriculture, Economic Research Service, Food Assistance and Nutrition Research Program Number 26-x, (forthcoming, late June 2002).

Hanson, Kenneth and Karen S. Hamrick. "The Impact of Welfare to Work on the Labor Market: Implications for the Food Stamp Program." U.S. Department of Agriculture, Economic Research Service, Manuscript under revision, June 2002.

Hanson, Kenneth, and Sherman Robinson, "Data, Linkages and Models: U.S. National Income and Product Accounts in the Framework of a Social Accounting Matrix." *Economic Systems Research*, 3(3), 1991, pp. 215-232.

Hanson, Kenneth, Sherman Robinson, and Gerald Schluter. "Sectoral Effects of a World Oil Price Shock: Economywide Linkages to the Agricultural Sector." *Journal of Agricultural and Resource Economics*, 18(1), 1993, 96-116.

Hanson, Kenneth, Sherman Robinson, and Stephen Tokarick (1993). "U.S. Adjustment in the 1990's: A CGE Analysis of Alternative Trade Strategies." *International Economic Journal*, 7(2) 1993:27-49.

Hanson, Kenneth, Stephen Vogel, and Sherman Robinson. "Sectoral and Economywide Impacts of Eliminating the Export Enhancement Program." Economic Research Service, U.S. Department of Agriculutre. Staff Paper No. AGES-9529. November 1995.

Hecker, Daniel E. "Occupational Employment Projections to 2010." *Monthly Labor Review*, Nov. 2001, 57-84.

Hertel, Thomas W., ed., *Global Trade Analysis: Modeling and Applications*. New York: Cambridge University Press, 1997.

Ho, Mun S. And Dale W. Jorgenson. "Trade Policy and U.S. Economic Growth." *Journal of Policy Modeling*, 16(2), 1994, 119-46.

Johansen, Lief. *A Multi-Sectoral Study of Economic Growth*. Amsterdam: North-Holland, 1960.

Joint Committee on Taxation. "Joint Committee on Taxation Tax Modeling Project and 1997 Tax Symposium Papers." (JCS-21-97), November 1997.

Jorgenson, Dale W. and Kun Yun. "The Excess Burden of Taxation in the United States."

Journal of Accounting, Auditing, and Finance, Fall 1991, 487-509.

Jorgenson, Dale W., and Peter Wilcoxen. "Energy, the Environment, and Economic Growth." In A. V. Kneese and J. L. Sweeney, eds., *Handbook of Natural Resource and Energy Economics, Vol III.* New York: North-Holland, 1993.

Kilkenny, Maureen and Sherman Robinson. "Computable General Equilibrium Analysis of Agricultural Liberalization: Factor Mobility and Macro Closure." *Journal of Policy Modeling*, 12(3), 1990, 527-56.

Lofgren, Hans, Rebecca Lee Harris, and Sherman Robinson. "A Standard Computable General Equilibrium (CGE) Model in GAMS." International Food Policy Research Institute, TMD Discussion Paper No. 75, May 2001.

Mauskopf, Eileen and Dave Reifschneider. "Dynamic Scoring, Fiscal Policy, and the Short-run Behavior of the Macroeconomy." *National Tax Journal*, 50(3), September 1997, 631-54.

McDaniel, Christine A. and Edward J. Balistreri. "A Discussion on Armington Trade Substitution Elasticities." U.S. International Trade Commission, Office of Economics Working Paper No. 2002-01-A, January 2002.

McKibbin, Warwick J. and Peter J. Wilcoxen. "The Theoretical and Empirical Structure of G-Cubed." *Journal of Policy Modeling*, 1995.

Powell, Alan, and Richard Snape. "The Contribution of Applied General Equilibrium Analysis to Policy Reform in Australia." *Journal of Policy Modeling*, 15(4), 1993, 393-414.

Randolph, William C. And Diane Lim Rogers. "The Implications for Tax Policy of Uncertainty About Labor Supply and Savings Responses." *National Tax Journal*, 48, September 1995, 429-46.

Robinson, Sherman. "Mulitsectoral Models." In H Chenery and T. N. Srimivasan, eds., *Handbook of Development Economics, Vol. II.* New York: North-Holland Pub., 1989.

Robinson, Sherman, Maureen Kilkenny, and Irma Adelman. "The Effect of Trade Liberalization in Agriculture on the U.S. Economy: Projections to 1991." In A. Stoeckel, S. Cuthbertson, and D. Vincent, eds., *Macroeconomic Consequences of Farm Support Policies.* Durham, NC: Duke University Press, 1989.

Robinson, Sherman, Maureen Kilkenny, and Kenneth Hanson. "The USDA/ERS Computable General Equilibrium (CGE) Model of the United States." Economic Research Service, U.S. Department of Agriculture. Staff Report No. AGES 9049, June 1990.

Seskin, Eugene P., and Robert P. Parker. "A Guide to the NIPA's." *Survey of Current Business*, March 1998, pp. 26-68.

Shoven, John B., and John Whalley. "Applied General Equilibrium Models of Taxation and International Trade: An Introduction and Survey." *Journal of Economic Literature*, 22, September 1984, 1007-1051.

Shoven, John B., and John Whalley. *Applied General Equilibrium.* New York: Cambridge University Press, 1992.

Taylor, Lance. Ed., *Socially Relevant Policy Analysis: Structuralist Computable General Equilibrium Models for the Developing World.* Cambridge MA.: MIT Press, 1990.

U.S. Department of Agriculture, Food and Nutrition Service. "Characteristics of Food Stamp Households: Fiscal Year 1996." March 1998.

U.S. Department of Agriculture, Food and Nutrition Service. "Characteristics of Food Stamp Households: Fiscal Year 2000." October 2001.

U.S. Department of Commerce, Bureau of Census, Current Population Survey, March 1997: Technical Documentation. Washington DC: GPO, 1997a.

U.S. Department of Commerce, Bureau of Census, Current Population Survey, 1997: Technical Documentation. Washington DC: GPO, 1997b.

U.S. Department of Commerce, Bureau of Economic Analysis, An Introduction to National Economic Accounting. Methodology Paper Series, MP-1. Washington DC: GPO, 1985.

U.S. Department of Commerce, Bureau of Economic Analysis, "Benchmark Input-Output Accounts for the U.S. Economy, 1992." *Survey of Current Business*, 77(11), November 1997.

U.S. Department of Commerce, Bureau of Economic Analysis, "National Income and Product Account Tables." *Survey of Current Business*, 79(12), December 1999.

U.S. Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation. "Initial Synthesis Report of the Findings from ASPE's "Leavers" Grants." Report prepared by Gregory Acs and Pamela Loprest of the Urban Institute, January 2001. http://aspe.hhs.gov/hsp/leavers99/synthesis.

U.S. Department of Labor, Bureau of Labor Statistics. *Consumer Expenditure Survey: Integrated Diary and Interview Survey Data, 1996*. Washington DC: GPO, 1997.

U.S. Department of Labor, Bureau of Labor Statistics. *Occupational Projections and Training Data*. Bulletin 2501, January 1998.

U.S. Department of Labor, Bureau of Labor Statistics. *Employment and Earnings*, 48(1), January 2001.

U. S. International Trade Commission. The Economic Effects of Significant U.S. Import Restraints: Second Update 1999. Investigation No. 332-325, Publication 3201, May 1999. http://www.usitc.gov

Vogel, Stephen and Kenneth Hanson. "Falling Prices and National Farm Policy: The Northern Great Plains." Agricultural Outlook, June-July, 2001, AGO-282, pp. 27-31.

Wash, Darrel Patrick, "A New Way To Classify Occupations by Education and Training," *Occupational Outlook Quarterly*, U.S. Department of Labor, Bureau of Labor Statistics, Winter 1995-96, pp. 29-40.

Appendix A: Placing the USDA-ERS U.S. CGE Model in Context of Other Models

CGE models are a natural application of general equilibrium theory and welfare analysis [Shoven and Whalley (1992)]. Early applications emphasized development policy and were generally limited to linear programming models or constrained input-output models [Blitzer, Clark, and Taylor (1975)]. Johansen's (1960) multi-sector growth model of the Norwegian economy was a pioneering development by using a linear approximation to a nonlinear model. In the late 1970's, the Industry Assistance Commission of Australia supported the development of a Johansen type CGE model to help identify losers, as well as winners, from proposed policy measures [Dixon, et al. (1982)]. Powell and Snape (1993) document the long and successful career of this project in Australia. More recent versions of the model have allowed nonlinear solutions to add greater flexibility [Dixon, et al. (1992)].

In the late 1970's, two schools of CGE models emerging in the United States developed the use of nonlinear solution algorithms, allowing for greater flexibility in behavioral responses to exogenous shocks [Shoven and Whalley (1984)]. Tax CGE models focused on tax incidence in neoclassical models for developed countries [Ballard, Fullerton, Shoven and Whalley (1985), Shoven and Whalley (1992)]. Development CGE models focused on trade and distribution effects of policies facing developing countries [Adelman and Robinson (1978), Dervis, DeMelo and Robinson (1982), Robinson (1989), Taylor (1990), Lofgren, Harris, and Robinson (2001)]. Both are static, single-country models.

Today, as an emerging field within applied policy analysis, the scope of CGE modeling has exploded beyond the confines of any particular school or economic sub-discipline. The proliferation of model development has been enhanced by advancements in numerical solution algorithms, conceptual devices bridging the mathematics of economic theory and numerical analysis, availability of software facilitating model development, increased computational power of personal computers, and the increased availability of data. Furthermore, a greater interest in the economywide implications of policy alternatives — the ability to determine winners and losers from a policy change — has fueled the demand for this kind of applied general equilibrium analysis [and Ginsburgh and Keyzer (1997), Devarajan and Robinson (2002)].

The structure of CGE models can be distinguished by two general characteristics. First, does it encompass a single country or a set of countries? Multi-country models are used for addressing issues of regional and multilateral trade policy, where the relative impact among a group of countries is of interest [Hertel (1997), Burfisher and Jones (1998)]. Single-country models allow for finer levels of disaggregation and greater complexity in tailoring models to domestic policy issues [Adams and Dixon (1997), Ballard and Goddeeris (1999), de Melo and Tarr (1992), Francois and Reinert (1997), Taylor (1990), U. S. International Trade Commission (1999)].

Second, is the model an exercise in comparative statics or dynamics? While comparative static models solve within a single period, dynamic models obtain solutions for multiple periods of time. Within the class of dynamic models, these CGE models differ by whether the dynamics involve intertemporal optimization or are captured by sequentially linking single period optimizations. Dynamic models involving intertemporal optimization are useful when policy changes have an impact on the intertemporal decisions of households and firms related to labor supply, savings and investment, which have an impact on economic growth. Growth effects occur when a trade policy change has an impact on foreign capital flows which affect investment [Devarajan and Go (1998), Ho and Jorgenson (1994), Mckibbin and Wilcoxen (1995)]. They occur when a change in tax policy has an impact on the labor supply and savings decisions of households and the investment decisions of firms [Bosworth and Burtless (1992), Engen, Gravelle and Smetters (1997); Joint Committee on Taxation (1997), Jorgenson and Yun (1991), Mauskopf and Reifschneider (1997), and Randolph and Rogers (1995)]. Environmental policy can also have an impact on growth in that investment for abatement may reduce investment for capacity expansion [Jorgenson and Wilcoxen (1993)].

'Quasi-dynamic' models, or sequential dynamic models, capture only some dynamic impacts of a policy change [El-Said, Lofgren, and Robinson (2001)]. In these models, changes in labor supply and investment are added to initial endowments, leading to growth effects. However, these models do not account for the impact of intertemporal decisions made by households or firms. The tradeoff between these two types of dynamic models represents a continuum between the extremes of needing a high level of dynamic general equilibrium theory and a detailed institutional structure of economic activity.

The USDA-ERS CGE model of the United States is a static, single country model, but with considerable detail on households and labor. This allows the model to be used in analysis of government programs to assist low-income households in addition to farm and trade policy.

Appendix B: SAM Data Sources
The data underlying the CGE model are derived from a number of sources. Development of the data into a consistent SAM database for the CGE model involved considerable data processing.

Macroeconomic data and Input-Output Accounts
Aggregate economywide data were taken from the National Income and Product Account (NIPA) [U.S. Department of Commerce, Bureau of Economic Analysis (1999)]. We used 1996 as the base year. Data on aggregate household consumption, other components of final demand, industry production, and interindustry transactions are derived from the 1992 Input-Output (IO) Account [U.S. Department of Commerce, Bureau of Economic Analysis (1997)]. We update the 1992 IO Account to 1996 using NIPA data and detailed commodity data for final demand. The IO Account is constructed every five years with a seven-year lag, and includes about 500 industries and commodities, but only one household aggregate. For our CGE model, we aggregated these industries to about 50 with considerable detail on the farm and food processing sectors, and disaggregated households into 99 categories. We also included data on labor demand by industry and supply by households for a set of occupational categories.

Household sources of income - CPS March Supplement
The 1996 data for household income are from the 1997 Current Population Survey (CPS) March Annual Demographic Supplement [U.S. Bureau of Census (1997a)]. Every year, the March CPS includes supplemental questions on sources and amounts of money received during the previous calendar year, in this case, 1996. The March Supplement survey of the CPS is written and conducted by the Bureau of the Census, which also releases the data. For more information on the March Supplement data, see *http://www.bls.census.gov/cps/cpsmain.htm* .

Households receive income from three main sources: earnings from wages, salaries, and self-employment; capital income from the ownership of assets–dividends, interest, and rent; and transfer income

from government programs. Eleven government transfer programs are distinguished and characterized as pre-welfare reform. Total income by each source, including government transfers, are adjusted for consistency with values reported in the National Economic Accounts. The CPS tends to undercount income by each source except earnings.

Household expenditures - Consumer Expenditure Survey

Households consume goods and services, pay taxes, and save. Household expenditure shares are derived from the 1996 Consumer Expenditure Survey [U.S. Department of Labor, Bureau of Labor Statistics (1997)]. For more information, see *http://stats.bls.gov/csxhome.htm.*

Savings and taxes are specified as fixed saving rates and tax rates specific to each household group. Tax rates are derived from the CPS March Supplement, while savings rates are derived from the Federal Reserve, Survey of Consumer Finances, as presented in Bosworth, Burtless, and Sabelhaus (1991).

Household work status and labor supply - CPS Earnings File

Data on labor force participation are from the 1996 CPS earnings file [U.S. Department of Commerce, Bureau of the Census (1997)]. The data provide detailed information on the labor force, employment, unemployment, and demographic characteristics of the population. The CPS derives estimates based on interviews of a national sample of about 47,000 households that are representative of the U.S. civilian noninstitutional population 16 years of age and over. Labor force information is based on respondents' activity during 1 week each month. The Earnings file includes information on earnings of the one-quarter of the sample each month that is in the outgoing rotation of the panel sample. These respondents are asked about the usual hours worked and earnings on their sole or primary job. The 1996 data file contains information on almost 430,000 persons. For more information on the CPS data, see *http://stats.bls.gov/cpshome.htm.*

Using the CPS Earnings file we identified and grouped families. We used the BLS definitions of family–"a group of two or more persons residing together who are related by birth, marriage, or adoption"–and added single-person families. A single-person family may not necessarily be a single person household. Only primary families are included, where the primary family contains the householder. Ninety-five percent of observations of

working-aged persons are included in our definition. Excluded are the spouse-equivalent in an unmarried couple and secondary families.

Hourly earnings are computed by dividing usual weekly earnings by usual weekly hours; included are tips, overtime, and commissions. In a married couple where both spouses are employed, the primary earner was defined as the earner with the larger weekly earnings. For the status of no work, we distinguish unemployed from not in the labor force (NILF). For NILF, we distinguish the reasons as retired, disabled, and other. The labor supplied by each household is classified into the education and training occupational categories.

Education and training occupational categories

We grouped occupations into the 11 education and training categories developed by the U.S. Department of Labor [Wash (1995-96)] and used in the occupational employment projections [U.S. Department of Labor, Bureau of Labor Statistics (1998); Hecker (2001)]. Using these education and training occupational groups required a mapping from the detailed occupations used in the CPS earnings file for household labor supply and from the detailed occupations used in the Occupational Employment Statistics (OES) for industry labor demand.

Industry labor demand - Occupational Employment Statistics

The Occupational Employment Statistics (OES) program is a survey of nonfarm establishments on nonfarm wage and salary workers conducted by the U.S. Department of Labor, Bureau of Labor Statistics. A sample of 1.2 million establishments is in the survey, with 400,000 surveyed each year, taking 3 years to collect the full sample. The reference month of the data is October, November, or December, depending on the industry. The survey covers full-time and part-time wage and salary workers, and estimates are produced for over 750 occupations in 400 industry classifications. We used data from the 1996 survey to group the detailed occupations into the 11 education and training categories defined above. For more information on the OES data, see *http://stats.bls.gov/oeshome.htm* .

Industry Fixed Capital Stocks and Land in Agriculture

The Bureau of Economic Analysis (BEA) provides data on Fixed Reproducible Tangible Wealth. This database includes data on capital stocks by industry,

depreciation of capital stocks by industry, and investment by destination or industry receiving the new capital stocks. The data are available on the BEA website (http://www.bea.gov), and are described in the *Survey of Current Business* (September 1998, September 1997, May 1997, and July 1997). Land data for the agricultural crop sectors are specified with data on harvested acres by crop from *Agricultural Statistics*, with a few minor crops extrapolated from the Agricultural Census.

Elasticities
The elasticity parameters are taken from the literature, and treated as "best judgment" values. There are five sets of elasticities: 1) producer input substitution elasticities, 2) household commodity demand (price and income) elasticities, 3) household labor supply (wage and income) elasticities, 4) import-domestic goods substitution elasticities, and 5) export-domestic sales substitution elasticities.

For producers, intermediate goods are used in fixed proportion to production, while factor substitution in production is characterized with a Constant Elasticity of Substitution (CES) value-added production function. A CES elasticity of substitution is specified for aggregate labor, capital, and land in agriculture. These elasticities are set at a value slightly greater than one for manufacturing, and less than one for services. Aggregate labor is a CES function of labor by occupation with an elasticity of substitution around one-half. Evidence for this treatment of producer elasticities is in Balisteri, et al. (2001), where they find that Cobb-Douglas is a reasonable starting point (CES elasticity of substitution equal to one). For an argument for flexible functional forms see Despotakis (1986). The CES substitution elasticities can also be adjusted to be consistent with the labor demand elasticities reported in Hamermesh (1993).

Where the scenario analysis relates to specific sectors or to specific inputs such as energy, then special treatment of the elasticities for these sectors or inputs are necessary [Despotakis (1988), Hanson, Robinson, and Schluter (1993)]. In our analysis of farm and food policies, the agricultural crop sectors take on a greater importance. For the crop sectors, we use a Constant Elasticity of Transformation (CET) land allocation function to control the response of agriculture to price changes. Traditionally, the production response of agricultural crops have been inelastic, and we calibrate the model response accordingly (CET land elasticity of 0.5 to 0.7).

Household price and income elasticities for expenditures across goods and services are characterized with a Linear expenditure System (LES), which is extended to include leisure so that the model can include a household's labor supply response to wage and income changes. The specification of these elasticities is discussed in Hanson, et al. (2002).

The import and export trade elasticities have less empirical support for their specification. In general they are considered to be elastic in response, though econometric estimates tend to be smaller and sometimes inelastic [McDaniel and Balistreri (2002)]. Over the years, researchers at the U.S. International Trade Commission have estimated the trade elasticities for the United States [Gallaway, et al. (2000)].

AGRICULTURAL SECTOR SCENARIO ANALYSIS AND THE ERS COUNTRY-COMMODITY LINKED S YSTEM

Ralph Seeley and Paul Westcott, Economic Research Service, USDA

Introduction

Changing conditions in foreign countries and in the United States often affect the agricultural sector. These changes may arise in macroeconomic conditions, agricultural policies, weather, and many other factors, and can be short-term or more permanent. Further, some changes may be small while others may represent significant variations in the structure of markets.

Economists are frequently asked to assess impacts of such changes on agricultural commodity markets as well as impacts on the farm sector as a whole. A useful approach in addressing these analyses is to use simulation models for the sector. This paper discusses one such large-scale model and some of the issues and challenges faced by economists when conducting analyses of changes in the sector.

The Country-Commodity Linked System (CCLS) is a partial-equilibrium agricultural economic modeling system with 24 commodity markets and 44 regions that is used for agricultural sector scenario analyses at the USDA's Economic Research Service (ERS). Scenario analysis using the CCLS often provides interesting challenges to analysts and modelers that involve increasing the capacities of the existing models or modifying the modeling system to address issues beyond the original model design. New approaches in using the modeling system are illustrated, for analysis of China's WTO accession, and for analysis of currency devaluation and financial crisis in Argentina. We begin with an overview of the modeling system and a discussion of general approaches used in scenario analysis.

Modeling system

Country/regional Models

The modeling system contains 44 country or regional models (table 1). A main area of interest in scenario analysis is the impact on the United States, so the system includes a detailed U.S. policy model (Fapsim). Several rest-of-region models and a rest-of-world model handle any countries or commodities missing from the individual country models. Further, for reconciliation with historical data, closure, and introduction of scenario shocks, the system also includes a residual "region" and an exogenous "region."

Table 1. Countries/regions in the Country-Commodity Linked System

Algeria
Argentina
Australia
Bangladesh
Brazil
Canada
Central America & Caribbean
China
Czech Republic
Egypt
European Union
Hong Kong
Hungary
India
Indonesia
Iran
Iraq
New Zealand
Pakistan
Philippines
Poland
Russia
Saudi Arabia
South Africa, Republic of
Japan
Malaysia
Mexico
Morocco
South Korea
Taiwan
Thailand
Tunisia
Turkey
Ukraine
United States (Fapsim)
Vietnam
Rest-of-region models:
    Asia
    Central and Eastern Europe
    Former Soviet Union
    North Africa and Middle East
    South America
    Sub-Saharan Africa
    Western Europe
    Rest-of-world

Other models:
    Residual "region"
    Exogenous "region"

Commodities

The CCLS also includes 24 commodities, such as various grains, animal products, and oilseeds and products (table 2).

Table 2. Commodities in the Country-Commodity Linked System

Coarse grains
    Corn
    Sorghum
    Barley
    Other coarse grains
Food grains
    Wheat
    Rice
Oilseeds
    Soybeans
    Rapeseed
    Sunseed
    Other oilseeds -- in certain models, may be disaggregated into Copra, Cottonseed, Groundnuts, and Sesame seed
Oil meals
    Soymeal
    Rapemeal
    Sunmeal
    Other oil meals -- in certain models, may be disaggregated into Copra meal, Cottonseed meal, Fish meal, Groundnut meal, and Sesame meal
Oils
    Soyoil
    Rapeoil
    Sunoil
    Other oils -- in certain models, may be disaggregated into Coconut oil, Cottonseed oil, Groundnut oil, Palm oil, Sesame oil, and Other tropical oils
Other crops
    Cotton
    Sugar
Animal Products
    Beef and veal
    Pork
    Poultry meat
    Eggs

Not all country/regional models contain all of the commodities shown in table 2, especially if a country supplies and uses a negligible amount of a commodity. In all, the system contains over 16,000 equations for a given year. (This count is approximate, because equations may be separated into individual terms by the modelers, while many equations are combined during the automated linking process.)

Equilibrium

The CCLS links the countries and commodities to solve for global equilibrium in prices and quantities, for each of about 10 projected years.

Formulation

The model behavioral equations typically use a linearized dynamic Cobb-Douglas functional form, which implies that the annual growth rate in a dependent variable equals the sum of annual growth rates in the explanatory variables, scaled by elasticities with respect to (w.r.t.) the explanatory variables. For example,

$[\text{quantity}_t]$
$= [\text{quantity}_{t-1}]$
$* \{1 + [\text{elasticity of quantity w.r.t. price1}]$
    $* ([\text{price1}_t] / [\text{price1}_{t-1}] - 1)$
    $+ [\text{elasticity of quantity w.r.t. price2}]$
    $* ([\text{price2}_t] / [\text{price2}_{t-1}] - 1)\}$

Scenario Analyses

A model **projection** provides a scenario about the future, conditional on assumptions that are specific to the scenario. (In contrast, a **forecast** would be an estimate of the future, based on our best estimate for every relevant assumption.) Scenario analyses address "what-if" questions compared to a reference scenario.

USDA Baseline

The annual USDA baseline is a major departmental project to which the CCLS contributes. The final baseline projections serve as the benchmark or reference projections against which we compare alternative scenarios.

Staff Analyses and Research Applications

Staff analyses are quick-turnaround, applied policy analyses answering specific questions that often are asked by the Administration and Congress. Staff analyses typically have deadlines of hours to weeks. The CCLS also contributes to longer-term research projects. A summary of short-turnaround applications and long-term projects using the CCLS is given in the appendix.

Four Scenario Analysis Steps

We go through four steps when conducting a scenario analysis with the modeling system; we

- analyze,
- implement,
- evaluate, and
- iterate.

Analyze

An analysis request is like a word problem in a math or science course, but often the problem statement is incomplete and data are missing. There is no textbook and no instructor. We are on our own to define fully the scenario to be analyzed, often filling in missing pieces of the underlying question being asked. We analyze the scenario dimensions, thinking of the variables and structure that would be in the ideal model for that scenario. Which variables will be part of the answer, the assumptions, or the mechanisms in between? We try to learn about the relevant institutional and market behavior. We acquire data and parameters. We may break the issue into components and analyze each separately to gain further insights and to make the problem more manageable.

Implement

Once we have analyzed and defined the issue and collected any needed additional data or parameters, we can incorporate the new assumptions into the modeling system. This at times may require modifications to the models. The models will generate an alternative outcome that can be compared to the reference scenario to derive impacts for the projections period.

Evaluate

We then want to evaluate the impacts. Did we get results that can be explained? The question is not whether the customer will like the results, but instead is whether the results make economic sense. The system is complicated, and it is possible to overlook some interactions when introducing a scenario. We may learn about the underlying economic behavior when we study counterintuitive results from this complex system. However, as we evaluate results, we also are evaluating whether the scenario was implemented correctly and completely.

Iterate

With a complicated scenario, some part of the scenario may not be implemented at first in a way that is fully consistent both with the question being addressed and with the model structure. The implementation may need to be fine-tuned and then we go back through the scenario analysis steps again.

Bridging the Gap: Scenario « Model

Typically, a model will not have all of the variables and behavioral relationships needed to handle a complex scenario directly. One way to bridge the gap is to translate the scenario impacts into terms of the existing model variables and structure. Alternatively, if time

permits, the model may be enhanced to handle the scenario more directly.

A number of approaches may be used to bridge the gap from the scenario definition to the model structure.

Borrow and Reuse

If time does not permit constructing a new model feature from scratch, it may be possible to borrow parameters, equations, and descriptions of behavior from comparable countries, commodities, and circumstances. Insight may be gained from historical situations with similar patterns.

To put into perspective this discussion of making do under deadlines, consider the words of Charles Babbage, sometimes called the father of modern computing: "Errors using inadequate data are much less than those using no data at all."

Economic Theory and Mathematics

Theory and math may assist with bridging the gap between scenario and model. As will be discussed later, we assumed in our analysis of China WTO accession that the new market participants (non-State Trading Enterprises) would behave like residual trade equations. That assumption allowed calculation of parameters to describe the non-State Trading Enterprises.

Translate Scenario

Another technique is to transform a scenario into the existing model dimensions. This works especially well if the model depends on widely applicable variables.

For instance, suppose that farmers are modeled as responding to net returns. Then, a proposed new policy can be analyzed, given calculation of its effect on net returns, by using the existing model structure and its responsiveness to net returns. The use of such a widely applicable variable makes usable many years of historical data.

For example, when monetary compensation payments to farmers were instituted in the European Union, we were able to simplify modeling of the effects of the payments on area by converting the payments into price equivalents. Assuming that farmers would respond to gross receipts in the same way that they had responded to market receipts, allowed reuse of the original elasticities, which were based on more observations.

In the Argentina crisis analysis, to be described later, the scenario macroeconomic and policy shocks were translated into cost indices, which were transformed into changes in various components of production costs and receipts, which ultimately affected areas and yields for crops.

Find Proxies

If the ideal variable cannot be measured, there may be some available variable that would be an adequate substitute.

As will be shown in more detail later, in the Argentina economic crisis analysis, we used the cost of production for each crop as a proxy for the reluctance of creditors to make loans, given changing contract rules.

Range of Assumptions

If a key parameter or series is unavailable, another strategy is to use a range of assumptions about the parameter, trying to bracket the plausible outcomes. In other words, present an alternative scenario for each level of the parameter. This approach works better with customers who understand the overall situation being modeled. When a variable's level is not known precisely, it is especially important that it be entered into equations as a variable, and not just as a number. If a better estimate of the variable becomes available, then that revised level should affect all of the relevant equations automatically.

Economic and Other Interrelationships

It helps to remain aware of interactions between variables, both when a scenario is being implemented, and when the results are being checked. We do not want to focus so hard on supply that we forget to look at demand. The same is true for quantities versus prices.

For example, when we implemented a "mad cow disease" (BSE) scenario for the European Union (EU), the first temptation was to reduce the preference for beef consumption directly. However, in our EU model, cross-commodity effects are evident primarily through prices. A direct shock to beef demand would have reduced the beef price, and caused pork demand to decline, through the positive elasticity of pork demand with respect to the beef price. The solution was to add a distinction between beef producer and consumer prices, and to introduce a gap between those two endogenous prices.

Aggregates « Details

We need to go back and forth between aggregates and details. Aggregate measures that do not balance or are beyond reasonable bounds can signal an unrealistic, infeasible answer, or provide an opportunity to apply constraints. Alternatively, laying out details may allow incorporation of micro data to build up to a solid answer.

Examples of aggregates are total calorie and protein requirements and availability, and total crop area. In the Argentina analysis, the addition of cost-of-production

detail allowed us to calculate the impact of scenario assumptions on net returns by crop.

Model Application Examples

Two examples are discussed to illustrate applications of the CCLS and the challenges faced by economists in conducting scenario analyses.

China's Accession into the World Trade Organization (WTO)

China joined the WTO in December 2001, too late to be included in the USDA 2002 baseline assumptions, so we needed to examine the implications of this major change in China's policies. When we divided the accession scenario into separate analytical pieces, it became evident that we would need to model the behavior of the new market participants, non-state trading enterprises (non-STE's).

State Trading Enterprises

The state trading enterprises (STE's) in China are governmental entities that, in the past, were given the licenses to import commodities into the country. The existing STE's were modeled using import functions that reflect institutional trade restrictions by applying small coefficients to the gaps between domestic and foreign prices:

$$\text{Imports}_t$$
$$= \text{intercept}$$
$$+ \text{coefficient} * (\text{domestic price}_t - \text{import price}_t)$$

The domestic price for each commodity adjusts to set supply and use equal.

Non-STE Market Clearing Behavior Assumed

Under the WTO agreement, non-STE's are to be given a portion of import licenses, but the behavior of the non-STE's is not yet clear. To allow modeling of non-STE behavior, which is expected to be more price-responsive, we assumed that the non-STE's would behave competitively, maximizing profits. In other words, we assumed that the non-STE's would not be an impediment to trade, that they would move a commodity from a low-price location to a high-price location, until the price gap no longer exceeded the cost of transferring the commodity. Market-driven non-STE behavior would imply that the non-STE's would clear markets. This behavior would be computationally comparable in the model to making imports the residual variable in the supply-use balance.

Simulating Residual Imports
If the non-STE behavior could be described as residual imports, then we could use the China model itself to simulate the price responsiveness of residual imports, which could serve as a proxy for non-STE price responsiveness.

We ran two scenarios with alternative prices, and obtained residual imports.

1. For each commodity expected to be affected, the import function was changed temporarily, from price-responsive to residual behavior:

   Imports$_t$ =
       Exports$_t$ + Consumption$_t$ + Ending stocks$_t$
       - Beginning stocks$_t$ - Production$_t$

2. The producer price was set equal to the border price, the model was solved, and the residual import level was saved.

3. The producer price was set equal to the border price times 1.10, the model was solved again, and the second residual import level was saved.

4. The original import and producer price equations were restored.

Total Imports
The import level at the border price gave the intercept for the new non-STE import function. The change in imports over the change in price gave the slope. Because import licenses unused by the STE's are to be reallocated to non-STE's in the last quarter of each year, the effective fraction of trade described by non-STE behavior was increased after consideration of transportation capacity. The total import function was the composite of the non-STE and STE behavior.

Total imports =
    non-STE imports * % of licenses for non-STE's
    + STE imports * (1 - % of licenses for non-STE's)

Argentine Economic Crisis: Macroeconomy and Policies

Modeling Argentina's current economic situation required handling shocks to Argentina's macroeconomy, such as Peso devaluation, and shocks to government policies, in particular to export taxes. (Modeling of limitations on the availability of credit is discussed in a later section.) Although the model responded in part to macroeconomic variables such as the exchange rate, the model needed to be enhanced to allow more complete responsiveness to the shocks in the scenario. Cost-of-production (COP) detail for crop year 2000/01 was added to the model, but it was necessary to allow the COP numbers to adjust with scenarios and across the projected years.

Macroeconomic and Policy Shocks
Modeling the economic crisis required endogenizing the responses of crop area to exogenous scenario shocks to the
1. GDP,
2. GDP deflator,
3. exchange rate, and
4. export taxes on
   a. processed ag commodities,
   b. primary ag commodities, and
   c. petroleum.

Cost Indexes
First, the exogenous shock series were translated into nominal cost indices, newly incorporated into the model.

Growth(Petroleum cost index)
    = Growth(international petroleum price
    * U.S. deflator
    * scenario nominal exchange rate
    * [1-scenario oil export tax])

Growth(Other imported input cost index)
    = Growth(scenario nominal exchange rate
    * U.S. deflator)

Growth(Wage cost index)
    = Growth(scenario real GDP / population
    * scenario domestic deflator)

Growth(Other domestic input cost index)
    = Growth(scenario domestic deflator)

Constructing Cost-of-production Changes
Changes in the cost indices above were used to derive changes for line items in the COP for a given crop. Shares were used to indicate the extent to which each cost index matters for a given line item in the COP. For example, the share of petroleum in the cost of planting corn is different from the share of petroleum in the cost of harvesting corn. In addition, cost shares for a particular input are different across different crops. The shares were based on COP data and on judgement.

Cost of production $_{2000}$ = Base period COP per hectare

Growth(Cost of production $_t$)

    = [Petroleum share
     * Growth(Petroleum cost index $_t$)

    + Other imported input share
     * Growth(Other imported input cost index $_t$)

    + Wage share
     * Growth(Wage cost index $_t$)

    + Other domestic input share
     * Growth(Other domestic input cost index $_t$)]

    * [Yield share
     * Growth(Yield $_t$)]

Costs of Production

The COP for a typical crop included line items such as
1. Production costs
    a. Plowing, planting
    b. Seed
    c. Fertilizer, Pesticides
    d. Interest
    e. Harvesting
2. Marketing costs
    a. Freight
    b. Handling, drying, etc.
    c. Administrative expenses

As mentioned above, each of the line items was endogenized across years and across scenarios.

Area Responses

Crop areas in the Argentina model were made functions of expected costs, based on endogenous production costs calculated as described above. Endogenous marketing costs determined as above were used to enhance the model's description of expected receipts.

However, when base period COP data from various sources were combined with country-wide average yields, it was noted that net returns varied across crops to a greater extent than appeared reasonable. Therefore, for the actual incorporation of endogenous expected receipt and cost information into area equations, net returns were not used directly. Instead, the growth rates for expected receipts and expected costs were kept distinct for each crop. The elasticity applied to growth in expected costs was made half as large (and of the opposite sign) as the elasticity applied to growth in expected receipts, for a given crop's influence on its own area or the area of a competing crop. The rationale for the use of an elasticity ratio of one half was that about half of costs would be fixed costs.

Argentine Economic Crisis: Limited Credit Availability

Anecdotal information from Argentina, obtained by ERS analyst Randy Schnepf, suggests that about half of field crop production is linked to credit, primarily offered by input suppliers. Credit availability in Argentina is expected to be severely limited during the crisis because of government changes in contract rules that have resulted in concern by creditors that they might not be repaid in full.

Time Gap

Farmers have short-term liquidity needs during the growing season, because of the time gap between their costs and their receipts. That time gap is bridged by loans, retained earnings, or other wealth. Any wealth must be distributed over production of the possible crops or animal products, and may need to last until stability is restored, so production that depends on retained earnings may respond similarly to production affected by credit limitations.

Risk of Not Receiving Loan

Interest accounts for the fact that the lender forgoes other allocation of the money lent, and risks not being paid back. Rather than the risk that the creditor would not be paid back, we now consider the risk that the borrower would not get the loan in the first place. The difference between credit cost (which shows up as interest), and credit availability, is analogous to the difference between a tariff and a quota.

Amount per Loan

First, the variable costs for producing a hectare of a given crop were calculated, and combined with the time gaps or "float" between costs and receipts.

Amount per loan = S (cost * months gap/12)

Fraction Borrowed

Second, the fraction of production costs that are borrowed was estimated depending on the cost of producing a given crop. The parameters are based on judgement.

    Fraction borrowed = 0.25 + 0.1 * amount per loan
    / scenario domestic deflator

Credit Availability Proxy

Third, we attempted to account for limited credit availability, by combining a fractional scenario risk premium with the cost of producing each crop. That is, the production cost was used as a proxy for the expected reluctance of creditors to extend themselves by making loans for crops that are more expensive. We

hypothesized that creditors would feel safer making small loans during the current crisis.

Loan size penalty = scenario risk premium
* amount per loan / scenario domestic deflator

Total "Interest" Rate
A total interest rate was calculated from the sum of the underlying interest rate, the inflation rate, a scenario-specific general risk premium, and the commodity-specific loan size penalty. This loan size penalty actually is not an interest rate, since it attempts to capture the risk of no loan being made in the first place. However, it was convenient to treat this proxy as an interest rate for inclusion in the existing model.

Total interest rate = 12 + inflation rate
+ scenario interest rate + loan size penalty

Total "Interest" Payment
The total "interest" payment combines the total loan amount and the total "interest" rate. It should be noted that the amount per loan appears three times in the construction of the total payment. Thus, the more costly crops to produce are penalized significantly, representing the higher probability of not receiving a loan.

Total interest payment = amount per loan
* fraction borrowed * total interest rate

Argentine Economic Crisis: Results

The results discussed here draw from the work of ERS's Argentina Crisis Evaluation Team. Members of the team included Randy Schnepf, Ralph Seeley, Richard Stillman, and David Torgerson. The scenario discussed here is one of several for the project.

Exchange Rates
Because exchange rates rise more rapidly than other costs in the first couple of years of the scenario, receipts keep up with costs. Consequently, projected total scenario crop area does not decline at first.

Expected Net Returns
Corn, being a higher-input commodity than wheat or soybeans, shows a greater decline in net returns than wheat and soybeans, largely through the higher "interest" expense faced, as a proxy for the decline in credit availability.

Crop Area
Corn area fell less than expected and less than soybeans in the intermediate years. With this unexpected result, in the results evaluation stage, we needed to re-examine the initial scenario implementation.

These counterintuitive intermediate-year results turned out to be explainable, being related to the relative sizes of area planted to the different crops, and relative cross price elasticities. Since corn area is relatively small, its elasticity with respect to soybean costs is larger than the elasticity of soybean area with respect to corn costs. The corn area equation is more aware that other crop costs are rising, so corn area does not fall as much. The soybean area equation is relatively oblivious to other crop costs; therefore, soybean area declines more between the reference run and the scenario in the intermediate years.

Concluding Comments

The Country-Commodity Linked System is a large-scale modeling system used in USDA's Economic Research Service to conduct "what if" scenario analyses of issues affecting the agricultural sector. Scenario analyses typically are conducted in 4 steps—we analyze and define the problem, implement the scenario in the model, evaluate the results for economic consistency particularly with regard to the appropriateness of the implementation approach, and iterate, if needed, to fine-tune how the scenario was implemented.

Scenarios often provide interesting challenges to analysts and modelers that require new approaches to adequately address an issue. This paper has illustrated these types of challenges by providing examples of scenario analysis approaches used for addressing effects of China's WTO accession and effects of Argentina's economic crisis. The reasoning behind the approaches taken in those scenarios may suggest solutions to other problems that analysts encounter in conducting scenario analyses in a changing world.

Appendix:
Varieties of Analyses Using the CCLS

Baseline
Baseline projections with scenario changes related to:
- European Union (EU) exchange rates;
- the soybean marketing loan and non-recourse loan;
- the Asia Crisis;
- the EU set aside rate;
- Conservation Reserve Program (CRP);
- EU wheat transportation costs;
- Acreage Reduction Program; and
- General Agreement on Tariffs and Trade (GATT).

Trade Liberalization
- China WTO accession scenario (multiple analyses).
- Implications of the General Agreement on Tariffs and Trade (GATT) on incomes, tariffs, wheat price differentials from changes in Export Enhancement Program (EEP) recipients, etc.
- Western Hemisphere trade liberalization with bilateral tariff reductions; trade flow changes modeled through applied Armington subsystem.
- Effects of removal of the Canadian rail subsidy (WGTA) for western grain.

Trade Policies
- Analysis of impacts of USDA's export programs, including the Export Enhancement Program (EEP), GSM credit guarantees, PL-480, and Section 416. The analysis helped to show the benefits of the programs, to allow comparison with the costs (2 analyses).
- Examination of U.S. coarse grain exports to Canada.
- Cotton Step 2 payments.
- Effects of imposing genetically modified organism (GMO) risk assessment costs on importers and inspection/testing costs on exporters of wheat, coarse grains, and oilseeds.

U.S. Commodity Policies
- Soybeans, no marketing loan or non-recourse loan program.
- Farm bill support, including GATT, EEP, CRP, and EU set aside scenarios.

Area, Yield, and Production Changes
- Potential new cropland in Argentina and Brazil, and impacts on area, yields, production, and world prices.
- Yield shocks in Russia and Ukraine.
- Analysis of the effects of foot-and-mouth disease on the Taiwanese pork industry.

- Baseline-related projections for Asia with historical pattern yield shocks.

Macroeconomic Changes
- Macroeconomic scenarios, including income, exchange rates, deflators, oil prices, etc.
- Income and exchange rate scenarios.
- Asia crisis, shocks to income and exchange rates (3 analyses).
- Slower income growth in developing countries.
- Russian income growth scenarios.

Country/Region Scenarios
- The following 24 countries/regions have been the focus of analyses: Argentina, Australia, Brazil, Canada, Central America and Caribbean, China, Czech Republic, European Union, Hungary, Indonesia, Japan, Korea, Malaysia, Mexico, Other South America, Paraguay, Philippines, Poland, Russia, Taiwan, Thailand, Ukraine, Venezuela, and the United States.

European Union
- European Union enlargement through accession of Czech Republic, Hungary, Poland, and Slovakia; movement to EU Common Agricultural Policy prices and policies (4 analyses).
- Effects of the EU meat and bone meal ban and BSE crisis.
- European Union (EU) export subsidy elimination.
- European Union implementation of Agenda 2000.

China
- Effects of newly-revised Chinese meat data and re-balanced feed demand, and new feed conversion coefficients.
- China scenarios including yield reductions from ozone, full trade liberalization, yield reductions from water shortages, faster declines in arable land, and moderation in income growth.
- Scenarios involving yield reductions and high Chinese demand, to simulate a high-food-price scenario alternative to the baseline projections for long-term food aid needs and availabilities.

Scenario Analyses Using A Global Dynamic Applied Equilibrium Model

Agapi Somwaru, Economic Research Service, U.S. Department of Agriculture and
Xinshen Diao, Trade and Macroeconomic Division, International Food Policy Research Institute (IFPRI)

Introduction

Applied (computable) general equilibrium (CGE) static models have been widely used as tools for trade reform and tax policy analysis for both developed and developing countries (see, for example, Shoven and Whalley, 1984; Cox and Harris, 1992; Brown, Deardorff, and Stern, 1992; Dervis, De Melo and Robinson, 1992). But traditional static CGE models cannot capture intertemporal economic behavior in a theoretically consistent manner. Past attempts to interject dynamics within static CGE models have been superficial: savings are assumed to be a fixed proportion of disposable income while investment is specified by "macro closures." The lack of theoretical foundation for *intertemporal* decisions and the failure to consistently capture the behavior of economic agents has not escaped the attention of many modelers (Bell and Srinivasan, 1984). The main features of a "basic" static CGE model is its circular flow; that is, it simulates the working of a market economy in which prices and quantities adjust to clear markets for products and factors. The model specifies the behavior of optimizing consumers and producers in the market economy (figure 1). If the model includes additional institutions such as the government, the capital account, as well as exports and imports, then we get the so-called full CGE model (figure 2).

In contrast with the static models, the key feature of the dynamic is its ability to capture economic adjustment behavior generated by a policy change in a theoretically consistent manner. Unlike the static model, consumption decisions are not made according to the household's current income. Consumption and saving decisions, which are jointly determined, are made intertemporally and savings are generated for future consumption (figure 3). The dynamic model by allowing for forward-looking behavior resembles closely the real economic behavior of agents and generates better results in policy changes.

On the supply side, producers make production decisions based not only on current but also on future prices. Investment decisions account for both current and future returns simultaneously and intertemporally. The ability to account for investment decisions is particularly important. In the static CGE model, resources (including capital) are fixed. Gains from change in policy are generated only from more efficient reallocation of current resources. As a result, the effects of the policy change may be underestimated as the investment response is not taken into account. To overcome this shortcoming, some static CGE modelers adjust the capital stock exogenously. Arbitrarily adjusting the capital stock not only departs from the economic theory but embodies the modelers' subjectivity. Our dynamic model is global in its specification. All countries and regions in the world economy are characterized by their intertemporal economic behavior. The model is also flexible in terms of sector coverage. Two features are critical in the model. First, the model's focus is in the real economy and it does not contain monetary features. The model's core on real economy makes it consistent with the neoclassical macroeconomic and growth theories. In other words, international financial markets, hence borrowing and lending, are captured by the real movements of commodities among countries. This makes the movement in the current account consistent with shifts in the trade account. Second, total factor productivity (TFP) and population growth are treated as exogenous in the neoclassical growth theory. For more on the model specification see Diao and Somwaru (2001).

Empirical Application of Scenario Analysis

The global model, both static and dynamic, was used to evaluate various scenarios of agricultural trade liberalization. In particularly, we evaluate the possible impacts of removing agricultural trade barriers, subsidies, and other trade distorting forms of support on the world economy. The model utilizes trade and domestic support instruments to capture worldwide adjustments to policy changes. It simulates changes in policy as scenario analyses in counterfactual or "what-if" comparison. Depending on the assumptions of factor mobility the model can assess medium to long-run changes.

Scenario Analysis --Static Approach
Effects on Agricultural World Prices

Using scenario analysis, we evaluate the effects of elimination of all tariffs (and tariff equivalents) on agricultural imports, export subsidies, and domestic support worldwide on world prices by country/regions (appendix I-1, and I-2; for more

details see Diao et al., 2001). The results indicate an increase in world agricultural price level of 11.6 percent relative to the level of world nonagricultural prices. Eliminating tariffs worldwide accounts for more than 50 percent of the 11.6 percent increase in world agricultural prices. That is, when other policy variables remain constant and only agricultural import tariffs are eliminated, world agricultural prices rise by 6 percent (relative to world nonagricultural prices) (figure 4). This result occurs because import barriers protect domestic producers by restricting imports. In many *import-protecting* countries, import restrictions raise domestic food prices higher than world prices while at the same time inducing these countries to employ too many resources in agriculture. Eliminating import tariffs raises the demand for agricultural imported goods, while contracting supply, thus placing upward pressure on world agricultural prices. This pressure in turn induces agricultural exporting countries to increase production. Tariffs are more distorting than other policies of agricultural support. The elimination of trade barriers by the developed countries, either as a group or individually, contributes the most to the increase in prices.

Effects by Policy Instrument

The study uses the following scenarios for evaluating the impact on trade by policy instrument: (1) eliminating agricultural import barriers (tariff equivalents) throughout the world; (2) eliminating agricultural export subsidies throughout the world; (3) eliminating domestic support in the developed countries; and (4) combining each of these scenarios. The study uses as two indicators to assess the effects of agricultural liberalization on the world economy, as well as on each country/region: (a) changes in world agricultural trade and (b) changes in a measure of social well-being, or welfare.

In general, freer trade results in more trade. The model results indicate that world agricultural trade is likely to increase substantially after liberalization. Removing all agricultural support and protection worldwide results in an increase in the value of world agricultural trade by about 30 percent. The corresponding volume of world trade rises 15 percent (table 1). Agricultural exports from *developed* countries rises by 32 percent, while exports from *developing* countries increase 27 percent in value. However, the corresponding increase in the volume of exports from the developing countries (16 percent) is larger than the increase from the developed countries (14 percent).

The scenario analysis indicates that the removal of import protection is a dominant stimulus to growth in world agricultural trade. When only agricultural tariffs worldwide are eliminated, world trade rises 26 percent in value and 17 percent in volume. Exports and imports both rise more in the developed country group than in the developing country group. Removing export subsidies or domestic support alone appears not to enhance world agricultural trade. When only agricultural export subsidies worldwide are eliminated, world agricultural trade falls 0.7 percent in value and 1.8 percent in volume. If only domestic support in the developed countries is eliminated, world agricultural trade rises 2.8 percent in value but falls slightly (0.7 percent) in volume (table 1). These results are consistent with the prediction of trade theory, in that, subsidies increase exports, albeit at the possible cost of reducing the exports of non-subsidized commodities.

Effects on Welfare

From a world perspective, more efficient allocation of resources yields higher global welfare. Typically, in a country with a high degree of agricultural support and trade protection, consumers pay relatively high prices for food and other agricultural goods, and/or their disposable income is taxed to cover the costs of agricultural policies. Our scenario analysis indicates that removing support or trade protection is expected to benefit consumers, however, welfare effect across countries vary and particularly when the world price is affected by agricultural policies.

This analysis uses the widely accepted equivalent variation (i.e., consumers' willingness to pay) concept to measure the social welfare gains or losses due to agricultural policy reform. Measurements consider both one-time welfare effects and welfare effects over time. The one-time effect measurements use status-quo (pre-reform) prices as the base and address the question: what income would be equivalent to the change brought about by agricultural policy liberalization (Varian, 1984). The welfare effects over time are measured by summing the discounted value of this measure over time.

The one-time effects of agricultural policy liberalization on a nation's social welfare appear relatively small among all countries/regions (table 2). Relative to nonagricultural sectors, agriculture accounts for a small share of Gross Domestic Product (GDP). Further, agricultural goods in consumers' consumption bundle in most countries, and particularly in the developed economies of the EU,

Canada, and the United States, are small as a proportion of total expenditures. Agriculture (including processed food products) accounts for less than 5 percent and 15 percent of the GDP of developed and developing countries, respectively. Consumption expenditures on food account for 5 percent of total expenditures for the developed country group and 17 percent for the developing country group. Thus, at a national level, agricultural policy liberalization alone is unlikely to have a large, one-time welfare effect on the aggregate economy in the short to medium-term.

The welfare effects are positive for the world aggregate. The sum of countries' equivalent variation due to worldwide agricultural policy reform is about $31 billion. This is equivalent to 0.1 percent of world aggregate GDP, and 1 percent of consumers' expenditures on agricultural and agriculture-related goods (table 1-7). Such welfare gains, however, are not equally distributed and some countries experience a negative welfare effect. Developed countries experience a $28 billion welfare gain, which is equivalent to 0.16 percent and 2 percent of their GDP and consumer expenditures on agricultural goods, respectively. Moreover, all developed countries in the model gain, with the largest gains shown by the EU ($9.3 billion), Japan and Korea ($8.6 billion), and the United States ($6.6 billion).

Among the three policy categories, removing tariffs generates positive welfare gains and for most countries and regions, while removing domestic support and export subsidies has negative effects for most developing countries (table 1-7). Holding other policy variables constant, removing tariffs results in a $25 billion welfare gain worldwide, $19.6 billion of which accrues to the developed countries and $5.7 billion to the developing countries. Removing domestic support or export subsidies results in a much smaller welfare gain worldwide, as export subsidy rates are much lower than the tariff rates in all countries/regions and the domestic support policies are mainly employed by the developed countries. The world aggregate welfare gain from the removal of domestic support is $2.8 billion, and the gain is $250 million from removal of export subsidies. Developed countries gain $4.7 billion from domestic support removal and $2.5 billion from export subsidy removal. Developing countries, however, experience welfare losses of $1.9 billion and $2.3 billion in the two scenarios, respectively.

Dynamic Welfare Results A Brief Overview of Method and Assumptions

The analysis earlier ignored the effect of reform on savings, investment, and the pattern of growth in a country's capital stock. To analyze these effects requires assumptions regarding households' willingness to forgo consumption and investment, the functioning of capital markets and international capital flows, as well as the technological spillovers that seem to accompany growth in countries' trade. These assumptions may be closely approximated for developed countries, but only poorly approximated for many developing countries. Nevertheless, for the most part, the analysis suggests changes in the long run that seem well within the realm of reason.

Numerous studies find empirically strong and positive linkages between growth and a country's total factor productivity (TFP) and the share of its economy involved in trade with more advanced nations (e.g., Coe and Helpman, 1995: Wang and Xu, 1997; and Coe et. al., 1997). Thus, a dynamic model should capture not only consumer saving and producer investment decisions but also the effects of trade liberalization on a country's growth in factor productivity. Such effects are modeled by increases in technological spillovers embodied in the trade between developing and developed countries. Specifically, if a developing country eliminates trade protection, it then tends to increase its rate of learning new skills, organizational methods, and the more advanced product and process technologies embodied in its imports of investment goods from developed countries. This process helps to increase labor productivity and returns to land and social capital (Grossman and Helpman, 1991; Romer, 1994). The spillovers of the advanced technology embodied in trade can also result from developed countries' reduction of agricultural protection. As developed countries increase imports of agricultural goods, their exports of capital goods may be enhanced. Thus, this longer-run type of analysis allows for agricultural trade reform to yield broader economy-wide benefits, which, as shown next, is found to be higher for developing countries.

This study calculates the change in the regional equivalent variation for three different years as well as the intertemporal welfare index, which measures the welfare gains in this dynamic setting. Changes in equivalent variation for the three different years are compared with the base year, while the intertemporal welfare index is the sum of the welfare change over time where future gains and losses are discounted relative to current gains and losses. The over-time welfare effects of the liberalization vary, depending

on whether technological spillover-growth considerations are included in the analysis. Thus, welfare changes are specified under the different assumptions and, hence, the technological spillovers and growth effect of the policy reform on welfare can be told from the differences in the two groups of results.

Large Dynamic Welfare Gains

Without taking into account the technological spillover-growth effects of liberalization, (that is, by considering only the investment incentives created by reform) the over-time welfare effect is still modest in the first five years (table 3). As production and investment adjustments take time, the welfare effect in a longer time period, for example, in the 15th year or after, is relatively large. The world welfare gain in year-10 doubles the gain accrued in year-5. More simply stated, this result suggests that the payoff to agricultural trade policy reform takes time.

However, if the technological spillover-growth effect of policy reform is taken into account for developing countries, the over-time welfare gains increase significantly, especially in developing countries. The developing countries are beneficiaries of the technological spillovers embodied in trade with developed countries. Such benefits are assumed to generate an additional annual growth rate of 0.02 percent in the developing countries. This annual growth rate further increases welfare gains among the developing countries. Moreover, all the developing countries/regions in the model are better off after agricultural support and trade protection are totally removed worldwide, and the greater the volume of trade between developed and developing countries, the larger the welfare gain.

Conclusions

Our CGE framework, both static and dynamic, uses scenario analysis to analyze change in policies. Scenario analysis is a useful tool to capture economic agents' decisions under alternative policies and time spans. The dynamic specification captures the payoff to reform over time. Although the results are counterfactual they can used to evaluate the direction and trends of changes in policy.

Of the three categories,-- tariffs, domestic support, and export subsidies-- the results suggest that tariffs are the major cause of distortions in world agricultural prices. The worldwide elimination of import tariffs would cause world agricultural prices

to increase about 6 percent. As the protection levels and trade patterns vary among countries, so do the effects on prices, trade and welfare.

The study also finds that the payoff to reform takes time. Over time, worldwide agricultural liberalization generates larger gains than the short-time gains for most countries. For example, the discounted present value of world welfare gains in year-10 doubles the gain accrued in year-5. Moreover, if the technological spillover-growth effect of reform is taken into account, the welfare gains increase significantly for all countries in the world.

References

Coe, D.T., and E. Helpman (1995). "International R&D Spillovers," *European Economic Review*. 39: 859-887.

Coe, D.T., E. Helpman, and A.W. Hoffmaister (1997). "North-South R&D Spillovers," *The Economic Journal*. 107: 134-149.

Diao, X., A. Somwaru and T. Roe "A Global Analysis of Agricultural Reform in WTO Member Countries," *Agriculture Economic Report 802,* Economic Research Service, USDA, 2001.

Diao, X., and A. Somwaru. "A Dynamic Evaluation of the Effects of a Free Trade Area of the Americas - An Intertemporal, Global General Equilibrium Model." *J. Econ. Integration* 16(first quarter, 2001):21-47

Grossman, G., and E. Helpman (1991). *Innovation and Growth in the Global Economy*: MIT Press.

McDougall, R., A. Elbehri, and T. Truong (1998). *Global Trade, Assistance, and Protection: The GTAP 4 Data Base*. Center for Global Trade Analysis, Purdue University. West Lafayette, IN.

Romer, P.M. (1994). "The Origins of Endogenous Growth," *Journal of Economic Perspectives*. 8(1): 3-22.

Varian, Hal R. (1984). *Microeconomic Analysis*, Second Edition. W.W. Norton & Company.

Wang, J., and B. Xu. (1997) "Trade, FDI, and R&D Spillovers in the OECD," University of Florida. Unpublished.

Figure 1. Main Attributes of a Basic CGE Model

- Goods & Services
- Households
- Firms
- Factors

Demand · Expenditure · Final Goods: Supply · Final Goods: Revenue · Intermediate Goods: Supply · Intermediate Goods: Revenue · Income · Supply · Rents · Demand
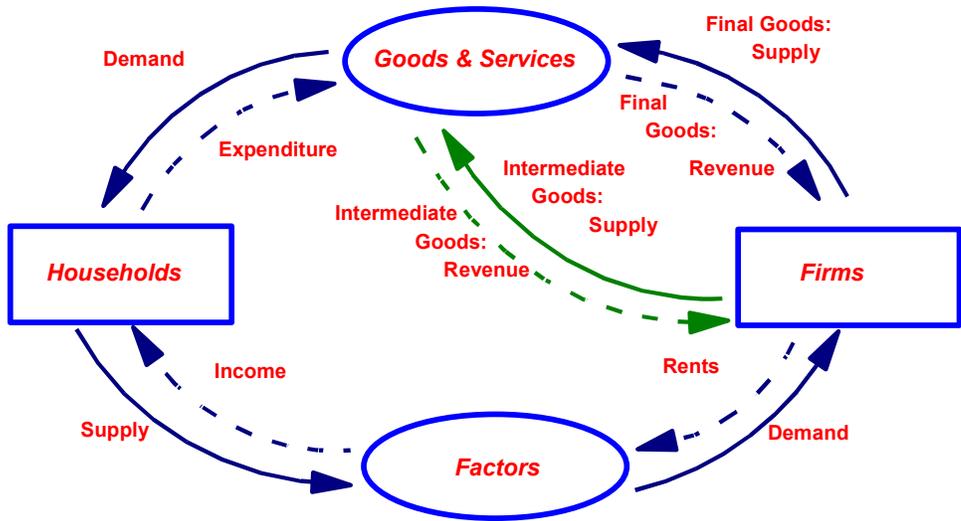
Figure 2. Main Attributes of a Full CGE Model

- Government
- Other countries / Rest of the World
- Basic CGE Model Structure
- Capital Account

Transfers & Business Subsidies · Taxes · Govt Purchases · Imports · Exports · Govt. Savings (Budget Deficit) · Private Savings · Investment · Savings from ROW (Net Capital Inflow)
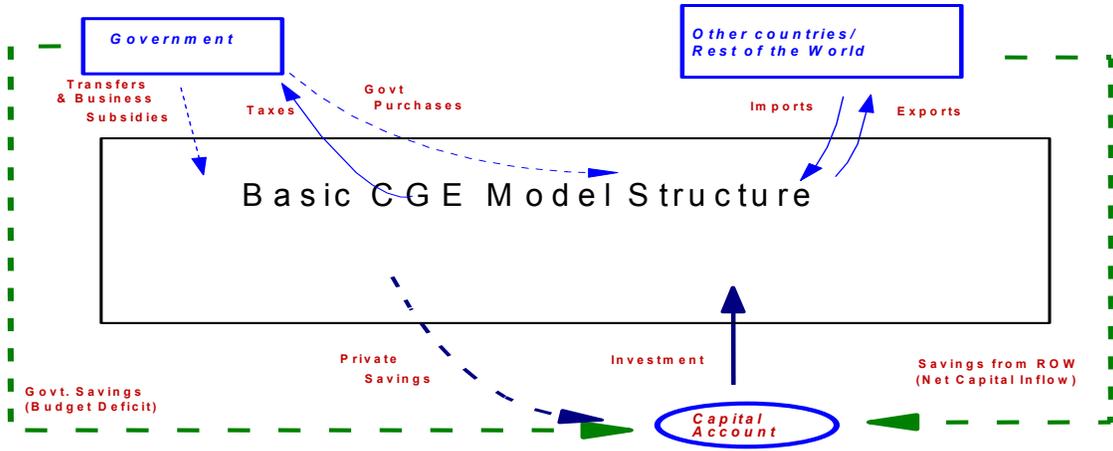
Figure 3. Attributes of a Dynamic Model

CURRENT INCOME

Consumers' Intertemporal Decisions

CONSUMPTION   SAVINGS

The elliptical box represents exogenous variables, the rectangular box represents endogenous variables. The dashed line is for revenue flows and solid line is for commodity flows.

World Capital Market

FOREIGN ASSETS/DEBT

World Commodity Market   Domestic Commodity Market

IMPORTS & EXPORTS   OUTPUT   DIVIDENDS   INTEREST

Firms' Intertemporal Decisions

ENDOWMENT INCOME   PRODUCTION   INVESTMENT   CURRENT PROFIT

LABOR, LAND   CAPITAL

Figure 4--Static analysis: Effects of trade reform on global prices
Percentage change in world agricultural price index from the base year

export subsidy 1.6

domestic support 3.7

tariffs 6.3

LDCs all 2.3

DCs all 9.1

Japan & Korea all 1.5

US all 1.8

EU all 4.4

LDCs tariffs 2.3

DCs tariffs 4.0

Scenarios
(source: estimated by ERS)

Table 1. Decomposition of World Agricultural Trade Effects of Global Agricultural Liberalization in the Model

-- Percentage change in total agricultural trade from the base year (1997)

Removing agricultural supports and protections by all regions		
World trade	29.71	14.66
Exports of developed country group	31.81	13.75
Imports of developed country group	35.93	19.03
Exports of developing country group	26.50	16.05
Imports of developing country group	20.02	7.85
Removing tariffs by all regions		
World trade	26.40	17.31
Exports of developed country group	31.28	20.79
Imports of developed country group	28.66	18.39
Exports of developing country group	18.93	11.97
Imports of developing country group	22.89	15.63
Removing domestic supports by developed regions		
World trade	2.70	-0.71
Exports of developed country group	0.85	-3.42
Imports of developed country group	5.43	1.82
Exports of developing country group	5.54	3.44
Imports of developing country group	-1.54	-4.70
Removing export subsidies by all regions		
World trade	-0.66	-1.76
Exports of developed country group	-1.43	-3.04
Imports of developed country group	-0.44	-1.25
Exports of developing country group	0.51	0.22
Imports of developing country group	-1.01	-2.54

Source: Estimated by ERS.

Table 2. Static Welfare Effects of Global Agricultural Liberalization

	EXP-1	EXP-2	EXP-3	EXP-4
		$ billion		
World	**31.06**	**25.22**	**2.80**	**0.25**
Developed country group	**28.48**	**19.56**	**4.74**	**2.53**
Australia and New Zealand	1.57	1.17	0.24	0.01
Japan and Korea	8.59	13.81	-3.66	-1.34
united States	6.57	3.83	0.97	-0.09
Canada	0.75	0.40	0.28	-0.09
European Union	9.28	0.14	6.06	3.72
EFTA	1.73	0.20	0.83	0.32
Developing country group	**2.60**	**5.66**	**-1.94**	**-2.28**
China	0.42	0.85	-0.28	-0.21
Other Asian countries	1.52	1.71	-0.09	-0.25
Mexico	-0.16	0.19	-0.27	-0.11
Latin America	3.65	2.71	0.68	-0.05
South African countries	0.25	0.60	-0.22	-0.22
Rest of the world	-3.07	-0.39	-1.76	-1.43

Source: Estimated by ERS.

Experiment-1 (EXP-1): Removing all agricultural supports and protections worldwide
Experiment-2 (EXP-2): Removing only tariffs worldwide
Experiment 3 (EXP-3): Removing only domestic supports in the developed countries
Experiment 4 (EXP-4): Removing only export subsidies worldwide

Table 3. Dynamic Welfare Effects of Global Agricultural Liberalization

	Without TFP growth			With TFP growth		
	In Year 5	In Year 10	In Year 15	In Year 5	In Year 10	In Year 15
		$billion			$billion	
World	**15.94**	**30.19**	**36.26**	**27.17**	**46.98**	**56.39**
Developed country group	**14.69**	**25.66**	**29.74**	**17.00**	**29.59**	**35.14**
Australia and New Zealand	3.26	3.34	3.40	3.32	3.43	3.52
Japan and Korea	-1.40	3.86	5.10	-0.85	4.70	6.17
United States	8.72	10.60	11.76	9.18	11.59	13.30
Canada	1.05	1.17	1.24	1.13	1.27	1.37
European Union	3.35	6.68	8.15	4.41	8.48	10.58
EFTA	-0.27	0.02	0.09	-0.18	0.12	0.21
Developing country group	**1.25**	**4.52**	**6.52**	**10.16**	**17.39**	**21.25**
China	1.24	1.68	1.83	1.48	2.02	2.23
Other Asian countries	-0.70	0.54	0.93	2.10	4.47	5.11
Mexico	-0.40	-0.22	0.09	0.53	0.99	1.60
Latin America	3.94	4.27	4.66	4.62	5.36	6.11
South African countries	0.16	0.33	0.50	0.35	0.59	0.81
Rest of the world	-3.00	-2.07	-1.49	1.07	3.97	5.39

Source: Estimated by ERS.

Appendix

I-1. Agricultural sectoral aggregation in the study

Sectors in the model	Sectors in GTAP data
Rice	Paddy rice, processed rice
Wheat	Wheat
Corn and other cereal grains	Corn and other cereal grains
Vegetable and fruits	Vegetable, fruits and nuts
Oil seeds and products	Oil seeds, vegetable oil
Sugar	Sugar cane and sugar beet, sugar
Other crops and products	Plant-based fibers, other crops
Livestock and products	Bovine cattle, sheep and goats and meats, other animal products, raw milk and dairy products, wool, and silk-worm cocoons
Other processed food sector	Beverages and tobacco products, and other processed food products

I-2. Countries and regions included in the study

1) Australia and New Zealand; 2) China, including Hong Kong; 3) Japan and Korea; 4) The other Asian countries;
5) Canada; 6) The United States; 7) Mexico; 8) Latin American countries; 9) the European Union;
10) the European Free Trade Area; 11) South African countries; 12) the rest of the world

Macroeconomic Issues

Chair: Jeffrey Osmint,   U.S. Geological Survey, U.S. Department of the Interior

Bankers or Macroeconomic Forecasters: Whose Interest Rate Forecast is Better?

David Torgerson, Economic Research Service, U.S. Department of Agriculture

This study examines the short-term interest rate forecasting record of bankers versus professional macroeconomic forecasters. In particular, over the period 1979 to 1990 we examine the one period ahead 6-month Treasury bill forecasting record of the bankers surveyed by the Kansas City Federal Reserve versus the Livingston Survey (the oldest panel of macroeconomic forecasters). The results generally favor the bankers.

Contingent Forecasting of Bulges In the Left And Right Tails of the Nonmetro Wage and Salary Income Distribution

John Angle, Economic Research Service, U.S. Department of Agriculture

A trend in the mean of nonmetro wage and salary income implies trends in the bulging of the left and right tail of the nonmetro distribution of wage and salary income. With four decades of data from the March Current Population Survey it is shown that when the estimated mean increases, the bulge in the left tail–the proportion of small incomes–diminishes and a bulge in the right tail–proportion of large incomes–appears. And vice versa for decreases in the mean. Proportionally, the bulges are mirror images of each other. A trend in one tail not only implies a trend in the other but in the mean as well.

Forecasting the Business Cycle With Polar Coordinates

Foster Morrison and Nancy L. Morrison, Turtle Hollow Associates, Inc.

The established methods of mathematical forecasting are special cases of noise-driven, linear difference equations. Nonlinear terms can be added directly at the cost of a more difficult stability analysis. An alternative is to apply a nonlinear transformation and then use a linear forecasting model on the new variables. In the case of the business cycle, polar coordinates provide an obvious choice. This provides a means to use statistical properties of the noise other than the mean square ("signal power"). This approach offers improved possibilities of forecasting the beginnings and endings of recessions.

This page intentionally left blank.

CONTINGENT FORECASTING OF BULGES IN THE LEFT AND RIGHT TAILS OF THE NONMETRO WAGE AND SALARY INCOME DISTRIBUTION

John Angle
Economic Research Service
1800 M Street, NW
Washington, DC 20036

Response to the Theme of the Conference[1]

The theme of this conference is the uncertain task of forecasting a variable's value after there has been a discontinuity in the time series caused by a major shift in the factors that generate and affect the variable. The present paper responds to the conference theme with a model that credibly forecasts how the nonmetro[2] distribution of wage and salary income would change should an adverse event cause mean nonmetro wage and salary income to fall much farther than it has in recent decades, e.g.., 40%. A fall of this magnitude is approximately eight times that of the greatest year to year fall in estimated mean nonmetro wage and salary income in the period 1963 through 1995 of about 5%.

Contingent forecasting of an event that has not occurred in the database must - to be credible - be forecast from a credible model of the dynamics of the variable in question. The credibility of this paper's model is demonstrated by its ability to parsimoniously reproduce three visually arresting but quite puzzling patterns in the dynamics of this distribution in the thirty-three year period, 1963 through 1995. These patterns are puzzling because while suggesting major systemic constraints on income distribution of some kind, they are not even commented on in the economic literature. Sahota (1978) asserts that there is no deductive route from micro-economics to a functional form for the distribution of wage and salary income.

There is a substantial public policy interest in what would happen to the nonmetro distribution of wage and salary income should an adverse event affecting the whole nation cause the mean of nonmetro wage and salary income to fall. Nonmetro residents have, historically, lower levels of educational attainment, lower mean wage and salary incomes, a greater relative frequency in the left tail of the wage and salary income distribution, and a greater probability of being categorized as in poverty than metro residents. See Fuguitt, Brown, and Beale, 1989. The public policy interest in the shape of the nonmetro wage and salary income distribution is on-going. For example, Karl Stauber (2001) deplored what he perceived to be the loss of middle income people in rural counties of the northern Great Plains states, an area of negative population growth. The present paper forecasts what would happen to the distribution of nonmetro wage and salary incomes if the nonmetro mean fell 40%. In particular, the present paper shows that a large, perhaps surprisingly large, proportion of less well educated nonmetro people will be reduced to a low level of wage and salary income when the overall mean of nonmetro wage and salary income drops precipitously.

Model-Based Forecasting

If a reliable model of the dynamics of the variable to be forecast is available, a general purpose forecasting tool, such as ARIMA, is usually discarded in favor of forecasting from the model. For example, consider forecasting the orbit of a comet. Optimal forecasting would rely on Newtonian mechanics, not ARIMA applied to a time-series of observations on coordinates. Forecasting the orbit of the comet as it approaches Jupiter and its orbit changes, a major shift of circumstance, requires Newtonian mechanics. An understanding of why change occurs, as represented in a model, is essential for forecasting past a major shift affecting the phenomenon to be forecast.

After describing the database, this paper establishes the credibility of the model to forecast how the nonmetro wage and salary income distribution changes if its mean falls 40%. The model's credibility is established by showing that the model parsimoniously reproduces the three puzzling

---

[2] 'Nonmetro' refers to the set of nonmetropolitan counties. A nonmetropolitan county is a county not in a Metropolitan Statistical Area (MSA) as defined by the Office of Management and Budget. MSA's include core counties containing a city of 50,000 or more people or having an urbanized area of 50,000 or more and total area population of at least 100,000. Additional contiguous counties are included in the MSA if they are economically integrated with the core county or counties. The metropolitan status of every county in the U.S. is re-evaluated following the Decennial Census. While there has been a net decline in counties classified as nonmetro over the decades, the definition of nonmetro has remained roughly constant

patterns in the distribution's dynamics from 1963 through 1995.

Data

The distribution of nonmetro wage and salary income is estimated with data from the March Current Population Surveys (CPS) of the years 1964 through 1996, which provide data on annual wage and salary income from the years 1963 through 1995. The March CPS is conducted by the U.S. Bureau of the Census. It has a substantial number of households in its nationwide sample. Studies of wage and salary income distribution vary in terms of the breadth of the population of earners examined. Economists often prefer a privileged subset of this population: full-time, year-round workers. Some economists limit their focus further to full-time year-round workers who are male heads of household, in a narrow age range, perhaps 30 to 45 years of age. The present study includes anyone earning at least $1 in wage and salary income, between the ages of 25 to 65 residing in a nonmetro county.

The measurement of education changed in the CPS after the 1990 Census from a count of years of school completed to a more degree oriented measure which better measures the diversity of post-secondary education. The present study reconciles the two categorizations of educational attainment by collapsing both sets of categories to an ordinal polytomy of five categories. The crudeness of this categorization obliterates the distinction between the two different categorizations of educational attainment. The categories of highest level of education attained used here are:

elementary school or less
some high school
completed four years of high school
some college
completed four or more years of post-secondary education

Pattern #1: Mirror Image Proportional Change in Two Bins, One in Left Tail, the Other in Right Tail

Figures #1, 2, and 3 are the time-series of three different statistics, but they have a common pattern which can be demonstrated via linear transformation or their intercorrelations. Figure 1 is the time-series of the median of nonmetro wage and salary income from 1963 through 1995. Figure 2 is the time-series of the relative frequency of

nonmetro wage and salary incomes in the range (or bin as statisticians say) of $1 to $8,000 in terms of 1995 dollars[3]. Figure 3 is the time-series of the relative frequency of the bin $36,001 to $44,000. This bin is to the right of the mode, median, and mean of the distribution and so is in the distribution's right tail.

---

[3] This paper estimates the distribution of annual nonmetro wage and salary income the traditional way, in terms of relative frequencies of observations falling into bins of fixed width. There are many ways to estimate a distribution. All of them involve a trade-off between parsimony of model and error of fit. Parsimony is expressed in the amount of smoothing of the estimate. In terms of fixed bins, the greater bin width, the fewer bins are used, and the greater the degree of aggregation and the smoother the estimate of the distribution. A wage and salary income distribution of a large population defined in geographic terms, is a distribution that is quite a familiar object and has been discussed and dissected for many years. It is known to be right skewed (Pareto's Law, broadly construed) and unimodal. Angle (1994) demonstrates the existence of a micro-structure of frequencies spikes over round income amounts in March CPS income data, indicative of respondents giving incomes to Census Bureau interviewers with fewer significant digits than the interviewers ask for. Census Bureau questionnaires ask for incomes to the nearest $1. Angle (1994) shows that this rounding of income amounts does not apparently introduce a net upward or downward bias. In published tabulations, the Census Bureau, traditionally, presents income distributions near their mode in terms of relative frequencies in bins of fixed length, e.g., $5,000, and in the right tail, in terms of bins of increasing width. This policy is intended to keep the standard errors of estimate of the right tail bins comparable to those of the bins near the mode. However, such presentation disguises how right skewed income distributions are because it is difficult to mentally adjust the relative frequency down for increasing bin length in the right tail. The present paper estimates a distribution with fixed length relative frequency bins, either $4,000 (in terms of constant 1995 dollars) wide or $8,000 wide. The choice is made to facilitate comparison between the more dense left tail and the less dense right tail. This paper makes many such comparisons.

All dollar amounts are converted to constant 1995 dollars using the PCE (personal consumption expenditure) price index numbers form Table B-7 Chain-type price indexes for gross domestic product, **Economic Report to the President**, February 2002 (Council of Economic Advisers, 2002).
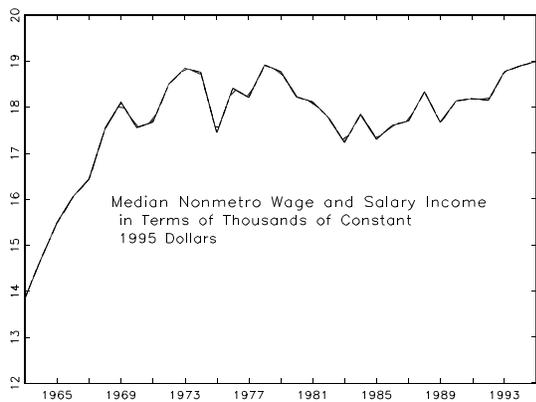
Figure 1
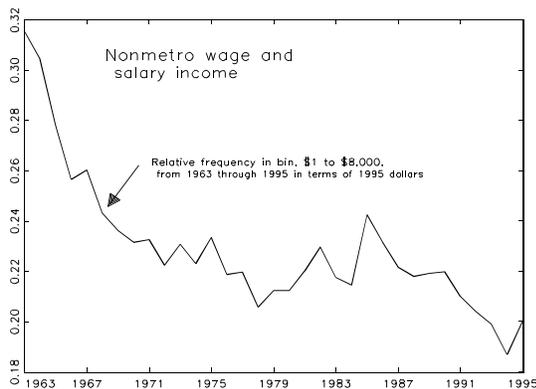
Figure 2
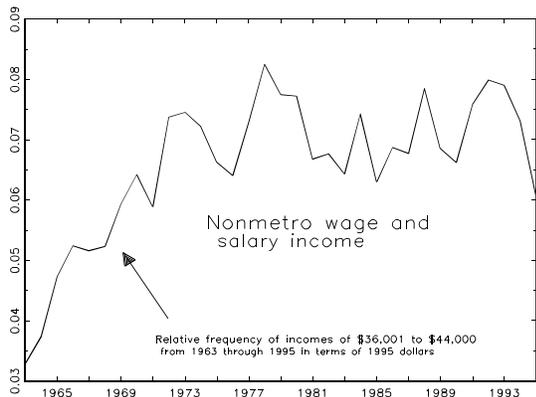
Figure 3

Notice that the relative frequency in the bin in the right tail (figure 3) rises as the relative frequency in the left tail bin (figure 2) falls, and vice versa in a minor way for a year or two following 1980. The two statistics are closely correlated, albeit inversely, over time: -.888. Comparison of figures 1 and 3 shows that these time-series rise together in a similar way with a simultaneous slight downturn at the end of the 1970's and beginning of the 1980's.

The median and the relative frequency in the right tail bin are correlated .872. The median is also closely, but inversely, correlated with the relative frequency in the left tail bin, but inversely: -.932. These intercorrelations are so high that they are indicative of near statistical equivalence.

Figure 4 shows that these three time-series, after transformation, largely overlap. The transformed relative frequencies are particularly close. Note that the time-series of the relative frequency in the bin in the right tail of the distribution, that of bin $36,001 to $44,000, appears to be centered on the time-series of the relative frequencies from the left tail bin, $1 to $8,000.

The similarities of figures #1, 2, and 3 might be, conceivably, accounted for by a rigid shifting of the nonmetro distribution of wage and salary income to the right as its mean increases. As figure 5 shows,
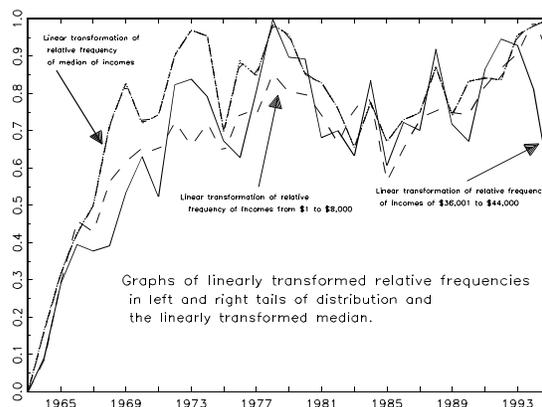
Figure 4
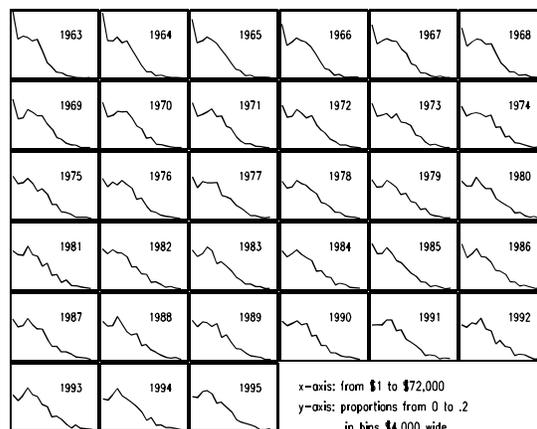
Figure 5

the distribution of nonmetro wage and salary income does not shift rigidly to the right or left over time.

The transformation of the time-series of figures #1 (median) and 3 (relative frequency in right tail) in figure 4, is:

$$\frac{y_t - \min(y)}{\max(y) - \min(y)}$$

The time-series of figure 2 (relative frequency in left tail) is transformed as:

$$\frac{\max(y) - y_t}{\max(y) - \min(y)}$$

All three transformations put the time-series on the same scale, that is, as the ratio of the positive difference between the statistic in a given year and an all-time extremum to the positive difference between the all-time maximum and all-time minimum. The transformation of the left tail relative frequency is the equivalent of flipping figure 2 as well as re-scaling it.
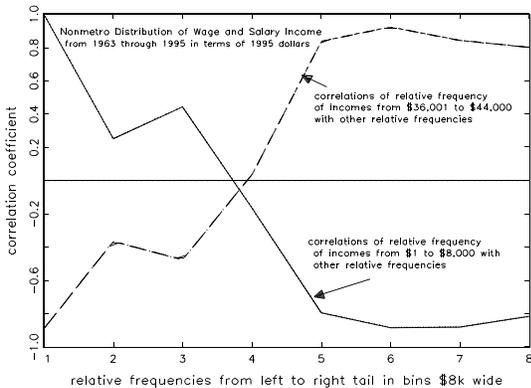
Pattern #2: Large Negative Correlation Between Relative Frequencies in Designated Left and Right Tail Bins But Near Zero Correlation Between Them and Relative Frequencies in the Central Mass of the Distribution

Figures #2 and 3 show that the time-series of relative frequencies in the bin $1 to 8,000 is highly but negatively correlated with the time-series of relative frequencies in the bin $36,001 to $44,000. Figure 6 displays the correlations of these two time-series with the time-series of relative frequencies in the other bins from $1 to $72,000. In absolute terms, the time-series of relative frequencies in the bins $1 to $8,000 and $36,001 to $44,000 are more closely correlated with each other

than with the time-series of relative frequencies of other bins. Relative frequencies in the tails are not correlated with relative frequencies in the central mass of the distribution. Figure 6 uses the bins $32,001 to $40,000 and $40,001 to $48,000 so the relative frequency in a bin defined on $36,001 to $44,000 is not correlated 1.0 with any one bin in figure 6.
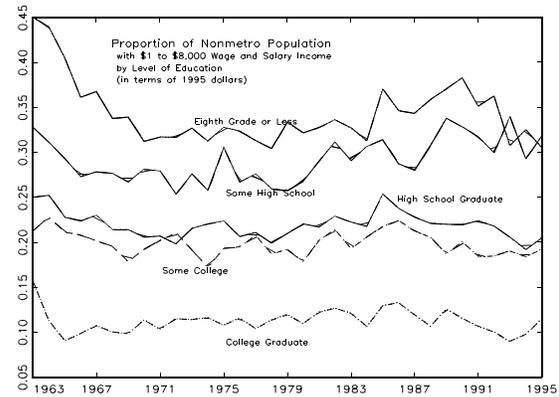
Figure 7

Pattern #3: Variability of Relative Frequencies is Proportional to Their Size

Figure 7 is the time-series of the relative frequency in the left tail ($1 to $8,000 bin) in the partial distributions of nonmetro wage and salary income conditioned on education. Figure 7 is the same as figure 2, except that in figure 7 there are five time series, one for each level of education distinguished. It is not surprising that the relative frequencies of the $1-$8,000 bin scale inversely with education. The standard deviations are:

highest level of education	mean relative frequency of incomes, $1 to $8,000	standard deviation of relative frequency of incomes, $1 to $8,000
elementary school or less	.3438	.0358
some high school	.2921	.0243
high school graduate	.2196	.0146
some college	.1988	.0136
college graduate or more	.1127	.0132

Figure 8

highest level of education	bin ranges	mean relative frequency of incomes in bin	standard deviation of relative frequency of incomes in bin
elementary school or less	$28,001-$36,000	.0640	.0141
some high school	$32,001-$40,000	.0612	.0118
high school graduate	$36,001-$44,000	.0634	.0115
some college	$40,001-$48,000	.0610	.0091
college graduate or more	$44,001-$52,000	.0691	.0109
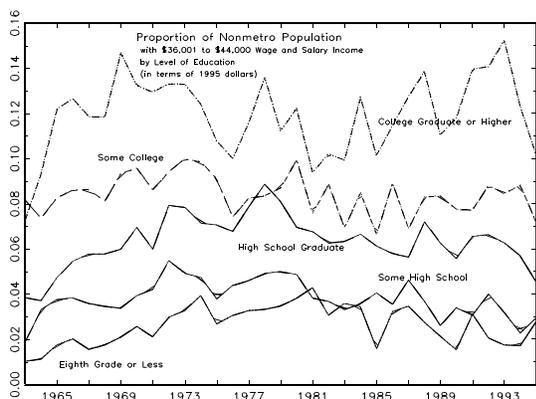
Figure 8 shows the comparable five time-series of the bin ($36,001-$44,000) relative frequencies. Not surprisingly, the right tail relative frequencies scale positively with level of education. The variability of each time-series appears to roughly scale with the size of the relative frequency. In fact, it is possible by judiciously choosing bins to roughly match both relative frequency and variability. See figure 9 and the following table. Note that the pattern in figure 9 requires for successively higher levels of education bins successively farther to the right. So variability appears to be proportional to relative frequency.

Figure 9

A Model of the Distribution of Nonmetro Wage and Salary Income

There is a model that reproduces patterns #1, 2, and 3 in the dynamics of the distribution of nonmetro wage and salary income as a function of change in its mean. The model requires each partial distribution of the conditional distribution, nonmetro wage and salary be modeled separately, then summed to an estimate of the unconditional nonmetro distribution. There are three propositions in this model:

1) Each partial distribution is gamma distributed as:

$$ f_{it}(x) \quad = \quad \frac{\lambda_t^{\alpha_i}}{\Gamma(\alpha_i)} \; x^{\alpha_i - 1} \; e^{-\lambda_t x} $$

2) Each gamma pdf modelling a partial distribution has an unchanging shape:

$$ \alpha_i \quad \equiv \quad \text{shape parameter of people} $$
$$ \text{at ith level of education} $$
$$ > \quad 0 $$

3) All the gamma pdfs modelling a partial distribution have the same scale parameter:

$$ \lambda_t \quad \equiv \quad \frac{\bar{\alpha}_t}{\bar{x}_t} $$
$$ > \quad 0 $$

where,

$f_{it}(x)$ ≡ pdf model of the

    distribution at the ith

    level of education at time t

x ≡ income > 0

λ_t ≡ scale parameter at time t

$\bar{x}_t$ ≡ unconditional mean of income at time t

$\bar{\alpha}_t$ = $w_{1t}\alpha_1 + ... + w_{it}\alpha_i + ... + w_{It}\alpha_I$

w_{it} = proportion of population

    at ith level of education at time t

and $\bar{x}_t$ and the α_i's and w_{it}'s are exogenous. The α_i's do not vary over time. So the only time-varying inputs to the model are $\bar{x}_t$ and the w_{it}'s. The w_{it}'s vary more slowly proportionally than $\bar{x}_t$. While making allowance for variation in w_{it}, the model takes $\Delta\bar{x}_t$ where:

$$\Delta\bar{x}_t \equiv \bar{x}_t - \bar{x}_{(t-1)}$$

as the stochastic shock and source of change from year to year.

    To see how this model of a partial distribution of the conditional distribution, nonmetro wage and salary income conditioned on education, changes as a function of a change in the unconditional mean of nonmetro wage and salary income requires taking the partial derivative:

$$\frac{\partial f_{it}(x)}{\partial \bar{x}_t}$$

Appendix A shows that:

$$\frac{\partial f_{it}(x)}{\partial \bar{x}_t} = f_{it}(x) \cdot \left(\frac{\bar{\alpha}_t x - \alpha_i \bar{x}_t}{\bar{x}_t^2}\right)$$

i.e., the partial of $f_{it}(x)$ with respect to $\bar{x}_t$ is the product of the pdf itself, $f_{it}(x)$, multiplied by a_{it}, where:

$$a_{it} \equiv \left|\frac{\bar{\alpha}_t x - \alpha_i \bar{x}_t}{\bar{x}_t^2}\right|$$

For a particular income amount x_0, $f_{i(t+1)}(x_0)$ can be approximated as:

$$f_{i(t+1)}(x_0) \approx f_{it}(x_0) + \left[\frac{\partial f_{it}(x_0)}{\partial \bar{x}_t}\right]\Delta\bar{x}_{(t+1)}$$

The partial is negative if:

$$\bar{\alpha}_t x < \alpha_i \bar{x}_t$$

A negative partial indicates that the pdf model becomes smaller in the part of the distribution where this condition obtains as $\bar{x}_t$ increases. Thus the difference

$$\bar{\alpha}_t x - \alpha_i \bar{x}_t$$

serves as a discriminant determining the sign of the partial over the pdf model, $f_{it}(x)$. The domain of x over which $f_{it}(x)$ decreases or increases as $\bar{x}_t$ increases is clearer if $\bar{\alpha}_t$ is factored out of the discriminant, leaving the difference:

$$x - \left(\frac{\alpha_i}{\bar{\alpha}_t}\right)\bar{x}_t$$

The partial derivative is negative where this difference is negative. This difference is negative for all income values x less than the product of $(\alpha_i/\bar{\alpha}_t)$ by the unconditional mean $\bar{x}_t$. Since in a gamma pdf the mean is a ratio of its shape to its scale parameter, the expression:

$$\left(\frac{\alpha_i}{\bar{\alpha}_t}\right)\bar{x}_t = \frac{\alpha_i}{\lambda_t} = \bar{x}_{it}$$

is, under the model, the conditional mean of $f_{it}(x)$, $\bar{x}_{it}$. The pdf model of the partial distribution of nonmetro wage and salary income in the ith education group, $f_{it}(x)$, decreases to the left of $\bar{x}_{it}$ as the unconditional mean, $\bar{x}_t$, increases. Vice versa to the right of $\bar{x}_{it}$.

    Since the partial derivative of $f_{it}(x)$ with respect to $\bar{x}_t$ is a product of the density by:

$$\left(\frac{\bar{\alpha}_t x - \alpha_i \bar{x}_t}{\bar{x}_t^2}\right) \equiv a_{it}$$

the expected amount of change in $f_{it}(x)$ given change in the unconditional mean depends on both how far from the conditional mean a particular income amount is and the current density at that point.

    It is hard to distinguish change in the extreme right tail from variation due to sampling error because the relative frequencies there are tiny. However, the model implies that an increase in unconditional mean nonmetro wage and salary income will change the extreme right tail more, proportionally, than any other part of the distribution because it is farther from the conditional mean than the extreme left tail or any other income amount, i.e., its a_{it}. Morris, Bernhardt, Handcock (1994) report thickening in the far right tail of the wage and salary income distribution of the U.S. as a whole from the mid-1960's through the mid-1980's, a finding

consistent with this paper's model. But while this paper's model implies the greatest proportional increase in the relative frequency of the extreme right tail as the mean increases, the model implies the greatest absolute increase in relative frequency, i.e., the most people appearing in this income range, at the extreme left tail of the distribution of the least well educated when the mean decreases. See figure 10 which shows that the largest relative frequency of any partial distribution of the distribution of nonmetro wage and salary income conditioned on education is the relative frequency in the $1- $8,000 bin of the distribution of people with at most an elementary school education. Empirically, figure 7 supports this implication.
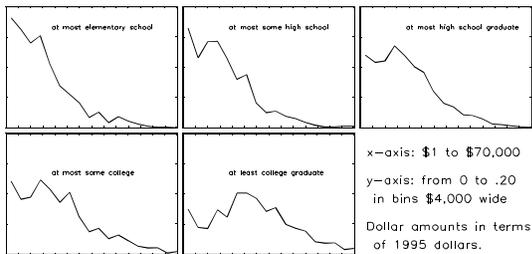
Relative Frequency Distribution of Annual Wage
and Salary Income of Nonmetro Workers in 1987

Workers Aged 25 to 65   Source: March Current Population Survey

Figure 10

Fit of Model to Data

There are 33 (1963-1995) years x 5 levels of education = 165 partial distributions of the conditional distribution, nonmetro wage and salary income conditioned on education. Each partial distribution is represented by the weighted relative frequency of incomes in each of 18 bins, from $1 to $4,000 in terms of 1995 dollars, to $68,001 to $72,000. There is a 19th bin, the relative frequency of incomes at least as great as $72,001. This bin is not fitted. So there are 165 x 18 = 2,970 observations, one for each bin. Each observation has a dependent variable, the relative frequency, and explanatory variables: the estimated bin mean, unconditional mean income in a year, and the level of education. Bootstrapped standard errors are not presented here because they are small and the length limitation on this paper severe. The model requires the estimation of the unconditional mean of incomes but only information on incomes in the range $1 to $72,000 is available. The unconditional mean is estimated by multiplying the mean of incomes $72,000 or less by the constant, k, that minimizes the squared error of the fitted model. So 6 parameters are estimated in fitting the model to the

data: 5 shape parameters, one for each level of education, and k.

The model is fitted via a stochastic search algorithm, classifiable as "simulated annealing", i.e., a stochastic search that homes in on an optimum with continual randomized but shrinking "back up's", the annealing phase of simulated annealing. This phase re-initiates the stochastic search with a less fine search grid at a distance away from the previous optimum. The metaphor is 'more random' = 'hotter'. Annealing is the strengthening of a metal by re-heating and cooling it, quenching.

There are six parameters estimated in the fit. k is estimated as 1.1974. The alphas of the five education groups are estimated as:

highest level of education	estimated α_i
elementary school or less	1.126
some high school	1.263
completed four years of high school	1.494
some college	1.631
four years of post-secondary education or more	2.331

The fit is measured as the sum of squared errors between the 2,970 observed relative frequencies and those expected under the 6 parameter model fitted. This fit is .991, about .00033 per relative frequency estimated. The sum of absolute deviations between observed and expected relative frequencies is 39.98 with a mean absolute deviation of .0135. The correlation coefficient between the observed and the expecteds is .933. Its square is .871. This 6 parameter fit to 165 partial distributions with quite different shapes, consisting of 2,970 relative frequencies, over 33 years of change, is remarkably close. It can be summed up qualitatively in the expression "lawlike". Figure 11 illustrates the fit of the model in a particular year.
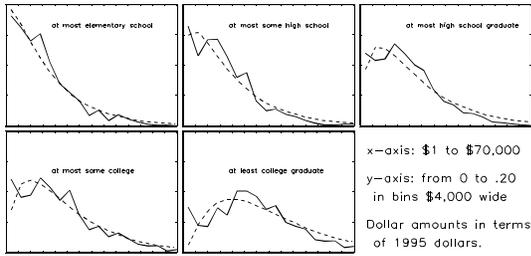
Relative Frequency Distribution of Annual Wage
and Salary Income of Nonmetro Workers in 1987
(solid curve) and Fitted Model (dotted curve)

Workers Aged 25 to 65   Source: March Current Population Survey

Figure 11

How the Model Reproduces Patterns #1,2, and 3

Pattern #1

Pattern 1 is the overlapping, i.e., approximate equality, of the time-series of three linearly transformed variables: a) the relative frequencies of incomes from $1 to 8,000, b) the relative frequency of incomes from $36,001 to $44,000, and c) the median. The left tail bin, $1 to $8,000, relative frequencies, y_{lt}, are transformed:

$$\frac{\max(y_l) - y_{lt}}{\max(y_l) - \min(y_l)}$$

The relative frequencies in the bin, $36,001 to $44,000, y_{rt}, are transformed:

$$\frac{y_{rt} - \min(y_r)}{\max(y_r) - \min(y_r)}$$

as are the observations on the median.

The transformations of the relative frequencies are approximately equivalent when $a_{ilt} = -a_{irt}$, as in:

$$a_{ilt} \equiv \left[\frac{\bar{\alpha}_t x_l - \alpha_i \bar{x}_t}{\bar{x}_t^2}\right]$$

where $x_l = \$4,000$, the midpoint of the left tail bin, and a_{irt} is the corresponding expression for the bin $36,001-$44,000 and x_r. a_{ilt} and $-a_{irt}$ are from the expression for the partial derivatives of $f_{it}(x)$ at $x = x_l$ and $x = x_r$, where $x_r = \$40,000$, and:

$$\frac{\partial f_{it}(x)}{\partial \bar{x}_t} = f_{it}(x)\left[\frac{\bar{\alpha}_t x - \alpha_i \bar{x}_t}{\bar{x}_t^2}\right] = f_{it}(x)\, a_{it}$$

Since the minimum relative frequency in the bin $36,001-$44,000 occurs in the first year, i.e., $y_{r0} = \min(y_r)$. Approximating all change in the relative

frequency in that bin as due to change in the unconditional mean, implies the following approximations to the time series of y_{rt}:

$$y_{r1} \approx y_{r0} + f'_{i0}(x_r)\Delta\bar{x}_1 \qquad f'_{i0}(x_r) = y_{r0}a_{ir0}$$

$$y_{r2} \approx y_{r1} + f'_{i1}(x_r)\Delta\bar{x}_2 \qquad f'_{i1}(x_r) = y_{r1}a_{ir1}$$

$$y_{r3} \approx \;......$$

$$......$$

or:

$$y_{r1} \approx y_{r0} + y_{r0}a_{ir0}\Delta\bar{x}_1 = y_{r0}(1 + a_{ir0}\Delta\bar{x}_t)$$

$$y_{r2} \approx y_{r0}(1 + a_{ir0}\Delta\bar{x}_1) + y_{r0}(1 + a_{ir0}\Delta\bar{x}_1)a_{ir1}\Delta\bar{x}_2$$

$$\approx y_{r0}(1 + a_{ir0}\Delta\bar{x}_1)(1 + a_{ir1}\Delta\bar{x}_2)$$

$$............$$

$$y_{rt} \approx y_{r0}(1 + a_{ir0}\Delta\bar{x}_1)(1 + a_{ir1}\Delta\bar{x}_2)\cdots(1 + a_{ir(t-1)}\Delta\bar{x}_t)$$

The transformation of the time-series is, since the y_{r0}'s are approximately monotonic increasing:

$$\frac{y_{r0}(1 + a_{ir0}\Delta\bar{x}_1)(1 + a_{ir1}\Delta\bar{x}_2)\cdots(1 + a_{ir(t-1)}\Delta\bar{x}_t) - y_{r0}}{y_{r0}(1 + a_{ir0}\Delta\bar{x}_1)(1 + a_{ir1}\Delta\bar{x}_2)\cdots(1 + a_{ir(T-1)}\Delta\bar{x}_T) - y_{r0}}$$

where T is the last time point and y_{rT} is either the maximum or close to it. The y_{r0}'s cancel out. The transformation of the left tail bin relative frequency, approximately monotonic decreasing, turns its maximum into its minimum, its decreases into increases. If its changes are, like those of the other bin, driven by changes in $\bar{x}_t$, the transformed time-series of the left tail bin relative frequency will equal, in the absence of measurement and sampling error, the transformed relative frequencies in the right tail bin when $-a_{ilt} = a_{irt}$, as was to be shown. This condition obtains where:

$$x_l - \left(\frac{\alpha_i}{\bar{\alpha}_t}\right)\bar{x}_t = -\left[x_r - \left(\frac{\alpha_i}{\bar{\alpha}_t}\right)\bar{x}_t\right]$$

$$x_l - \bar{x}_{it} = -\left[x_r - \bar{x}_{it}\right]$$

i.e., x_l and x_r are equidistant around the conditional mean. Given $x_l = \$4,000$, the weighted mean of the x_r's of the five conditional distributions fall into the bin $36,001 to $44,000 from 1963 through 1995 15 out of 33 years as $\bar{x}_t$ and $\bar{\alpha}_t$ change. $\bar{\alpha}_t$ changes, proportionally, more slowly than $\bar{x}_t$. x falls within $2,000 of this bin in an additional 16 years. See figure 12.
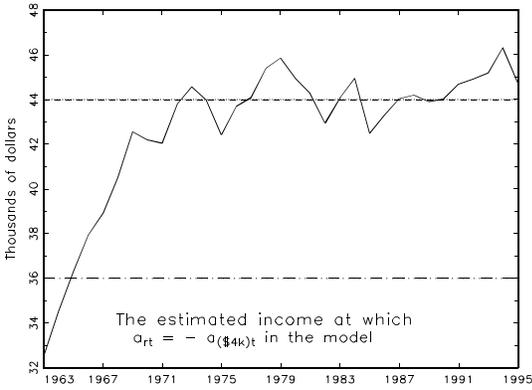
Figure 12

The transformation of median nonmetro wage and salary income overlaps the previous two transformations in figure 4. The median of each $f_{it}(x)$ is approximately, following the approximation formula for the median of a gamma pdf (Salem and Mount, 1974):

$$m_{it} \approx \left(\frac{3\alpha_i - 1}{3\lambda_t} \right) = \left(\frac{3\alpha_i - 1}{3\bar{\alpha}_t} \right) \bar{x}_t$$

The weighted mean of the m_{it}'s is used as an estimate of the median of the unconditional distribution, m_t. Since:

$$\frac{\partial m_{it}}{\partial \bar{x}_t} \approx \left(\frac{3\alpha_i - 1}{3\bar{\alpha}_t} \right)$$

the following approximations can be made:

$$m_{i1} \approx m_{i0} + \left(\frac{3\alpha_i - 1}{3\bar{\alpha}_t} \right) \Delta \bar{x}_1$$

Given a minimum at m_{i0} and that m_{iT} is approximately the maximum, the transformation:

$$\frac{m_{it} - \min(m_i)}{\max(m_i) - \min(m_i)} = \frac{\sum_{i=0}^{i=t} \frac{\Delta \bar{x}_i}{\bar{\alpha}_i}}{\sum_{i=0}^{i=T} \frac{\Delta \bar{x}_i}{\bar{\alpha}_i}}$$

This expression is not equal to the relative frequencies but like the expression for the transformed relative frequencies is a function of a ratio of sums involving $\Delta \bar{x}_t$.

Pattern #2

It has been shown that the relative frequencies of bins in the left and right tail where $a_{irt} = - a_{ilt}$ are linear combinations of each other varying in opposite directions, i.e., one goes up, the other goes down, and vice versa. Where the

condition, $a_{irt} = - a_{ilt}$, is met exactly, they are correlated -1.0. The correlations of relative frequencies at either x_l in the left tail or x_r in the right tail with a relative frequency near the conditional mean, $\bar{x}_{it}$, will be close to zero since a_{it} will be close to zero.

Pattern #3

Pattern #3 follows immediately from the expression for the partial derivative of a relative frequency with respect to $\bar{x}_t$:

$$\frac{\partial f_{it}(x)}{\partial \bar{x}_t} = f_{it}(x) \left[\frac{\bar{\alpha}_t x - \alpha_i \bar{x}_t}{\bar{x}_t^2} \right]$$

Change in the relative frequency at x_0 from time t to t+1 is approximated as:

$$f_{it}(x_0) \left[\frac{\bar{\alpha}_t x_0 - \alpha_i \bar{x}_t}{\bar{x}_t^2} \right] \Delta \bar{x}_{(t+1)}$$

The shock comes from $\Delta \bar{x}_{(t+1)}$ and is multiplied by $f_{it}(x_0)$. Consequently, the variance of the relative frequency at x_0 is proportional to the relative frequency at x_0.

One-Step Ahead Out-of-Sample Forecasts

Figure 13 shows the mean absolute error per relative frequency forecasted in one year ahead forecasting of the nonmetro distribution of wage and salary income by level of education using this paper's model. The first forecasted year is 1973 and its forecast is done with data from 1963 to 1972 inclusive. Each partial distribution of the conditional distribution has 18 relative frequencies, from bin $1 to $4,000 through bin $68,001 to $72,000. There are five partial distributions, one for each level of education. After the initial forecast, each year's data are added to the database. Each forecast of the next year's distribution requires as input, besides parameters estimated from its database to date, a forecast of the next year's mean nonmetro wage and salary income. Perfect foreknowledge of next year's mean is estimated to be next year's mean of incomes between $1 and $72,000 multiplied by the parameter (estimated from database to date: k = 1.197) that maximizes the fit between distribution and the fitted gamma pdf. The curve of mean absolute error per relative frequency forecasted given perfect foreknowledge of next year's mean is plotted in figure 13 along with two other curves. One is the mean absolute error per relative frequency with an estimate of next year's mean nonmetro wage and salary income that is 5% too high. The other curve is the mean absolute error per relative frequency with an estimate of next year's mean nonmetro wage and

salary income that is 5% too low. The maximum year to year increase in the mean as estimated in the period 1963 to 1995 is 5.4%. The maximum year to year decrease in the estimated mean in the period 1963 to 1995 is 5.1%. Notice that even with the over or underestimates, forecasts based on this paper's model have small errors, almost as small as that of the model fitted to all 33 years of data simultaneously, with many fewer years of data. The last forecasts in figure 13, those for 1995, are based on parameters estimated with 32 years of data. The fitting of the model for forecasting is done the same way as the fitting of the model to all 33 years of data.
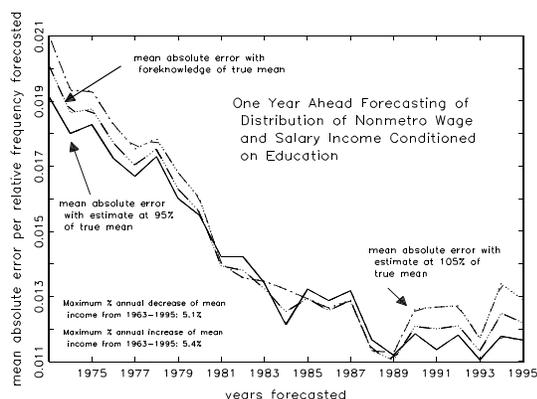
Conclusions and the Scenario of a 40% Fall in the Unconditional Mean

Because this paper's model is able to reproduce patterns #1, 2, 3 in the dynamics of the tails of the distribution of nonmetro wage and salary income, and because it shows that foreknowledge of mean nonmetro wage and salary income is tantamount to foreknowledge of the distribution, this paper's forecast of how the nonmetro wage and salary income distribution changes shape if there is a major shift in the time-series of mean wage and salary income is credible. Figure 14 shows the differences, bin by bin, to the 1995 distribution, nonmetro wage and salary income conditioned on education if the unconditional 1995 mean is decreased 40%. The right tails of all the partial distributions become thinner and the left tails bulge. Figure 14 shows that this response is most pronounced in the right tail among the most educated and most pronounced in the left tail among the least well educated. Indeed the increase at the extreme of the left tail, the smallest incomes, is quite large, indicative of a shift of a substantial fraction of the nonmetro population of the least well educated in the bin of the smallest incomes, $1 to $4,000. Note the part of the nonmetro population who are at least college graduates is, as a whole,

considerably more buffered from this shift to the lowest income bin than less well educated people.

This paper's model has been about the part of the U.S. population with a residence in a nonmetro county because that is the focus of interest of the Economic Research Service. This paper's model is not specialized in any way for this population though and works equally well for the metro part of the U.S. population.
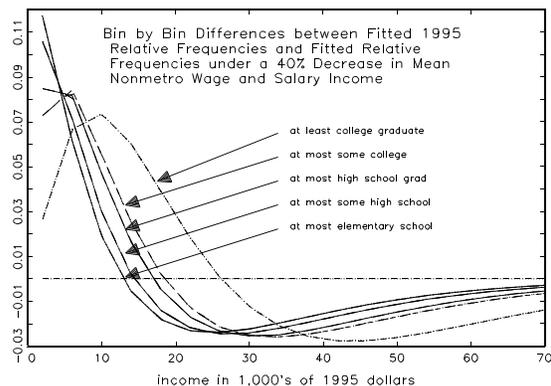
APPENDIX A: The Partial Derivative of $f_{it}(x)$ With Respect to $\bar{x}_t$

The partial derivative of $f_{it}(x)$ with respect to $\bar{x}_t$ gives an expression for how $f_{it}(x)$ changes as a function of $\bar{x}_t$.

$$
\begin{aligned}
f_{it}(x) &= \frac{\lambda_t^{\alpha_i}}{\Gamma(\alpha_i)} x^{\alpha_i-1} e^{-\lambda_t x} \\
&= \exp\bigl(\alpha_i \ln(\bar{\alpha}_t) - \alpha_i \ln(\bar{x}_t) \\
&\quad -\ln(\Gamma(\alpha_i)) + (\alpha_i-1)\ln(x) \\
&\quad -\bar{\alpha}_t\left(\frac{x}{\bar{x}_t}\right) \bigr)
\end{aligned}
$$

where,

$$
\lambda_t = \frac{\bar{\alpha}_t}{\bar{x}_t}
$$

and,

$$
\bar{\alpha}_t = w_{1t}\alpha_1 + w_{2t}\alpha_2 + \ldots + w_{It}\alpha_I
$$

So,

$$\frac{\partial f_{it}(x)}{\partial \bar{x}_t} = \exp\Big(\alpha_i \ln(\bar{\alpha}_t) - \alpha_i \ln(\bar{x}_t)$$
$$-\ln(\Gamma(\alpha_i)) + (\alpha_i - 1)\ln(x)$$
$$-\bar{\alpha}_t\left(\frac{x}{\bar{x}_t}\right)\Big) \quad \cdot$$
$$\Big(\alpha_i \ln(\bar{\alpha}_t) - \alpha_1 \ln(\bar{x}_t)$$
$$-\ln(\Gamma(\alpha_i)) + (\alpha_i - 1)\ln(x)$$
$$-\bar{\alpha}_t\left(\frac{x}{\bar{x}_t}\right)\Big)'$$

and,

$$\Big(\alpha_i \ln(\bar{\alpha}_t) - \alpha_1 \ln(\bar{x}_t)$$
$$-\ln(\Gamma(\alpha_i)) + (\alpha_i - 1)\ln(x)$$
$$-\bar{\alpha}_t\left(\frac{x}{\bar{x}_t}\right)\Big)' =$$
$$= \left(\frac{-\alpha_i}{\bar{x}_t} + \bar{\alpha}_t\left(\frac{x}{\bar{x}_t^{\,2}}\right)\right)$$
$$= \frac{\bar{\alpha}_t\, x - \alpha_i\, \bar{x}_t}{\bar{x}_t^{\,2}}$$

Thus,

$$\frac{\partial f_{it}(x)}{\partial \bar{x}_t} = f_{it}(x) \cdot \left(\frac{\bar{\alpha}_t\, x - \alpha_i\, \bar{x}_t}{\bar{x}_t^{\,2}}\right)$$

REFERENCES

Angle, J. 1994. "Frequency spikes in income distributions." **1994 Proceedings** of the Business and Economic Statistics Section of the American Statistical Association. 265-270.

Council of Economic Advisers. 2002. **Economic Report of the President**. Washington, DC: U.S. Government Printing Office.

Fuguitt, Glenn, David Brown and Calvin Beale. 1989. **Rural and Small Town America**. New York: Russell Sage.

Morris, Martina, Annette Bernhardt, and Mark Handcock. 1994. "Economic inequality: new methods for new trends". **American Sociological Review** 59: 205-219.

Sahota, Gian. 1978. "Theories of personal income distribution". **Journal of Economic Literature** vol. 16, #: 1-55.

Salem, A.B.Z. and T.D. Mount. 1974. "A convenient descriptive model of income distribution: the gamma density". **Econometrica** 42: 1115-1128.

Stauber, Karl. 2001. "Why Invest in Rural America-And How? A Critical Public Policy Question for the 21st Century". **Economic Review of the Federal Reserve Board of Kansas City**. Second Quarter, 2001: 33-63.

This page intentionally left blank.

FORECASTING THE BUSINESS CYCLE WITH POLAR COORDINATES

Foster Morrison and Nancy L. Morrison
Turtle Hollow Associates, Inc.
PO Box 3639
Gaithersburg, MD 20885-3639 USA
Phone: 301-762-5652   Fax: 301-762-2044
Email: 71054.1061@compuserve.com

1. The Dynamics of Forecasting

Forecasting is a specialized part of the broad category of predicting the future state and behavior of dynamic systems. What defines forecasting is the particular dynamic systems in question, namely economic, business, social, and even political conditions.

Prediction in the physical sciences may be quite accurate and reliable, with the orbits of the major planets being the most ancient example, as well as one of the more accurate ones. But accuracy is the exception rather than the rule, even in the physical sciences. Weather forecasting still is not as reliable as desired, despite great improvements from radar observations, satellite photography, and computer models.

Climatology is the specialty within meteorology that deals with the changes in aggregated variables over long periods of time. In other words, has the weather been getting hotter or colder over the past 500 years? This is now a serious debate known as "global warming." The questions raised by possible climate change have important implications for economic, business, social, and political issues, so the interests of the geophysical and forecasting communities have a point of convergence there.

Economics and ecology are really two parts of the same system, the terrestrial biosphere. It would be absurd to attempt to forecast for one and not the other, at least for time periods of more than a few years. The famous "limits to growth" debates of the mid-1960s were inconclusive because the real issues were clouded by the novelty of large-scale computer models. If there is one thing that can be said with virtual certainty, it is that exponential (constant rate) growth is not sustainable. But predicting or planning for the end of growth is horrendously complicated, just like the questions about global warming.

The growing interaction of the areas of traditional forecasting with geophysics and ecology is one of the factors producing the problems addressed by the theme of this Twelfth Federal Forecasters Conference: "Major Shifts: Discontinuity, Uncertainty, and Forecasts." (Another is the collapse of the global balance of power that came with the end of the "cold war.") The novelty

factor now might be chaos theory (Gleick, 1987), but forecasters (and scientists) should not allow themselves to be distracted by something that really is of marginal significance. Most of the data series in question are some mix of deterministic components and filtered noise.

To get back to basics, consider the noise-driven linear difference equation with constant coefficients

$$\mathbf{x}_{i+1} = \mathbf{A}\mathbf{x}_i + \mathbf{n}_i \qquad (1)$$

The column vector $\mathbf{x}$ consists of the variables being forecast, with the subscript representing time in uniform (or nearly uniform) steps. The square matrix $\mathbf{A}$ has constant elements. The noise vectors $\mathbf{n}_i$ consist of one or more components which are "random" numbers, usually with a zero mean and other constant statistics.

The various types of forecasting models consist of constraints on the elements of $\mathbf{A}$. When more than one previous value of a variable are used in the forecast, they must be on the "right-hand side" of (1), and hence all but the earliest must be on the "left." This requires off diagonal elements in $\mathbf{A}$ equal to 1 (to shift earlier values in the "right-hand" vector down by one in the "left-hand" vector), with the rest in that row being 0. There is good reason for doing this rather than having $\mathbf{A}$ not be square and the vectors of variables differing in dimension, especially for purposes of stability analysis.

For (1) to be dynamically stable it is necessary that all the (possibly complex) eigenvalues of $\mathbf{A}$ be of absolute value less than 1. Note that if all the elements of $\mathbf{A}$ are real, some (or all) of its eigenvalues may be pairs of complex conjugates. Hence, it would be useful to have all the popular forecasting models expressed in the form (1), along with a subroutine for finding eigenvalues. (For details on established forecasting methodologies see Makridakis & al., 1998.)

2. Trend Modeling

For most applications of (1) as a forecasting model, the data have the trend removed. The dynamical properties of this difference equation tell one why this should be done. For one thing, the forecast is generated by extrapolating (1) with $\mathbf{n}_i = \mathbf{0}$. This is justified, since the expected value of the noise is generally assumed to be zero [$\mathbf{E}(\mathbf{n}_i) = \mathbf{0}$] and usually it is. The signal power of

$\mathbf{n}_i$ is not zero $[E(\mathbf{n}_i^T\mathbf{n}_i) > 0]$ and this generates the generally increasing variances of a forecast that approaches zero asymptotically.

Polynomial regressions can be very good trend models, except for the fact that they are not suitable for extrapolation unless they are first order (linear in time). However, this criterion can be met for some data series that are sufficiently short. In other cases, using the logarithms of the data may make a linear trend model adequate.

Low-pass filters have much to recommend them as trend models. The simplest example, the moving average, creates problems at the beginning and end of the data, since the proper reference point for it is in the middle of the time span being averaged. Otherwise, the trend of a straight line of data would lie above or below the data. This difficulty can be avoided through use of the ramp filter, whose proper reference point is the end of the data set (Morrison and Morrison, 1997).

Using the ramp filter requires discarding enough data at the beginning of the series to generate the first filtered point. If the data span is short, there may be fewer data points left than one needs or wants to get the coefficients in the matrix **A** in (1). In that case it may be advantageous to use a higher-order polynomial as a trend model for the data points otherwise discarded, adding a constraint to make it match the first ramp filtered point.

In any case, extrapolating the ramp filter as the trend is dynamically stable. The forecast does not soar or sink like a higher-order polynomial.

3. Advanced Forecasting Methods

There are two ways to apply (1) to forecasting. The independent elements in **A** may be determined by the method of least squares, which will then provide components of $\mathbf{n}_i$ from the residuals. When statistical tests show that relevant components of $\mathbf{n}_i$ are sufficiently "random," the forecast is considered optimized.

Constructing the best possible forecast still is judgmental, despite the variety of statistical tests available. Less well known to practitioners is the fact that randomness is not a rigorous scientific or mathematical concept (Kac, 1983, 1984). In fact, the acceptable level of rigor in mathematics is not totally settled (Bishop, 1975; Bishop and Bridges, 1985), though it is not a critical factor in forecasting accuracy.

Linear filtering forecasts also can be constructed without ever determining **A**. Filter coefficients can be estimated using the correlation function of the data and cross correlation functions, where there are several data series. These statistical functions can be obtained using

fast Fourier transforms (FFTs are more efficient, if less flexible, than computing Fourier transforms by brute force). Error estimates are obtained with almost no extra effort (Morrison, 1991).

Such linear filtering methods are widely used in geodesy and geophysics (Jordan, 1972; Morrison, 1977; Morrison and Douglas, 1984). When combined with the ramp filter as the trend model, they can be effective with long data series such as stock market indices, the GDP, and the indices of leading, lagging and coincident indicators.

Stability can be guaranteed by having the statistical functions asymptotically approach zero for long times. Since time lags longer than 32 steps are rarely used, this is achieved *de facto*. Of course, it is a good idea to have at least 64 points for the FFTs and another 40 to 60 points for the ramp filter. This has been no problem with the economic data we have analyzed, but it could be in some cases, say the average monthly exchange rates for the euro. Using a polynomial trend model for the first 40-60 points would solve that problem.

For higher-order polynomial models, appropriately scaled Legendre functions have much to offer, since they are orthogonal over the interval $[-1, +1]$ and almost orthogonal over uniformly spaced points covering that interval (Jahnke and Emde, 1945; Abramowitz and Stegun, 1964). This means that the matrix to be inverted to solve for the coefficients is almost orthogonal and hence very well determined, even for higher orders.

Nonlinear terms could be added to (1) quite easily, but the stability analysis becomes more challenging. In this case one might want to consider an ordinary differential equation (ODE) model, starting with the linear system

$$d\mathbf{x}/dt = \mathbf{Bx} + \mathbf{n}(t) \qquad (2)$$

This would also have some advantages when the data series have different sampling rates, say, both quarterly and monthly, or irregular sampling rates (like daily stock market closings).

A model for nonzero components of $\mathbf{n}(t)$ could be an expansion in orthogonal functions, such as the Chebyshev polynomials. Runge-Kutta (single step) methods probably would suffice for solving (2) numerically, but it might be informative to try CNC (continuous numerical continuation) which also is an application of Chebyshev approximation (Morrison and Morrison, 2000).

Very large, nonlinear ODE models have been used for economic and ecological modeling (Forrester, 1961), but the applications have been long-term policy analysis rather than forecasting. Such efforts created a new discipline known as System Dynamics (Wils, 1988). Further research in this area might prove fruitful if

practitioners used stability analysis (Bellman, 1969; Luenberger, 1979; Morrison, 1991) rather than just numerical solutions. Another caveat is that the models should be small in dimension and not attempt to simulate $\mathbf{n}(t)$ deterministically.

An alternative to using nonlinear forecasting equations is doing a nonlinear transformation. For some problems in celestial mechanics, nonlinear transformations have simplified perturbation theories (Brouwer and Clemence, 1961), but they still have been displaced by numerical methods.

Orbits with small inclinations or eccentricities may be modeled by converting from Keplerian elements that are essentially polar coordinates to similar Cartesian coordinates. For example, the eccentricity e and the argument of perigee may be replaced by

$$= e \sin \qquad (3.1)$$

$$= e \cos \qquad (3.2)$$

This is necessary because the angle becomes increasingly ill-defined as e gets smaller and the elliptical orbit collapses into a circle. One symptom of this problem is perturbation terms that have e as a divisor.

4. The Business Cycle: a Basic Macroeconomic Model

There are many factors limiting the optimal dimension of computer models of complex systems. One is the fixed precision of floating point arithmetic. The popular Intel CPUs have been supporting 80-bit floating point, but some others support only the 64-bit version. Higher precisions can be obtained only by resorting to assembly language codes that require a high level of programming sophistication and produce a great increase in execution time (maybe a factor of 100).

Increased floating point precision cannot solve all the problems of computer modeling and forecasting, but there also is no fixed level of precision that is always adequate. The new Intel 64-bit (instruction set) CPUs will evict the floating point operations and put us back where we were in the days of the 8086, 80286, 80386, and some of the 80486 models. Hopefully, Intel or some independent vendor, will provide a floating point coprocessor (recall the old 8087 and 80287 models) and maybe one that will support a lot more than 80-bit precision.

Other factors include the limited accuracy of the data sets, the inadequacy of the data collection processes, and the modeling errors. Computers can speed up computations enormously and eliminate annoying errors (as well as add annoying errors), but they can do nothing to alleviate these deficiencies. Other new technologies, such as the Internet, can be helpful, but the traditional approaches used in celestial mechanics and some other physical sciences where the phenomena are deterministic (or nearly so) cannot be generalized for applications to things like economic forecasting or policy analysis.

What can be done to develop the best feasible models of large, complex systems? Aggregation is essential, first to reduce the noise level in the data, and also to keep the dimension of the model small enough so that it is as numerically stable and controllable as possible. The most common form of aggregation is a linear combination, such as the popular stock market indices. Economists have created others, such as the GDP (gross domestic product).

For example, the modeler should know what the eigenvalues are for the matrices $\mathbf{A}$ or $\mathbf{B}$, if a linear model is used. Nonlinear models, when linearized, produce a state-transition matrix (partial derivatives of state variables with respect to initial conditions), and its eigenvalues should be determined.

All equilibrium points in nonlinear models should be identified, as well as their Liapunov exponents (eigenvalues from the models linearized around these points). There also are methods that can identify whether the flow of aggregated solutions is in compression or of constant volume (or hypervolume) (Morrison, 1991; Thompson and Stewart, 1987). In this context aggregation means looking at the evolution of sets of solutions, rather than a single one. This approach, called topological dynamics, offers a number of practical tools, if fewer than one might desire.

In the case of the US economy, a set of three aggregated variables has been created through decades of efforts by economists in government, the private sector, and the academic community. These are the indices of leading, coincident, and lagging indicators (*Handbook*, 1984). They are not ideal, but they are the best thing available, if not the best thing possible.

To create a model of the business cycle we detrended these indices and then created a phase plane model. The first attempt used only the leading and coincident indices (Morrison and Morrison, 1997). Then we added the lagging index and projected the three-dimensional trajectory onto a best-fitting plane (Morrison and Morrison, 2001). For forecasting we have been using time series methods rather than a difference equation (1).

5. Forecasting Polar Coordinates

Forecasts are constructed using (1) by setting $\mathbf{n}_i = \mathbf{0}$ and then simply executing the matrix-vector multiplication for as many steps as desired. If all the

eigenvalues of **A** are indeed of absolute value less than 1, then the vector $\mathbf{x}_i$ decays asymptotically to **0**. At the same time the error estimates (a matrix of variances and covariances) will approach constant values; this is self-evident in the time series equations.

The solutions of (1) are often "cyclical" functions, but with irregular periods. This characteristic is found in most *detrended* economic time series and has led some investigators to conclude that the average period is something other than a statistical coincidence. But this is rarely the case, since a spectral analysis usually reveals that the power spectrum is a fairly smooth function that gradually decays for ever higher frequencies. This is a defining property of *filtered noise*.

Exceptions include a few things like seasonality. Many economists smooth their data to eliminate these effects.

Cycles with highly stable periods are extremely predictable, exhibiting very gradual deterioration in forecasts for the phase angle. This is why planetary orbits could be predicted quite accurately by the Ptolemaic method, basically a three-dimensional spectral analysis approach to curve fitting.

Filtered noise has variable periods, so the variance of the phase angle grows rather rapidly. Forecasts often lose useful precision after only one cycle and sometimes in much shorter periods. Making the model (1) larger or more sophisticated (by adding nonlinearities, *e.g.*) may improve the results slightly, but after a certain point uncontrollable numerical instability takes over.

Using weighted (or unweighted) averages is one way to improve the precision of forecasts. The S&P 500 index is more predictable (and stable) than most of its component common stocks. But is this information useful to investors? It is now, since index mutual funds are available.

Nonlinear transformations offer even more possibilities for improving forecasting precision. In the case of the business cycle, the polar coordinates can be used as the forecasting variables rather than the Cartesian-like percent deviations from the trend of the three composite indices.

The first step is to rotate the x-y plane about the y-axis to achieve a least squares fit to the data, $\mathbf{r} = (x, y, z)^\mathrm{T}$

$$\mathbf{r}_2 = \mathbf{Gr} \tag{4}$$

where x, y, and z are the detrended indices of leading, coincident, and lagging indicators; **G** is the appropriate rotation matrix. We have been using a 60-point ramp filter for the trend.

The z_2 coordinate is discarded and the phase plane

variables are defined by

$$= (x_2{}^2 + y_2{}^2)^{\frac{1}{2}} \tag{5.1}$$

$$= \tan^{-1}(y_2/x_2) \tag{5.2}$$

For purposes of forecasting, it should be noted, the phase angle, (in radians), should not be constrained to the range [0, 2). If this were done, the phase angle would look like a sawtooth function to any forecasting algorithm instead of something rather close to a linear function of time.

A long-term forecast for   and   would not spiral into the origin, like (1), but would have the asymptotic properties

$$_i \tag{6.1}$$

$$_i \tag{6.2}$$

(The   operator indicates the change between successive values.)   In other words, the asymptotic forecast (i) does not decay to the origin, but uniform circular motion. This is a result of the fact that the expected value of the input noise in (1) is zero, but the expected values of the "energy" of the noise (the squares of the components) are positive.

Unless you believe that the business cycle will be eliminated by financial policy, monetary policy, the wonders of just-in-time delivery, or as yet unknown benefits of information technology, the long-term forecast, (6.1) and (6.2), is much better than one of spiraling into the origin ($_i$ 0). However, this does not necessarily mean that short-term forecasts will be more accurate.

The graphs of the business cycle (Morrison and Morrison, 1997, 2001) strongly suggest that the noise is nonstationary, *i.e.*, it varies in both amplitude and frequency content. This is one reason that forecasting can be an extremely frustrating activity. In fact, the conference theme of "discontinuity and uncertainty" reflects the reality that the noise can be violent impulses as well as a hum of varying pitch. And it should be reassuring that the scope of the problem fits so well into the jargon of the mathematical modeling of dynamic systems.

The discipline that studies violent impulses is history. There is more than 5000 years of history, but very little in terms of economic time series. Extensive data collection did not really begin until after World War II. However, some data are available from earlier centuries. It might be possible to construct business cycle models for some countries over long periods of time.

Nonlinear difference (or differential) equations are the

proper tool for studying the effects of large impulses. Where the System Dynamics school went astray was in making the models much too large and also, perhaps, in trying to predict the reversals of long-term trends.

A second principle to follow is that the approach should be simulations rather than forecasts. In other words, if the business cycle is energized by a large impulse, how long will it take to settle back into its normal range of amplitudes. The Great Depression, which was global in scope and started earlier than 1929 in Europe, provides some quantitative data and a lot of qualitative and anecdotal material.

The future of forecasting is not attaining ever higher accuracy, because that is not possible. Following in the steps of the pioneers of System Dynamics and creating methods for quantitative policy analysis is feasible, but one must recognize the limits to modeling imposed by nonlinearity and inadequate data sets.

6. Preliminary Forecasts of the Business Cycle

For our first forecasts of the business cycle using polar coordinates, we adapted our time series computer program to the altered mathematical behavior of the new variables. The trend of the radial coordinate was modeled with an 82-point ramp filter to match the average length of a cycle. This just about eliminates any dynamical effects due to imperfect scaling in the

three indices. A 32-point ramp filter was used for the phase angle because that is adequate for something so close to a linear function of time. For the phase angle 32 points were used for the prediction filter and 16 for the radial coordinate. To forecast the three indices we have used a 32-point ramp filter for the trend and 32 points in the prediction filters.

Forecasts for the most recent data available are shown in Figure 1. The polar coordinate forecast moves more rapidly than the Cartesian alternative, suggesting a faster recovery from the recession. Our intuitive view is that this is overly optimistic, since a double-dip recession may be in progress.

Our preliminary conclusion is that the nonstationary nature of the noise may make the polar coordinate forecast more accurate in some cases, but that the Cartesian variable forecast may work better when the cycle is in a stall. The best feasible forecast perhaps may be attained by using nonlinear difference (or differential) equations for the polar coordinates, but with a nonstationary model of the noise. A forecast of the variance-covariance matrix for the noise is achievable and maybe even one for the components of the noise.
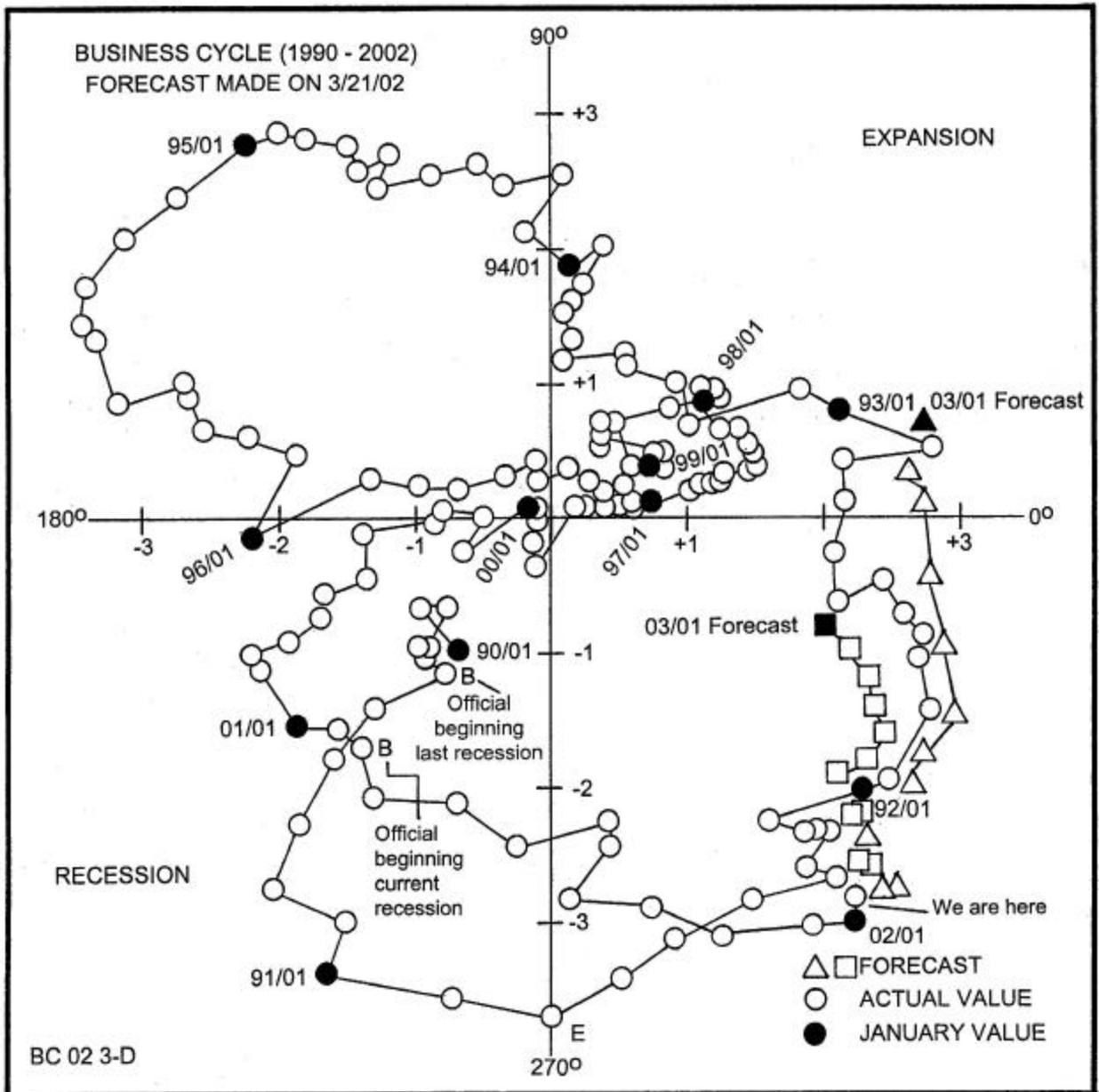
Figure 1. The business cycle model is a phase plane plot of a weighted mean of the detrended leading and detrended lagging indicators as x-coordinate and detrended coincident indicator as y-coordinate. Normal cycles follow a counterclockwise roughly circular path with occasional stalls and reversals. Time is indicated along the cycle path. The data have a 2-month lag. Expansions occur between 0° and 90° and recessions between 180° and 270°. Other angles denote transition (90°-180°) and recovery (270°-360°=0°) periods. An "official" (NBER) beginning of a recession is indicated by a label "B" and an end by "E".

The polar coordinate forecast (- triangle) moves more rapidly than the Cartesian alternative (- square), suggesting a faster recovery from the recession. Our intuitive view is that this is overly optimistic, since a double-dip recession may be in progress. Unless you believe that the business cycle will be eliminated by financial policy, monetary policy, the wonders of just-in-time delivery, or as yet unknown benefits of information technology, the long-term polar coordinate forecast is much better than one of spiraling into the origin. However, this does not necessarily mean that short-term forecasts will be more accurate.

7. References

Abramowitz, M. and I.A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, NBS Applied Math. Series 55, US Govt. Printing Office, Washington, DC, 1964.

Bellman, R., *Stability Theory of Differential Equations*, Dover Publications, New York, 1969.

Bishop, E., The crisis in contemporary mathematics, *Historia Math.*, *2*, 507-517, 1975.

Bishop, E. and D. Bridges, *Constructive Analysis*, Springer-Verlag, Berlin, 1985.

Brouwer, D. and G.M. Clemence, *Methods of Celestial Mechanics*, Academic Press, New York, 1961.

Forrester, J.W., *Industrial Dynamics*, The M.I.T. Press, Cambridge, MA, 1961.

Gleick, J., *Chaos: Making a New Science*, Viking Penguin, New York, 1987.

Handbook of Cyclical Indicators: a Supplement to the Business Conditions Digest, 1984, US Dept. of Commerce, Bureau of Economic Analysis, Washington, DC.

Jahnke, E. and F. Emde, *Tables of Functions: with Formulae and Curves*, 4th edition, Dover, New York, 1945.

Jordan, S.K., Self-consistent statistical models for the gravity anomaly, vertical deflections, and undulations of the geoid, *J. Geophys. Res.*, *77*, No. 20, pp. 3660-3670, 1972.

Kac, M., What is random?, *American Scientist*, *71*, 405-406, 1983.

Kac, M., More on randomness, *American Scientist*, *72*, 282-283, 1984.

Luenberger, D.G., *Introduction to Dynamic Systems: Theory, Models and Applications*, Wiley, New York, 1979.

Makridakis, S., S.C. Wheelwright and R.J. Hyndman, *Forecasting: Methods and Applications*, third edition, John Wiley & Sons, Inc., New York, 1998.

Morrison, F., Azimuth-dependent statistics for interpolating geodetic data, *Bulletin Géodésique*, *51*, No. 2, 105-188, 1977.

Morrison, F., *The Art of Modeling Dynamic Systems: Forecasting for Chaos, Randomness, and Determinism*, Wiley-Interscience, New York, 1991.

Morrison, F. and B. Douglas, A comparison of gravity prediction methods on actual and simulated data, *Geophysics*, *49*, No. 10, 1774-1780, 1984.

Morrison, F. & N.L. Morrison, A phase plane model of the business cycle, *The 8th Federal Forecasters Conf. 1996 & The 7th Federal Forecasters Conf. - 1994: Combined Papers & Proceedings*, US Dept. of Education, NCES 97-341, pp. 93-112, Debra E. Gerald, editor, Washington, DC, 1997.

Morrison, F. & N.L. Morrison, Continuous numerical continuation, *The 10th Federal Forecasters Conf. - 1999*, US Dept. of Education, NCES 2000-017, pp. 183-190, Debra E. Gerald, editor, Washington, DC, 2000.

Morrison, F. & N.L. Morrison, An improved phase plane model of the business cycle, *The 11th Federal Forecasters Conf. - 2000*, US Dept. of Education, NCES 2001-036, pp. 183-191, Debra E. Gerald, editor, Washington, DC, 2001.

Thompson, J.M.T. and H.B. Stewart, *Nonlinear Dynamics and Chaos: Geometrical Methods for Engineers and Scientists*, Chichester, UK, 1987.

Wils, W., *Overview of System Dynamics the World Over*, Croon DeVries, Parkweg 55, 3603 AB Maarssen, Netherlands, 1988.

This page intentionally left blank.

Industry Issues

Chair: Annette Clauson, Economic Research Service, U.S. Department of Agriculture

Two Measures of Induced Employment

Art Andreassen, Bureau of Labor Statistics, U.S. Department of Labor

The industry employment that is generated in the creation of a particular industry's output can at present be calculated. However, this employment does not include the additional amount generated when employees spend their wages, i.e., the induced employment. This paper presents two possible approaches to derive a measure of this employment. Both rely on using certain national income account relationships to convert industry value added to PCE, personal consumption expenditures. The first approach assumes average relationships while the second marginal ones in their applications to the 2000 I/O tables. The marginal approach will be extended to offer insights into the present economic downturn and recovery.

Economic Implications of Future Years Defense Purchases

Douglas S. Meade, INFORUM, University of Maryland
Ron Lile, Office of the Secretary of Defense, U.S. Department of Defense

Both inside and outside the Pentagon, defense policy analysts are interested in the economic implications of planned defense purchases. The Defense Employment and Purchases Projections System (DEPPS) was designed to help analysts in government and business understand how industries, States, and occupational groups are affected by changes in the defense budget. DEPPS consists of an interindustry model, a State model and an occupational model. The interindustry model consists of the Inforum detailed interindustry model *Iliad*, joined with the defense translator, a matrix that translates outlays on detailed defense budget programs to the industries that directly supply these programs. The State model distributes defense spending by industry to the state level, based on state shares derived from historical data. The occupational model translates defense related employment by industry to the occupational level. The DEPPS projections are made for each Future Year Defense Program (FYDP), and published on the DoD web site. This presentation will show samples of results produced by DEPPS, and describe how they were produced.

"This time we've got it right": Forecasting Asbestos Claims Against U.S. Corporations, 1985-2002

Timothy Wyant, Ravenstat, Inc.

Litigation of asbestos disease claims commenced more than 25 years ago. Asbestos exposure in the workplace had by then effectively ended. In the last 15 years, experts have repeatedly forecast future disease claims in numerous bankruptcy proceedings, and for the numerous trusts set up to distribute funds to claimants. Discontinuities in claims trends have obliterated most of these forecasts, and continue to plague claims forecasts being made today. The Manville Trust, in operation since the 1980s, in 2001 doubled its future claims forecasts. Nine major corporations declared bankruptcy in 2000-2001 due to projected asbestos liabilities. Pressure for a federal resolution to asbestos litigation is continuing, and will likely increase.

This page intentionally left blank.

Two Measures of Induced Employment

Arthur Andreassen
Bureau of Labor Statistics
U.S. Department of Labor

Introduction

Employment produces output to sell to final users and to intermediate users as inputs in their production process. The most obvious connection of spending to employment involves direct purchases, e.g., automobiles, and the employment necessary to produce the steel, rubber, glass, etc., which go into the auto. Less so is the employment that produces the inputs necessary to produce that glass, etc. In fact, almost one half of the economy's output is sold as inputs for further processing. Input/output analysis, an extension of the National Income Accounts, measures all of these levels of employment. Specialized tables calculate the employment effects of demand spending and are used by planners to compare the employment impact of different programs. Knowing how much employment a program generates in production begs the further question of how much employment is generated as these employees then spend their wages, the induced employment. This paper presents two possible solutions.

Although both approaches differ in important aspects they are similar at their beginnings. The only sources of the industry level data that are necessary for this study are input/output tables and their use necessitates certain assumptions to get past data shortcomings, especially as to timeliness. Among the specialized tables that compose the input/output structure is a Use Table that measures both the material and the factor inputs which go into an industry's output. Basic is the derivation of industry Personal Consumption Expenditures (PCE) due to the spending of industry wages. To get to PCE by industry its relationship to industry output must be derived. First, a ratio of industry compensation to industry value added is calculated. Appropriate measures of the compensation and value added by industry are available only from benchmark I/O tables, the last of which was published in 1992. From this table the ratio in 1992 was obtained. Inevitably, this ratio of compensation to value added, especially at the industry level, would have changed from 1992 to 2000, but the extent this is so will not be known until the publication of an updated benchmark table. However, and maybe surprisingly, the ratio of total

compensation to total value added is a constant 58% in both years offering some level of assurance that the assumption of continuity is not disastrously wrong. Updated I/O tables for 2000 contain all the remaining required data and so from this point forward the relationships are timely. A 2000 industry ratio of value added to output was calculated and applied to the previous industry compensation to value added ratio giving industry compensation to output ratios. From these ratios that get us the portion of industry compensation in industry output we then must go PCE by industry. By applying certain relationships in the National Income Accounts at a national level such ratios are derived. Compensation consists of wages and salaries and benefits and it is from wages that PCE are made. Further, part of wages goes for taxes, savings and other payments that does not enter into their PCE and such leaks must be removed. Ratios of wages and salaries to compensation and of PCE to personal income are calculated to do just that. All values are in current dollars and, when possible, have been adjusted to remove the affect of imports.

Calculating Average Induced Employment

This first approach use the relationships of employment, output and demand that existed in 2000 and tries to determine what portion of the PCE in that year was determined by the spending of wages. This approach takes the level of employment and output as given and allocates it to the established demand sectors. It assumes the goods purchased by wages were already available and did not require additional production and employment. This handling fits nicely with the present use of employment requirements tables. To carry out this exercise a second major input/output table, the total requirements table, is necessary. This table converts the Use Table to a measure of the output called forth in every industry to satisfy a dollar of industry demand. It measures both the direct (i.e., the auto and the glass), and the indirect (i.e., the inputs into the glass industry) effects of demand purchases. Each industry in the economy has its own column representing the direct and indirect output its purchase of inputs is responsible for. Each cell in the column is a value of industry output and so a potion of each cell is factor income. Each cell in the row represents the output of the same

industry. To change this total requirements table to one representing factor income the ratios that were previously derived are applied. Each row is scaled by that industry's compensation to output ratio giving a compensation requirements table. Each cell of this table is then scaled by the PCE to compensation ratio to create a PCE requirements table, the columns of which represent the value of PCE generated by the spending of wages paid to satisfy a dollar of that industry's demand.

Finally, this PCE requirements table must be converted from dollars to employees to calculate induced employment. Total requirements tables already are routinely converted from dollars to employment by scaling each row by that industry's employment output ratio. This is the table that is used to generate the direct and indirect industry employment per dollar of demand to measure the employment impact of spending. A variation of this procedure will be followed to convert the PCE requirements table to an PCE induced employment requirements table. A single scalar, representing the employment required per dollar of PCE is calculated. The PCE bill of goods for 2000 is multiplied by an employment requirements table to generate the employment necessary to satisfy that demand. This employment is divide by total PCE to get an employee per dollar of PCE that scales every cell of the PCE requirements table converting it to employment. The information contained in this table is similar to the standard employment requirements table and can be used in the same fashion. Column sums are the total of the induced employment in all industries in the production of a dollar of demand. The diagonal represents the induced employment of that industry's wages.

The results of these calculations are contained in table 1. Column 1 is the usual industry sum of the 2000 employment requirements table and is the direct and indirect employment generated in the production process by one million dollars of demand from that industry. Column 2 is the induced employment generated when workers spend their wages. Interesting insights into industries can be obtained from this table. Large relative values in column 1 can be a result of either a large portion of inputs going to compensation or low wages rates. A large relative value in column 2 indicates high relative wages. For example, industry 1, agricultural production, has a moderate value in column 1 and a low value in column 2 indicating a moderate portion of low wage employees. On the other hand, industry 9, oil and gas field services, has a low relative value in column 1 and a high value in column 2 indicating high wage

employees. To get the total employment impact of demand spending by industry both columns 1 and 2 should be added together giving the employment generated both in production and induced. The level of induced may strike some as being low since PCE is two thirds of GDP, however, the portion of PCE that induced spending represents is much less after imports, benefits, taxes and other payments are removed.

Calculating Marginal Induced Employment.

Data calculated in the previous exercise can be used to expand the normal Use Table to make it more of a dynamic analysis. Adding a row to represent PCE spending by industry and a column representing the goods purchased by PCE allows the measurement of output responses as the economy produces for induced spending. The previous analysis was static in the sense it just allocated the actual 2000 output and employment. This analysis assumes the introduction of a change and measures its effects. Incorporating induced PCE spending as a material input puts it on the same footing as other inputs in that it calls forth production from the "PCE" industry column which in turn calls for direct and indirect production throughout the economy. Production then results in additional wages and additional induced purchases. This additional row is the culmination of the values obtained by multiplying the ratio of industry compensation to value added (1992) times the ratio of industry value added to output (2000). This row is then scaled by the single value of PCE per industry derived from the wage to compensation ratio (2000) multiplied by the PCE to PI (2000) ratio. The "PCE" column is a dollar distribution of PCE final demand equal in total to the "PCE" row total. The PCE final demand is lowered, and so to is total final demand, by the portion thus shifted into the table. This expanded Use Table is then inverted creating an expanded total requirements table. This table now calculates the added output throughout the economy that is needed to satisfy induced spending. An employment requirements table then derived by scaling each of the original rows by that industry's employment output ratio and the new row by the previously calculated employment per dollar of PCE ratio. Column 3 in table 1 represents the employment per million dollars of demand of that industry. These values are higher than those in the first two columns for two reasons. The first is due to the added production and employment called forth by the induced spending. The second is because two thirds of PCE is no longer considered demand and removed from is the divisor but that PCE did not disappear since it is now within the table.

Which is better? Neither seems perfect, the calculation of average excludes any matching production that may be necessary and so is probably too low. On the other hand, the marginal approach generates 50% more jobs in 2000 than actual, table 2, which means that the full effects of introducing added demand into a table to include induced spending continues for more than one year. The real values are between the two so depending on the use, it is up to the analyst to decide.

Cyclical Insights

Because the expanded employment inverse is dynamic and measures marginal responses it can offer insights into the present cyclical condition. Discussions revolve around the possible speed and strength with which the economy will recover from the present slowdown. The economic bedrock for the past 3 years has been the consumer whose spending is expected to be the major determinant of the path into the future. As this study has shown total PCE as a demand component may represents two thirds of GDP but over 60% of it is really dependent on wages generated by the other demand components. Table 2 compares the employment generated by each sector of demand in 2000. Column 1 is the typical calculation of demand generated employment without the effect on PCE of induced employment. Column 2 is the employment effect with induced PCE removed from final demand and distributed to the demand sectors that actually generate it. The impact of imports has been removed from both. Comparing the two columns shows the importance to employment of the portion of PCE that is dependent on wages and where production will have to increase if PCE is to do so also. Column 1 shows PCE being responsible for over half of employment and 4 to 10 times more important than the other demand sectors. Taking induced employment into account lowers the importance of PCE while raising that of the others. A consumer with only 30% of PCE dependent on the spending of non-wage income may already be tapped out and not have the where with all on his own to keep the economy expanding. State and local government spending is tied into tax receipts that are stagnant or declining. Further, if the housing market cools and a high dollar and foreign recessions keep exports from reviving the sources of more consumer spending become even scarcer. All of which stresses the necessity for investment to contribute to growth.

Further Extensions to the Calculations

Modifications to the total requirements tables that have previously been described can be used to give insights into the specific impact each demand component has on wages and salaries and PCE. Specifically, each component's bill of goods can be used to generate the amount of wages and the amount of PCE that bill of goods will induce. This will more completely show the true affect that demand categories have on total GDP growth.

Presently the National Accounts allocates 68% of demand to PCE but obviously some portion of that is a result of of spending of wages by employees hired to fulfill the other demand components. Table 3, column 1, is the result of applying each bill of goods to the Wages and salaries requirements table earlier created. This table emphasizes the contribution of each specific demand component to total wages and can be used to get some idea of the increase that will flow through to wages of increases in specific demands. PCE as a demand component still generates 55% of wages from the spending of non-wage income sources such as transfer payments, interest receipts, etc. Column 2 is the PCE generated by each demand component after its multiplication times the PCE requirements table previously created. This shows the dollar amount of PCE that is a result of the spending of wages received in the satisfaction of each demand component. By these calculations of the $6,708 billion of PCE in 2000 the spending of wages was responsible for $3,970 billion or 59%.

Table 3 converts this spending from dollars to employment with the application of an employment to dollar of PCE ratio. Column 1 is the employment that is generated with the present allocation of all PCE and no accounting for the induced spending on the part of the other demand components. Column 2 reallocates PCE to the other demand components. Here it is seen that induced spending by the other components is responsible for 24 million of the 53 million jobs that induced spending generates. Column 3 is the redistribution of jobs to the demand categories that produce them emphasizing that 53% of jobs are dependent on demand components other than PCE.

Table 1.
Jobs per million dollars of demand: 2000
Actual, induced and multiplier
(current dollars)
number

	Jobs from Production (number)	Induced Jobs (number)	Multiplier Jobs (number)
1. Agricultural production	15.9	2.9	26.1
2. Veterinary services	18.0	5.3	36.6
3. Landscape and horticultural services	30.5	6.2	52.3
4. Agricultural services, n.e.c.	24.8	5.2	43.2
5. Forestry, fishing, hunting, and trapping	15.0	4.0	29.1
6. Metal mining	8.0	5.0	25.7
7. Coal mining	7.1	4.7	23.6
8. Crude petroleum, natural gas, and gas liquids	3.0	2.2	10.7
9. Oil and gas field services	7.9	7.9	35.7
10. Nonmetallic minerals, except fuels	9.4	4.5	25.4
11. Construction	14.4	5.1	32.4
12. Logging	14.2	3.1	25.3
13. Sawmills and planing mills	12.9	4.4	28.5
14. Millwork, plywood, and structural members	15.8	5.0	33.3
15. Wood containers and misc. wood products	13.9	4.6	29.9
16. Wood buildings and mobile homes	15.9	5.0	33.6
17. Household furniture	14.7	5.0	32.2
18. Partitions and fixtures	11.1	5.3	29.6
19. Office and misc furniture and fixtures	10.1	4.8	27.1
20. Glass and glass products	10.1	4.8	27.1
21. Hydraulic cement	6.6	4.1	21.2
22. Stone, clay, and misc mineral products	10.4	4.9	27.6
23. Concrete, gypsum, and plaster products	10.3	5.0	28.0
24. Blast furnaces and basic steel products	8.3	4.8	25.3
25. Iron and steel foundries	10.5	6.0	31.5
26. Primary nonferrous smelting and refining	7.9	4.0	21.9
27. All other primary metals	8.8	3.7	21.8
28. Nonferrous rolling and drawing	8.1	4.0	22.1
29. Nonferrous foundries	10.3	5.4	29.2
30. Metal cans and shipping containers	9.1	4.5	25.1
31. Cutlery, handtools, and hardware	10.6	5.2	28.8
32. Plumbing and nonelectric heating equipment	10.6	5.1	28.5
33. Fabricated structural metal products	11.2	5.3	29.7
34. Screw machine products, bolts, rivets, etc.	9.5	5.8	30.0
35. Metal forgings and stampings	10.2	6.0	31.3
36. Metal coating, engraving, and allied services	12.7	5.3	31.6
37. Ordnance and ammunition	9.9	5.7	29.9
38. Miscellaneous fabricated metal products	10.1	5.2	28.5
39. Engines and turbines	8.5	5.3	27.2
40. Farm and garden machinery	8.9	4.9	26.2
41. Construction and related machinery	9.7	5.6	29.6

42. Metalworking machinery and equipment	12.1	6.4	34.6
43. Special industry machinery	8.5	5.7	28.6
44. General industrial machinery and equipment	10.4	5.8	30.8
45. Computer and office equipment	7.1	4.0	21.2
46. Refrigeration and service industry machinery	9.8	5.1	27.9
47. Industrial machinery nec	13.9	6.4	36.3
48. Electric distribution equipment	9.5	4.7	26.1
49. Electrical industrial apparatus	9.3	5.3	27.8
50. Household appliances	10.4	4.8	27.2
51. Electric lighting and wiring equipment	10.2	4.9	27.6
52. Household audio and video equipment	12.7	4.2	27.4
53. Communication equipment	6.5	4.6	22.5
54. Electronic components and accessories	8.0	5.0	25.5
55. Miscellaneous electrical equipment	9.3	5.0	26.7
56. Motor vehicles and equipment	9.3	4.8	26.1
57. Aerospace	8.2	5.7	28.4
58. Ship and boat building and repairing	13.4	5.6	33.2
59. Railroad equipment	8.8	5.1	26.7
60. Miscellaneous transportation equipment	10.4	4.3	25.7
61. Search and navigation equipment	7.8	5.6	27.4
62. Measuring and controlling devices	10.1	5.5	29.4
63. Medical equipment, instruments, & supplies	8.6	4.6	24.7
64. Ophthalmic goods	12.6	6.1	34.0
65. Photographic equipment and supplies	6.5	3.6	19.1
66. Watches, clocks and parts	11.1	4.2	26.1
67. Jewelry, silverware, and plated ware	12.5	4.0	26.4
68. Toys and sporting goods	12.7	5.0	30.4
69. Manufactured products, nec	14.0	5.3	32.5
70. Meat products	16.4	3.8	29.9
71. Dairy products	12.2	3.6	24.9
72. Preserved fruits and vegetables	10.2	3.8	23.7
73. Grain mill products, fats and oils	11.8	4.0	25.9
74. Bakery products	13.5	4.9	30.8
75. Sugar and confectionery products	11.0	4.0	24.9
76. Beverages	7.1	3.2	18.3
77. Miscellaneous foods and kindred products	11.3	3.8	24.8
78. Tobacco products	4.4	2.2	11.9
79. Weaving, finishing, yarn and thread mills	13.5	4.9	30.9
80. Knitting mills	15.5	5.4	34.5
81. Carpets and rugs	11.0	4.1	25.4
82. Miscellaneous textile goods	10.7	5.2	28.8
83. Apparel	16.2	5.5	35.5
84. Miscellaneous fabricated textile products	15.4	5.3	34.0
85. Pulp, paper, and paperboard mills	8.5	4.3	23.7
86. Paperboard containers and boxes	9.7	5.1	27.7
87. Converted paper products except containers	8.6	4.3	23.6
88. Newspapers	12.3	5.2	30.4
89. Periodicals	8.2	4.3	23.2
90. Books	8.9	4.2	23.9
91. Miscellaneous publishing	8.5	3.6	21.1
92. Commercial printing and business forms	12.3	5.4	31.4
93. Greeting cards	8.6	3.3	20.3

94. Blankbooks and bookbinding	12.9	5.4	31.9
95. Service industries for the printing trade	12.0	6.2	33.7
96. Industrial chemicals	6.6	4.0	20.8
97. Plastics materials and synthetics	7.6	4.4	23.2
98. Drugs	6.8	3.8	20.3
99. Soap, cleaners, and toilet goods	7.0	3.6	19.8
100. Paints and allied products	6.7	4.2	21.4
101. Agricultural chemicals	7.5	3.9	21.0
102. Miscellaneous chemical products	8.0	4.5	23.9
103. Petroleum refining	3.4	2.2	11.2
104. Miscellaneous petroleum and coal products	5.8	3.5	17.9
105. Tires and inner tubes	9.9	5.0	27.7
106. Rubber products, plastic hose and footwear	11.6	5.2	30.1
107. Miscellaneous plastics products, nec	10.6	4.9	27.8
108. Footwear except rubber and plastic	16.6	4.3	31.6
109. Luggage, handbags, and leather products, nec	11.7	4.3	26.8
110. Railroad transportation	10.1	5.7	30.0
111. Local and interurban passenger transit	25.8	8.6	56.1
112. Trucking and courier services, except air	13.6	5.3	32.1
113. Warehousing and storage	16.9	4.8	33.8
114. Water transportation	10.0	4.4	25.4
115. Air transportation	13.8	4.7	30.5
116. Pipelines, except natural gas	4.7	2.3	12.7
117. Passenger transportation arrangement	17.9	5.4	36.9
118. Miscellaneous transportation services	18.0	5.5	37.4
119. Telephone and telegraph communications	6.7	3.8	20.1
120. Cable and pay television services	10.1	4.1	24.7
121. Radio and television broadcasting	12.8	5.7	32.8
122. Electric utilities	4.9	3.4	16.7
123. Gas utilities	6.1	3.0	16.6
124. Combined utilities	0.0	0.0	0.0
125. Water and sanitation	11.8	4.7	28.5
126. Wholesale trade	10.9	4.9	28.1
127. Retail trade, exc eating and drinking places	20.5	5.0	38.1
128. Eating and drinking places	29.1	5.6	48.8
129. Depository institutions	8.1	4.4	23.6
130. Nondepository institutions and finance, nec	13.0	7.5	39.2
131. Security and commodity brokers	7.0	6.3	29.3
132. Insurance carriers	13.7	7.1	38.5
133. Insurance agents, brokers, and service	15.2	5.2	33.3
134. Real estate	5.8	2.2	13.7
135. Royalties	0.0	0.0	0.0
136. Owner-occupied dwellings	1.5	0.6	3.5
137. Hotels and other lodging places	19.8	5.5	39.2
138. Other lodging places	20.3	5.3	38.9
139. Laundry, cleaning, and shoe repair	24.3	5.1	42.3
140. Personal services, nec	15.2	4.2	30.1
141. Beauty and barber shops	39.1	5.1	57.1
142. Funeral service and crematories	12.9	5.0	30.4
143. Advertising	11.6	5.0	29.3
144. Services to buildings	24.4	5.1	42.3
145. Miscellaneous equipment rental and leasing	10.9	4.0	25.1

146. Personnel supply services	35.7	8.0	63.9
147. Computer and data processing services	11.1	5.5	30.6
148. Miscellaneous business services	15.2	4.2	30.1
149. Automotive rentals, without drivers	10.8	5.5	30.1
150. Automobile parking, repair, and services	17.2	4.5	33.2
151. Electrical repair shops	11.6	4.9	28.9
152. Watch, jewelry, and furniture repair	19.2	4.3	34.4
153. Misc repair shops and related services	12.7	4.5	28.6
154. Motion pictures	14.5	5.4	33.5
155. Video tape rental	24.6	3.8	38.0
156. Producers, orchestras, and entertainers	17.5	5.2	35.9
157. Bowling centers	25.9	4.7	42.3
158. Commercial sports	13.5	5.6	33.0
159. Amusement and recreation services, nec	18.4	4.9	35.7
160. Offices of health practitioners	13.0	6.2	34.7
161. Nursing and personal care facilities	25.5	6.7	49.1
162. Hospitals	16.8	6.2	38.6
163. Health services, nec	17.1	5.6	36.7
164. Legal services	10.4	5.5	29.6
165. Educational Services	23.2	6.8	47.2
166. Individual and misc. social services	25.2	6.2	47.0
167. Job training and related services	29.4	6.9	53.7
168. Child day care services	48.0	5.5	67.5
169. Residential care	35.3	6.6	58.6
170. Museums, botanical and zoological gardens	25.6	6.7	49.1
171. Membership organizations	26.8	6.7	50.4
172. Engineering and architectural services	12.5	6.3	34.8
173. Research and testing services	13.5	6.3	35.8
174. Management and public relations	12.9	5.6	32.5
175. Accounting, auditing, and other services	14.6	5.8	35.0
176. Private households	53.4	9.1	85.6
177. US Postal Service	14.7	8.9	46.0
178. Federal electric utilities	0.0	0.0	0.0
179. Federal government enterprises, nec	10.6	4.4	26.0
180. Federal general government	7.8	9.1	39.9
181. Federal government capital services	0.0	0.0	0.0
182. Local government passenger transit	0.0	0.0	0.0
183. State and local electric utilities	0.0	0.0	0.0
184. State and local government enterprises, nec	10.9	4.6	27.3
185. State and local government hospitals	18.9	9.1	51.1
186. State and local government education	27.6	9.1	59.7
187. State and local general government, nec	25.5	9.1	57.7
188. State and local government capital services	0.0	0.0	0.0
189. Noncomparable imports	0.0	0.0	0.0
190. Scrap, used and secondhand goods	0.0	0.0	0.0
191. Rest of the world industry	0.0	0.0	0.0
192. Inventory Valuation Adjustment	0.0	0.0	0.0

Table 2.
Jobs Generated by Each Demand Component: 2000
Actual and multiplier
(current dollars)

	Actual Jobs 2000 (thousands)	Multiplier Jobs 2000 (thousands)	Jobs/mill $ Actual Demand (persons)	Jobs/mill $ Multiplier Demand (persons)
Total	143,734	203,827	13.6	32.7
PCE	86,290	57,988	13.9	29.3
Equipment and software	9,126	25,333	10.5	29.1
Nonresidential structures	4,448	10,611	13.7	32.7
Residential structures	5,455	12,254	13.4	30.1
Inventory change	684	1,784	15.8	41.2
Exports	8,339	22,661	8.4	22.7
Defense	2,635	9,878	7.4	33.3
Nondefense	2,071	7,123	9.7	35.4
State and local education	11,513	25,368	22.9	50.4
Other state and local government	12,577	29,290	17.7	41.3

Table 3.
Generated Wages and Generated PCE: 2000
(billions of current dollars)

	Wages and Salaries	Personal Consumption Expenditures
Total	4,837	3,970
PCE	2,677	2,198
Equipment and software	421	345
Nonresidential structures	160	131
Residential structures	176	145
Inventory change	29	23
Exports	371	305
Defense	188	154
Nondefense	131	108
State and local education	317	261
Other state and local government	266	300

Table 4.
Jobs Generated by Each Demand Component
Before and After the Reallocation of Induced Employment: 2000
(millions of jobs)

	Induced Employment Unallocated	Induced Employment	Induced Employment Reallocated
Total	143.7	53	143.7
PCE	86.3	29	62.3
Equipment and software	9.1	5	14.1
Nonresidential structures	4.4	2	6.4
Residential structures	5.5	2	7.5
Inventory change	.7		.7
Exports	8.3	4	12.7
Defense	2.6	2	4.8
Nondefense	2.1	1	3.1
State and local education	11.5	4	15.5
Other state and local government	12.6	4	16.6

This page intentionally left blank.

ECONOMIC IMPLICATIONS OF FUTURE YEARS DEFENSE PURCHASES
AN ILLUSTRATION OF DEPPS

Douglas S. Meade, Inforum and Ron Lile, Department of Defense, OSD PA&E[*]

1. Introduction

Despite the downsizing of the Defense Department that started in the late 1980s, defense spending still comprises 3.2 percent of GDP, and this share may be expected to grow, according to the recently released fiscal year 2003 defense budget. Both inside and outside the Pentagon, defense policy analysts, businessmen and economists are interested in the economic implications of these defense purchases. Since the distribution of spending among industries and states is by no means uniform, and since many of the economic effects are indirect, an analytical tool is needed to determine these spending implications. Also of interest is the effect of defense spending on the demand for skilled and professional labor.

The Defense Employment and Purchases Projections System (*DEPPS*) was designed to help analysts understand how industries, states and occupational groups are affected by changes in the defense budget. *DEPPS* consists of three major components: an interindustry model, a state model and an occupational model. The interindustry model (*IDEPPS*) consists of the Inforum detailed interindustry model *Iliad*, joined with the defense translator, a matrix that translates outlays on detailed defense budget programs to the industries that directly supply these programs. The state model (*RDEPPS*) distributes defense spending by industry to the state level, based on state shares derived from historical data. The occupational model (*LDEPPS*) translates defense related employment by industry to the occupational level.

The *DEPPS* projections are made for calendar year outlay estimates derived from the Future Year Defense Purchases (FYDP), as published in *National Defense Budget Estimates*. The projections are also informed by recent historical industry and state spending patterns derived from various published and unpublished sources.

In this paper, we'll take a tour through the highlights of the *DEPPS* projections. The sample tables in this paper are from the fiscal year 2001 projections. (There were no projections in 2002, as there was no FYDP released, and the 2003 projections are now in progress.) We'll discuss each of the main parts of *DEPPS* in turn, and then conclude with some general observations. Along the way, we'll try to provide some insight into the calculations that lie behind the projections.

2. *IDEPPS,* The Interindustry Component of *DEPPS*

The purpose of *IDEPPS* is to determine defense-related production needed to supply the bill of goods and services specified in the FYDP. Defense-related production includes both direct purchases by DoD, such as an Abrams tank or a Commanche helicopter. It also includes indirect purchases, such as the semiconductors used to make the electronic systems in tanks, helicopters, ships and aircraft. Using this information, one can easily see if the planned defense budget contributes to growth or decline in a given industry. One can also see the projected share of total output comprised of defense-related production.

The *IDEPPS* projections can be summarized as follows:

- They are produced at a level of 320 industries, the same used for the detailed Inforum model of the U.S. economy.

- They are made in constant (inflation-adjusted) dollars, by calendar year, for the interval defined by the FYDP.

- They reflect planned expenditures or outlays, not appropriations or budget authority.

- They reflect DoD expenditures for military programs only.

- They exclude expenditures for pay.

For each of the 320 industries that supply directly or indirectly to defense, several tables of information can be compiled from the projections. We'll look at three sample tables for the Electronic components industry.

---

[*] Douglas S. Meade, Inforum, c/o Department of Economics, University of Maryland, College Park, MD 20742, (301) 405-4607. Ron Lile, Office of the Secretary of Defense, Program Analysis & Evaluation, 1800 Defense Pentagon, Washington, DC 20301-1800, (703) 614-3840.

Table 1. Projected Defense Purchases of Electronic Components, 2000-2005
(In Millions of 2001 dollars)

	2000	2001	2002	2003	2004	2005
Summary of Defense Purchases						
Direct	1,837	1,883	1,884	1,940	1,944	1,949
Indirect	1,653	1,676	1,705	1,760	1,795	1,809
Total	**3,489**	**3,559**	**3,589**	**3,700**	**3,740**	**3,758**
Indirect Defense Purchases by Purchasing Sector						
Missiles	54	57	60	64	62	62
Ammunition	32	33	35	36	36	36
Tanks and Tank Components	0	0	0	0	0	0
Other Ordnance	1	1	1	1	1	1
Communications Equipment	638	652	665	689	706	707
Other Electronic Equipment	269	273	272	276	277	276
Motor Vehicles	3	3	4	4	4	4
Aircraft and Parts	86	90	94	102	108	110
Aircraft Engines and Parts	7	7	7	7	7	7
Shipbuilding	4	4	4	5	5	5
All Other	558	555	563	577	588	599
Total	**1,653**	**1,676**	**1,705**	**1,760**	**1,795**	**1,809**

Table 2. Sources of Projected Direct Plus Indirect Defense Purchases of Electronic Components, 2000-2005
(In Millions of 2001 dollars)

	2000	2001	2002	2003	2004	2005	00-05
Military Personnel	6	6	6	6	6	7	4.14
Operations & Maintenance + Revolving Funds	1,354	1,367	1,344	1,360	1,367	1,368	0.21
Procurement	1,489	1,547	1,614	1,727	1,774	1,801	3.80
Aircraft	278	291	309	332	345	359	5.08
Missiles	381	379	409	418	419	411	1.54
Weapons and Tracked Vehicles	3	2	2	3	3	3	1.61
Ships and Conversions	21	17	17	19	17	13	-9.17
Ammunition	1	1	1	1	1	1	1.92
Other	805	856	875	953	988	1,014	4.60
RDT&E	579	581	572	559	545	530	-1.76
Military Construction	46	44	39	34	33	36	-4.67
Family Housing	15	14	14	14	14	14	-0.45
Total	3,489	3,559	3,589	3,700	3,740	3,758	1.48

Table 3. Projected Domestic Production, Defense Purchases, and Imports for Defense Production of Electronic Components, 2000-2005
(In Millions of 2001 dollars, except as noted)

	2000	2001	2002	2003	2004	2005	00-05
Total U.S. Domestic Production	39,384	42,039	44,201	46,673	49,289	52,160	5.62
Plus Imports	6,158	6,467	6,670	6,930	7,212	7,554	4.09
Less Exports	15,354	16,776	18,111	19,518	21,011	22,550	7.69
= Domestic Use	30,188	31,730	32,760	34,085	35,490	37,163	4.16
Import Share of Domestic Use (percent)	20.4	20.4	20.4	20.3	20.3	20.3	-0.07
Defense Purchases	3,489	3,559	3,589	3,700	3,740	3,758	1.48
Less Imports	337	341	347	358	365	368	1.75
Domestic Defense Purchases	3,152	3,218	3,242	3,342	3,375	3,390	1.46
Domestic Defense Purchases as a Share of Domestic Production (percent)	8.0	7.7	7.3	7.2	6.8	6.5	-4.16

In each of these tables, projections are shown for each year of the FYDP. In several tables, the last column shows the average annual growth rate.

Table 1 shows how total defense-related purchases are divided between direct and indirect purchases. For the indirect purchases, it also indicates from which major direct purchasing sector they are derived. For example, Table 1 indicates that in 2000, an estimated 638 million *indirect* expenditures for Electronic components was needed to supply the *direct* expenditure of Communications equipment to DoD. Also note that Electronic components is an industry for which a large share of defense purchases are indirect. In 2001, DoD was estimated to spend about $1,883 million directly, and $1,676 indirectly.

Table 2 shows the origins of defense-related demand for Electronic components from the major headings of the DoD budget. This table can help to understand how the demand for an industry will shift as purchases are reallocated from one major budget category to another. From this table we can see that operations and maintenance, procurement and RDT&E (Research, Development, Test and Evaluation) comprise almost all of the defense-related demand for this industry. Within the procurement budget, the largest sources of demand are aircraft, missiles and other procurement.

Table 3 is useful for comparing trends in defense and nondefense purchases. Shown in the first block of items in the table are projections made by Inforum of economy-wide domestic production, net imports (imports less exports) and domestic use. (Domestic use is the sum of domestic production and net imports). Also shown is the projected share of domestic use supplied by imports.

The middle part of the table presents projections (for comparison to the estimates of total domestic production) of defense purchases from domestic suppliers. *"Domestic* defense purchases" are defined as *total* defense purchases less *imports* for defense production. In the example used, total projected defense purchases of electronic components amount to about $3,559 million, with $341 million supplied by imports. Defense purchases from domestic producers were therefore calculated to total about $3,218 million in 2001.

Shown at the bottom of the table are estimates of the share of total domestic production accounted for by defense purchases. Again using electronic components as an example, defense purchases are projected to account for about 7.7 percent of the industry's output in 2001 and for 6.5 percent in 2005.

How the IDEPPS Projections are Made

Figure 1 summarizes how the *IDEPPS* projections are computed. The Future Years Defense Plan or FYDP is the starting point. This is essentially the defense part of the published federal budget, except the nonmilitary functions. Projected outlays by major program in constant prices are made available on a fiscal year basis in the publication *National Defense Budget Estimates*.

The next step in *IDEPPS* starts with the constant price outlays and converts these to implied direct purchases from each of 320 industries, using what is called the "defense translator". The translator is a matrix that embodies information on many detailed defense programs. Any particular program may purchase inputs from a dozen or more industries. Table 4 illustrates how the translator for one of the budget accounts listed earlier--Aircraft procurement--would allocate outlays, in the year 2001, among various industries. Note that, in this example, about 77 percent of the outlays go to the three aircraft-related industries.

Figure 1. IDEPPS Projection Flow
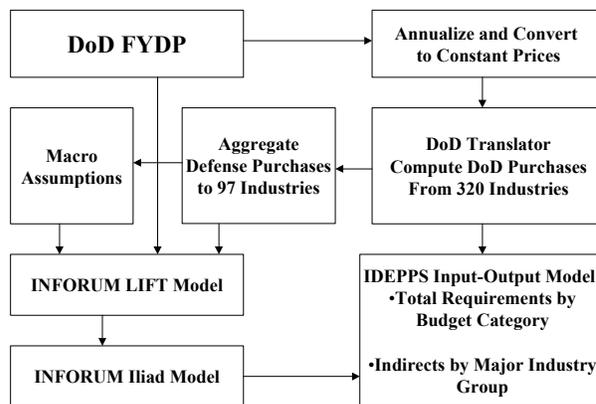
The translators for the major accounts allow the computation, from the budget data described above, of direct defense purchases from each of the 320 industries in the system. These projections are computed in constant dollars for the upcoming budget year.

The *IDEPPS* projections of total defense purchases are made using the 320-sector interindustry Inforum model. This model is used to calculate the indirect requirements of the expenditures indicated by the translator, as well as determine what proportion of total requirements in each industry is satisfied by imports.

The interindustry model is used several times in *IDEPPS*, for the direct DoD purchases associated with:

• the DoD budget as a whole;

• each of the major aggregate DoD budget accounts; and

Table 4. Estimated Distribution Among Industries of Outlays from the Aircraft Procurement Account, 2001

	2001 Value	Share (%)
22 Ammunition, except small arms	1491.6	8.3
220 Communication equipment	279.2	1.5
235 Aircraft	7332.8	40.6
236 Aircraft and missile engines	706.0	3.9
237 Aircraft and missile parts	5804.5	32.1
238 Ship building and repairing	439.1	2.4
246 Search and navigation equipment	633.9	3.5
290 Research laboratories and management consulting	288.5	1.6
295 Engineering and architectural services	759.9	4.2
296 Other professional services, including accounting	341.9	1.9
Total	**18077.6**	**100.0**

Table 5. New Mexico Summary

	2000	2001	2002	2003	2004	2005	00-05
AGGREGATE MEASURES							
Total Direct Expenditures (Purchases and Pay)	2,431	2,394	2,368	2,365	2,375	2,391	-0.33
Indirect Defense Purchases Resulting from Direct Purchases	1,013	975	996	1,013	1,034	1,062	0.95
Indirect Defense Purchases Resulting from Pay	603	592	585	584	586	592	-0.37
Total Nondefense Expenditures	87,090	89,819	91,204	93,416	95,977	98,626	2.49
Total Output	91,138	93,780	95,154	97,378	99,971	102,672	2.38
Government Industry Compensation	1,390	1,372	1,351	1,343	1,344	1,352	-0.55
LARGEST PURCHASES BY INDUSTRIAL SECTORS							
Total Direct Expenditures (Purchases and Pay)							
Research labs, and other professional services	458	458	450	445	438	430	-1.28
New construction	115	117	113	113	113	114	-0.22
Air transport	95	94	95	102	112	117	4.13
Trucking, highway passenger transit	57	58	60	61	63	65	2.79
Gas utilities	38	35	34	34	34	34	-1.92
Indirect Defense Purchases Resulting from Direct Purchases							
Research labs, and other professional services	148	147	148	150	151	153	0.62
Gas utilities	105	102	99	99	96	95	-2.09
Other business services	104	104	106	109	112	114	1.96
Real estate and royalties	74	72	71	72	76	79	1.14
Crude petroleum	74	55	61	59	61	67	-2.04

Table 6. Top 10 States in Direct Purchases of Communication Equipment (Millions of 2001 Dollars)

	2000	2001	2002	2003	2004	2005	00-05
Florida	1,141	1,156	1,154	1,154	1,155	1,138	-0.05
California	878	901	913	938	963	963	1.85
Massachusetts	706	728	743	767	792	792	2.30
Washington	517	540	565	593	621	624	3.77
Indiana	518	534	544	562	581	583	2.34
Maryland	395	404	407	418	429	429	1.67
Texas	381	392	399	413	426	427	2.28
New York	317	325	328	336	345	344	1.64
Virginia	274	277	274	276	278	275	0.12
Iowa	197	205	212	221	230	231	3.14
Top 10 Total	5,324	5,460	5,540	5,677	5,817	5,806	1.73
Total U.S.	6,315	6,471	6,553	6,709	6,868	6,854	1.64

• each of 11 aggregate industrial sectors.

The first application of the table yields projections of total, direct, and indirect defense purchases (Table 1). Indirect defense purchases are calculated by subtracting direct defense purchases from total domestic requirements from defense. The remaining applications disaggregate defense purchases by budget category (Table 2).

The *IDEPPS* reports include projections, made by Inforum, of total domestic production (Table 3). The projections are derived from two types of data: 1) the DoD budget data used in *IDEPPS* and 2) other assumptions underlying the projections are those made by Inforum in its published baseline forecasts. The projections reported for domestic defense purchases (also shown in Table 3) are calculated by subtracting estimates of imports used to produce defense purchases from total defense purchases. The import share of total apparent consumption for each year in the forecast period is computed from Inforum projections of imports and consumption.

3. *RDEPPS*, The State-Level Projections Component of *DEPPS*

IDEPPS addresses the question: "What industries produce defense goods and services?" *RDEPPS* addresses the question: "Where will defense-related production occur?" In this component of *DEPPS*, the geographical distribution of the industry level purchases from *IDEPPS* is determined. Due to limitations in the available data, these projections are made at a level of only 97 industries, which corresponds to the sectoring of the Inforum *LIFT* model. The projections are made for each of the 50 states and the District of Columbia. Unlike *IDEPPS*, *RDEPPS* also determines spending impacts of active duty and military retirement pay that is spent in the U.S.

The expenditure projections are presented in two formats; one designed to show the level and composition of potential expenditures in individual states, and the other to show the geographic distribution of purchases from given industrial sectors.

Expenditure Tables. Table 5 illustrates the format of the state-expenditure projections, using the forecast for New Mexico as an example. The first block of the table shows aggregate measures, in dollar value, of projected direct and indirect defense expenditures in the state during each of the projection years. A projection of nondefense economic activity and total output, prepared by Inforum is also provided. The second and third blocks of the table show the industrial sectors projected

to lead in defense or defense-related sales over the projection period.

Starting at the top of the table "Direct Defense Expenditures" ($) are the monies disbursed by DoD to pay for purchases of goods and services and to cover payroll expenses. Purchases of magnetic recording tape by the Defense Logistics Agency and the wages of military and civilian personnel at Kirtland Air Force Base are two examples of such expenditures. Direct purchases, in turn, trigger subsequent rounds of transactions, referred to collectively as "indirect defense expenditures from direct purchases." These expenditures represent purchases by DoD's prime contractors (and their suppliers) of parts and materials used in producing items ordered by DoD. Fuel bought by a trucking company for transporting a shipment of goods to DoD would be examples of this type of expenditure. "Indirect purchases resulting from pay" ($592 million in 2001) represent purchases by DoD's military and civilian employees of goods and services for their personal use. The purchases of a clock radio by a DoD employee would be an example of this category of expenditure. The personal consumption expenditures of military and civilian employees may be taken as a measure of the indirect effects of the pay portion of the DoD budget. These are included in *RDEPPS* (but not other parts of *DEPPS*) because they are often a focus of attention in local development efforts.

In 2001, some $2,394 million in direct expenditures was projected to be disbursed by the Defense Department in New Mexico to pay its employees and reimburse its direct suppliers for the goods and services they provide. Pay to military members and civilian government workers accounts for a large share of DoD's expenditures in the state ($1,372 million). Indirect purchases of $975 million dollars are projected to result from DoD Purchases and from purchases made by DoD employees. In terms of defense-related expenditures, sales of research labs and other professional services absorb the largest share of indirect defense dollars ($147 million).

Tables 6 and 7 illustrate the format of the industry projections, using estimated purchases from the Communications equipment sector as an example. Two tables are provided for each of 97 industrial sectors, the first showing the top 10 states in which the sector is projected to make the bulk of its direct defense sales over the forecast period and the second showing the top 10 states in which indirect defense sales resulting from direct purchases are projected to be concentrated. Altogether, the 10 states represented in table 6 are estimated to account for 84 percent of the total direct purchases of Communications equipment. The top 10

states in table 7 comprise 69 percent of total indirect spending.

How the RDEPPS Projections are Made

The state level estimates cover expenditures originating from the following aggregate accounts of the defense budget: (1) military personnel; (2) procurement; (3) research, development, test, and evaluation (RDT&E); (4) operations and maintenance (O&M); and (5) military construction and family housing. For each of these accounts, total defense-related expenditures can be classified into three categories:

- Pay projections, both for active-duty and retired military personnel, and for the DoD civilian workforce;

- Projected direct defense purchases derived from IDEPPS; and

- Projected indirect defense purchases derived from IDEPPS.

DoD Pay. Historically, the distribution of DoD pay among states has differed significantly from the distribution of direct purchases. Consequently, in estimating future levels of defense expenditures, it is useful to treat pay and purchases separately. This requires some transformation of the budget data because pay expenditures are not grouped into a single account. With the exception of the retired pay account, which consists entirely of pay, several of the budget accounts cover both purchases and pay.

RDEPPS separates, for each budget account, non-pay outlays from pay components. The pay portions cover the wages and salaries of military and civilian DoD personnel, whether they are stationed in the United States or abroad. Because the state-level estimates consider only expenditures made in the United States, the aggregate pay data must be adjusted to remove the fraction of pay disbursed outside the country.

This adjustment is quite substantial. In 2001, about 15.8 percent of the active-duty force was stationed overseas, in U.S. territories, or aboard ships in foreign waters. An estimate of these individuals' pay is subtracted from total military pay in order to arrive at an estimate of the amount of pay going to military personnel stationed in the United States. (Though service members stationed outside the country do not necessarily receive all of their pay abroad, there is no simple way to determine what proportion is received by dependents living in the United States, or how those funds are distributed among the individual states.) Some civilian personnel are also stationed overseas or in U.S. territories, and some military retirees live abroad. Small adjustments to

civilian (6.2 percent) and retired pay (1.5 percent) are therefore made as well.

Direct Defense Purchases. Direct defense purchases for the nation as a whole are aggregated from *IDEPPS* 320 industries to the RDEPPS 97 industry sectoring. The result is projections, for each account, of domestic direct defense purchases from each of the 97 industrial sectors. After the purchases have been allocated by sector, they are distributed at the state-level on the basis of state shares of direct purchases arising from each budget account. Note that the state shares differ for each of the major accounts. Furthermore, pay is distributed using pay shares, as described below.

This procedure has the very important advantage of reflecting the effects of changes in the composition of defense purchases, but it requires very detailed information on historical state shares of direct defense expenditures. This information is derived primarily from historical data on contracts awarded.

Indirect Defense Expenditures. Indirect purchases are triggered by purchases made directly by DoD. Each indirect purchase, in turn, typically generates a series of subsequent purchases. Although indirect defense purchases constitute a sizable share of total defense spending, only fragmentary data on their geographical distribution are available. Moreover, assembling a reasonably complete data series would be a very large undertaking. Intead, such purchases were estimated using assumptions typical of regional analysis.

Industries are divided into two categories: "basic" and "non-basic". "Basic" industries are those for which the national market is considered the relevant one. All manufactured goods are assumed to fall within this category. "Non-basic" industries, on the other hand, are industries for which the state is considered the relevant market. Indirect purchases for "basic" industries are assumed to be distributed according to the same state shares as total production for that industry. This distribution is assumed to be that projected by the Inforum STEMS model, which forecasts economic activity by state. Indirect purchases for non-basic industries are assumed to be distributed according to the distribution of the direct spending which generates these purchases.

The results are estimates of the indirect defense purchases that arise from the nonpay portion of the DoD budget. The pay portion of the budget also has indirect effects, which arise from the consumption expenditures of DoD employees. Consequently, in making the state-level estimates, indirect defense purchases are defined as the sum of: (1) indirect purchases stemming from the purchases component of the DoD budget;

Table 7. Top 10 States in Indirect Purchases of Communication Equipment
(Millions of 2001 Dollars)

	2000	2001	2002	2003	2004	2005	00-05
California	112	117	122	126	128	130	2.84
Florida	73	76	81	84	85	87	3.51
Texas	71	72	74	76	77	78	1.90
Massachusetts	59	61	63	65	65	65	1.93
Illinois	56	57	58	60	60	61	1.65
New Jersey	32	34	35	37	37	38	3.34
Nevada	26	28	29	30	30	30	2.78
Virginia	26	27	29	30	30	30	3.11
Ohio	26	27	27	28	28	28	1.10
New Mexico	26	26	27	27	27	28	1.57
Top 10 Total	507	524	544	563	567	574	2.47
52 All U.S.	733	756	783	809	815	826	2.37

Table 8. Top 5 Industries Employing Aeronautical and Astronautical Engineers
Total U.S. Employment and Defense-Related Employment
Thousands of Workers, Ranked by Level in 2001

Defense-Related Employment

	2000	2001	2002	2003	2004	2005	00-05
100 Federal government, defense	9.1	8.8	8.7	8.6	8.5	8.6	-1.10
51 Aerospace	5.1	5.3	5.4	5.6	5.6	5.6	2.05
54 Search & navigation equip.	0.2	0.2	0.2	0.2	0.2	0.2	-2.96
47 Communication equipment	0.2	0.2	0.2	0.2	0.2	0.2	-1.42
77 Professional Services	0.1	0.1	0.1	0.1	0.1	0.1	-4.47

Total U.S. Employment

	2000	2001	2002	2003	2004	2005	00-05
51 Aerospace	26.8	27.9	28.9	29.7	31.1	32.6	3.92
100 Federal government, defense	9.1	8.8	8.7	8.6	8.5	8.6	-1.10
101 Federal government, nondefense	4.6	4.7	4.7	4.8	4.9	5.0	1.57
80 Other business services	2.8	2.8	2.8	2.8	2.8	2.8	-0.03
53 Other transportation equipment	2.2	2.3	2.4	2.4	2.5	2.6	2.92

Table 9. Share of Defense-related Employment by Occupation
Thousands of Workers

	Total	Defense Related	Percent Share
TOTAL EMPLOYMENT	149322	2864	1.92
74 Shipfitters	12	7	59.67
3 Aeronautical and astronautical engineers	59	16	26.71
60 Aircraft mechanics and engine specialists	174	36	20.44
64 Aircraft assemblers, precision	21	4	18.73
19 All other physical scientists	32	6	17.81
6 Electrical and electronics engineers	399	68	16.94
15 Operations research analysts	57	9	15.78
8 Mechanical engineers	230	33	14.29
5 Civil engineers, including traffic engineers	170	22	12.91
14 Mathematicians and all other mathematical scientists	35	4	10.63

and (2) consumption expenditures (indirect purchases resulting from the pay of military and civilian personnel) of defense personnel. Indirects from pay are calculated by first calculating the share of consumption out of total income, and then allocating the consumption by type of good. The shares used are the same as projected at the national level. As with the indirects from purchases, there are both basic and non-basic industries, so not all consumption goods and services are purchased in the given state.

Once indirect purchases have been estimated for each of the 97 industries, the computations proceed in much the same way as those for direct defense purchases. Since there is no basis for estimating how state shares of indirect purchases vary by budget account, indirect defense purchases from each of the 97 sectors are presented only in total.

Estimation of State Direct Purchases Shares. State shares of direct purchases are calculated on the basis of historical data showing how those expenditures have been distributed at the state level in recent years. Since adequate historical data on the distribution of indirect defense purchases are not available, a somewhat different method is used to calculate state shares of those purchases. This section describes how state shares are established for each category of expenditures, and notes the potential limitations of the methods.

State Shares of Pay. Estimated outlays for military pay are allocated among the states on the basis of their shares of total military pay in the most recent year for which this information is available. These shares are held constant over the projection period. Military retired pay and civilian pay likewise are distributed among the states on the basis of the distribution in the base period.

Because the state distributions are fixed at historical levels, increases in military or civilian pay (or in military retirement annuities) over the projection period only affect the estimated amount of pay going to each state, not each state's share relative to other states. That is, if the amount of military pay disbursed in state "x" in the base period were twice that disbursed in, state "y," the estimates for each future year would show twice as much military pay being disbursed in state "x" as in state "y."

This "fixed shares" assumption will lead to serious distortions in the estimates only if there are major changes in the number of personnel within given states (or in the distribution of personnel among pay grades) over the projection period.

4. *LDEPPS,* The Defense-Related Employment and Skilled Labor Component of *DEPPS*

This component of DEPPS tracks employment generated by DoD direct hire, and from direct and indirect purchases. It also uses the projected occupational matrix from BLS to show the employment for each of 100 occupational groups by industry.

Questions of the affect of defense purchases on the demands for labor of various occupational groups is interesting for a number of reasons. Defense-related employment is an important segment of employment for several professional and skilled occupations. This is particularly true for certain types of scientists and engineers. Forecasting demand for these occupational categories can help individuals decide whether this is a good field of study in which to invest in education. For policy makers, it is helpful to know if certain occupations may be in relatively short supply, thus leading to bottlenecks or excessive wage costs.

LDEPPS employment projections are based on projections of defense-related production combined with projected changes in labor productivity. Employment by occupation is then calculated using the occupational shares matrix. The BLS occupational shares describe, for example, what share of employment in the motor vehicle industry will be mechanical engineers. Labor productivity is the ratio of gross constant dollar output divided by total hours worked, in each industry. *LDEPPS* relies on the productivity and employment projections calculated in the Inforum *LIFT* model.

For each occupation, both total and defense-related employment are broken down among 89 industries comprising total GDP. These 89 sectors are essentially the sectors in the Inforum LIFT model which have employees, with a few special definitions such as education, hospitals, domestic servants, and government employees.

Presented in Table 8, as an example, are the *LDEPPS* projections of employment of aeronautical and astronautical engineers. The first 6 columns show the year by year projections in thousands of persons. The last column shows the average growth rate over the period, 2000 to 2005.

The top half of this table shows what is called "defense-related employment." Defense-related employment of people in an occupation is defined as the sum of:

- Employment in that occupation by DoD;

- Private sector employment in that occupation directly engaged in defense production; and,

- Private sector employment in that occupation indirectly engaged in defense production (i.e.,

engaged in production of inputs to goods bought by DoD; production of inputs to those inputs; and so on through the chain of production).

Shown in the lower half of the table is projected total employment of aeronautical and astronautical engineers. Nondefense employment (not shown separately) is the difference between total and defense-related employment in the category.

The format of the projections is the same for all of the 100 occupations included in *LDEPPS*. The "aeronautical and astronautical engineers" occupation is a convenient example because employment is concentrated in comparatively few industries. It is, however, unrepresentative in two respects. First, employment in most occupational categories is much more widely distributed among industries. Second, defense-related employment is about 27 percent of total employment of aeronautical and astronautical engineers. (This is not surprising, as DoD and defense-related purchases account for over half of the output of the domestic aerospace industries). For most occupations, including other engineering specialties, the defense-related share of total employment is much smaller.

Table 9 shows total employment and defense-related employment for the top 10 occupations, ranked by the share of defense-related employment in the total. Overall, defense-related employment makes up only 1.9 percent of total employment in 2001. However, for the occupations presented in this table, defense-related employment is a much larger share, ranging from 10 percent to almost 60 percent.

How the LDEPPS Projections Are Made

The *LDEPPS* projections are computed in two main parts:

1. projecting employment in each of 89 sectors; and

2. estimating (sector by sector) employment in each of 100 occupational categories.

The first part relies on the employment projections from the Inforum *LIFT* model. The second part relies heavily on projections and data published by the Bureau of Labor Statistics (BLS) of the U.S. Department of Labor.

Total Employment by Industry. LDEPPS takes as its point of departure *IDEPPS* projections of purchases from each of 320 SIC industries. For the base year the coefficients in the *LDEPPS* are ratios of employment to industry output. The projected values of the labor input coefficients reflect expected trends in labor productivity. (Note that employment per dollar of output is the reciprocal of average labor productivity.)

Employment by Occupation. The 100 occupational categories in *LDEPPS* are aggregations of more detailed categories established by BLS. Definitions of the occupational categories used by BLS change somewhat from one survey to the next. BLS maintains detailed definitions of the categories used.

The BLS National Industry-Occupational Matrix gives, for each industry, the shares of employment in the industry accounted for by various occupations. This matrix, which covers wage and salary workers, is prepared biennially by BLS. BLS generates projections of occupational distribution by industry by analyzing the factors expected to influence trends in the staffing patterns of industry as technologies change. Currently, the BLS projected matrix that is in *LDEPPS* is for 2008. Inforum has enhanced the BLS matrix by estimating the distribution of the self-employed and family workers, as well as filling in numerous cells not disclosed in the original table.

LDEPPS uses (for each year of the forecast horizon) the appropriate linear interpolation between the National Industry-Occupational matrix for the most recent year and the projected table for 2008. For each industry, the estimated shares of employment accounted for by the different occupational categories are multiplied by total projected employment in the industry. Projected employment for an occupation is the sum across industries of employment in the occupation in question.

DoD direct employment is, however, handled somewhat differently. The total number of civilian employees in the Department of Defense is derived from the FYDP. Distributions of employees into occupational categories are based on special tabulations developed by BLS from Office of Personnel Management reports. It should also be noted that in *LDEPPS* teachers and other educational workers employed by state and local governments are included in Sector 87 (Private and public education, and non-profit organizations) rather than in Sector 102 (State and local government). Also, state and local hospital workers are combined with private hospital workers in sector 83 (Private and public hospitals). Sector 102 excludes hospitals and education. The reason for this is that no separate occupational employment information is available for these sectors.

5. Concluding Comments

DEPPS was designed as an analytical tool to understand the economic implications of planned defense purchases. This paper has described how the three main components of DEPPS work together to produce estimates of defense-related spending by industry, by state, and employment by occupational group. DEPPS is used by DoD after the release of each FYDP to

produce a projections book entitled *Projected Defense Purchases: Detail by Industry and State*, which can be found on the DoD web site (see the References). Additional detail on DEPPS can be found in the *DEPPS* Primers at that site.

Other applications for DEPPS include: (1) comparing the implications of two or more alternative defense budgets; (2) tracking the historical contribution of defense spending to U.S. economic growth (Meade, 1998); (3) determining likely bottlenecks of increased defense spending in time of conflict (Meade, 1999); (4) deriving alternative measures of defense deflators (Meade and Lile, 2001); and (5) assisting state governments to determine the impacts of defense budgets on state economies.

References

Hecker, Daniel, "Occupational Employment Projections to 2010," *Monthly Labor Review*, November 2001, pp 57-84, http://www.bls.gov/opub/mlr/2001/11/art4full.pdf.

Institute for Defense Analyses, *The Revised Defense Translator*, October 1989.

Meade, Douglas S., "Defense Spending in the Context of the U.S. Economy," Inforum Working Paper WP98001, http://inforumweb.umd.edu/WorkPaper/INFORUM/wp98001.pdf, April, 1998.

Meade, Douglas S., "Constraints to U.S. Defense Spending," Inforum Working Paper WP99002, http://inforumweb.umd.edu/WorkPaper/INFORUM/Constraints.pdf, June, 1999.

Meade, Douglas S. and Ronald Lile, "An Analysis of Historical and Projected Cost Deflators for Budget Components," Inforum Working Paper WP01003, http://inforumweb.umd.edu/WorkPaper/INFORUM/wp01003.pdf, July, 2001.

U.S. Department of Commerce, Bureau of Economic Analysis, *Benchmark Input-Output Accounts of the United States, 1992*, http://www.bea.doc.gov/bea/articles/NATIONAL/Inputout/Meth/io1992.pdf.

U.S. Department of Defense, *National Defense Budget Estimates for FY 2003,* http://www.dtic.mil/comptroller/fy2003budget/FY03GBpdf.pdf, March, 2002.

U.S. Department of Defense, *Projected Defense Purchases: Detail by Industry and State,* http://www.ra.pae.osd.mil/depps/files/DEPPS2000.pdf, July, 2000.

U.S. Department of Defense, Supporting Documentation for DEPPS, http://www.ra.pae.osd.mil/depps/support.html, October, 2000.

CONCURRENT SESSIONS II

This page intentionally left blank.

Evaluating Projections and Dealing with NAICS

Chair: Charles Bowman, Bureau of Labor Statistics, U.S. Department of Labor

Evaluating the BLS Labor Force Projections to 2000

Howard Fullerton, Bureau of Labor Statistics, U.S. Department of Labor

Beginning from 1986 and continuing through 1996, BLS prepared five labor force projections for 2000. The overall errors were greatest for the 1986 and 1990 projections (1.5 percent low or high); except for these two projections, the errors were less than 1 percent. For those users for whom the error in the annual growth rate is most important, the error in the annual growth rate from 1988 was 0.02 percent. For four other projections, the error in the annual growth rate was either -0.1 or 0.1 percentage points.

An Evaluation of the 2000 Industry Employment Projections

Arthur Andreassen, Bureau of Labor Statistics, U.S. Department of Labor

The Bureau of Labor Statistics carries out a biennial program of medium-term projections of all phases of national-level economic growth. This paper compares actual 2000 industry employment with the BLS projections of industry employment published in 1989. Emphasis is placed on how well the projections process performed at an industry level and the analysis examines the sources of error. An attempt is made to classify error due to process error versus error due to incorrect assumptions, thus allowing the projections program to evolve and, hopefully, to improve over time.

Evaluating the 2000 Occupational Employment Projections

Jill Auyer and Andrew Alpert, Bureau of Labor Statistics, U.S. Department of Labor

The final phase of any projections process is the evaluation of the projections once actual data for the target year become available. The Office of Occupational Statistics and Employment Projections periodically evaluates the results of past projections in order to gauge how accurately the projections tracked actual employment growth. This study examines the accuracy of the 1988-2000 occupational employment projections. The analysis provides insight into why projections for certain occupations were particularly accurate or fell short of the mark.

NAICS Conversion Issues in Occupational Statistics and Employment Projections

Norman Saunders, Bureau of Labor Statistics, U.S. Department of Labor

The replacement of the Standard Industrial Classification (SIC) with the North American Industry Classification System (NAICS) will have major impacts on all forecasting programs depending upon coherent industry-based time series. Not only are many underlying concepts changed dramatically but the long historical time series necessary for optimal statistical estimation will not be available for some time into the future. This presentation proposes approaches within the BLS projections program, a large-scale examination of industry and occupational employment trends, that allow for an ordered transition to the new classification system and at the same time maintain the quality and usability of the historical and projected industry time series data.

This page intentionally left blank.

Evaluating the BLS Labor Force Projections to 2000

Howard N Fullerton Jr.
Bureau of Labor Statistics, Office of Occupational Statistics and Employment Projections

The Bureau of Labor Statistics (BLS) prepared five labor force projections to or through 2000. Three of those projections had errors of a million or less; the most extreme errors ranged 1.5 percent above or below the actual 2000 labor force of 140.9 million. The growth rate of the labor force is also crucial to the BLS employment projection program. The error in the growth rate varied by a tenth of a percentage point above or below the actual growth rate for the periods over which the projection was made. At the same time, projections of the civilian noninstitutional population were uniformly low. Thus the labor force participation rate projections were generally too high.

BLS has made labor force projections since the late 1950s. These projections were made for several demographic groups: age, sex, race and Hispanic origin. Beginning in 1968, BLS has reviewed past labor force projections. Such evaluations help both those making and those using the projections understand the sources of error the and the accuracy of specific components.

Until recently BLS projections focused on years ending in five, so evaluations took place at five-year intervals. This paper is an evaluation of the BLS labor force projections to 2000. Beginning in 1987 and continuing to 1995, BLS prepared five projections either to or through 2000.[1] This article examines the difference between the projections and the labor force as estimated in the Current Population Survey (CPS) using weights from the 1990 census. The differences or errors are calculated by sex for detailed age groups of the white, black, Asian and other, and Hispanic origin population and labor force. (Earlier of these projections did not have as much age detail for Hispanics as for the other groups.) Each of the five projections to 2000 had three alternatives: high, moderate, and low. This analysis, for the most part, focuses on the middle or "moderate" growth projection in each series. Where appropriate, the accuracy of the five 2000 projections are compared with evaluations of BLS projections to 1985, 1990, and 1995. Each of the projections is identified by the year from which the projection was made (1986, 1988, 1990, 1992, and 1994).[2]

One of the challenges in evaluating projections is that the estimates are not strictly comparable to the data projected. After the 1990 census, extensive changes to the CPS were implemented in 1994. These changes included an adjustment for the undercount, as well as changes in the questions asked. The latter resulted in a greater proportion of women and older persons being counted in the labor force. It is not possible to quantify the effect of these improvements in the survey, so it is not possible to know how much they affect projection accuracy. However, it is clear that projections made before 1994 did not anticipate the effects of the redesign and that projections made after 1994 did not immediately incorporate all the changes.

Another challenge in evaluation is the different uses made of the labor force projections. Some use the total labor force—indeed, the growth rate of the labor force—not needing any of the components. For many users, some part of the labor force is vital, for example, youth workers or older workers. Others use the projected labor force participation rates for market research or to project state populations. Another group of users focus on the distribution of the labor force by race and sex. No one measure of error or quality satisfies all these users. Further, there are two sources of error, projected population and projected labor force participation rates. It would be helpful to know how these combine to produce the errors in the labor force projections.

Evaluation of the aggregate 2000 projections

The following tabulation shows the projections to 2000 in millions and the numerical and percent error made in each year the projections were based:

Projection for 2000 made from:	Labor force		
		Error	
	Millions		Percent
1986	138.8	-2.1	-1.5
1988	141.1	0.3	0.2
1990	142.9	2.0	1.5
1992	141.8	1.0	0.7
1994	140.0	-0.9	-0.6
Actual:			
2000	140.9		

The overall errors were greatest in 1986 and 1990; except for these two years, the errors were less than 1 percent. The first three projections were also evaluated for 1995. It is interesting to note that the numerical errors are less for 2000 than for 1995 with the 1988 and 1990 projections. It is possible for a projection to improve with age. The error information above indicates that short versus long time-span does not seem to be a factor improving the accuracy of labor force projections. A similar conclusion would be inferred from earlier analysis.

For some users, the absolute error or the percent error is not relevant but the error in the growth rate is. The following tabulation displays the growth rates for the civilian labor force historically with the projected annual growth rate and the actual growth rate. All three rates are in a row are measured over the same number of years. The historic rate is calculated over the same number of years *before* the date of the projection as 2000 is *after* the date of the projection:

Projection for 2000 made in:	Historical rate	Projected rate	Actual rate	Error
1986	2.2	1.2	1.3	-0.1
1988	2.0	1.2	1.2	0.0
1990	1.6	1.3	1.1	0.1
1992	1.5	1.3	1.2	0.1
1994	1.2	1.1	1.2	-0.1

The error in the annual growth rate from 1988 was 0.02 percent. For four other projections, the error in the annual growth rate was either –0.1 or 0.1. For those using the projections to forecast either employment or economic growth, this level of error would be minor. For growth rates, BLS projected variously that the rate of growth would slow significantly from past rates of growth (by a full percentage point in the 1986 projection) to not much different from past rates of growth (by a tenth of a point in the 1994 projection). Except for the 1994 projection, when BLS projected a decrease in the growth rate and the labor force continued to grow at past rates, the change was in the correct direction and the

error in the growth rate was less than the projected change in the growth rate.

Population projections

BLS labor force projections are prepared using the incidence method: age-sex-race or -Hispanic origin labor force participation rates are multiplied by comparable projections of the population prepared by the Bureau of the Census.[3] For all these projections, BLS adjusted the projection to provide the civilian, noninstitutional population. Although errors were made in making this adjustment, they are not considered to be sufficiently large to incorporate into this analysis. Some sense of the size of this type of error may be garnered by seeing how the errors in the adjusted population varies for the first two labor force projection. For the projection from 1994, the projected population was also adjusted for the 1990 undercount since the CPS itself was so adjusted.

Population projections have three components: births, deaths, and net immigration. Each of these may be a source of error as well as the initial population from which the projection is made. Because these projections spanned a period of less than 16 years, errors in births did not affect the size or composition of the labor force. Although it is true that there were fewer deaths than projected, most of those extended lives occurred at older ages. The source of the discrepancy must be net immigration either over the projection or as part of the estimate of the base year population. If so, then errors would be larger for Hispanics and Asian and others. The Bureau of the Census prepares its own evaluation of their population projections; this paper only looks at the population projections as they affect the size and composition of the labor force.[4]

For the past decade, population growth has accounted for more labor force growth than has the labor force participation rate change. Thus the accuracy of population projections should be crucial to the accuracy of the labor force projections. The following tabulation shows the 2000 projections for the civilian, noninstitutional population aged 16 and with the errors associated with the total population projections:

Projections for 2000 made in:	Total	Error	
	(in millions)	(in percent)	
1986	204.7	-5.0	-2.4
1988	204.6	-5.1	-2.4
1990	208.0	-1.7	-0.8
1992	208.0	-1.7	-0.8
1994	208.8	-0.9	-0.4
Actual			
2000	209.7		

Unlike the labor force projection, all the population projections were low. Unlike the labor force projections, the population projections show steady improvement. The difference between the percent errors in the first tabulation and this one indicate that BLS made offsetting errors in labor force participation rates, reducing the errors in the aggregate labor force. The following tabulation presents hypothetical labor force projections using the projected population and the actual 2000 labor force participation rates:

Projections for 2000 made in:	Total	Error	Difference from actual error:	Percent error:
	(in millions)			
1986	136.9	-3.9	-1.8	-2.8
1988	136.8	-4.0	-4.3	-2.9
1990	139.5	-1.3	-3.4	-1.0
1992	139.2	-1.6	-2.6	-1.9
1994	140.4	-0.5	0.4	-.3

The numerical errors made in this hypothetical projection are less than for the population. Except for the projection from 1994, these projections would have a larger error than the projections that were made: the labor force would have been even smaller. The percent errors for these hypothetical labor force projections were different from that for the population projection and, except for 1994, greater.

To trace the error, the Mean Absolute Percent Error (MAPE) may be calculated at differing levels of aggregation. The following tabulation provides the MAPE's for various aggregations (in percent):

Projection for 2000 made in:	1986	1988	1990	1992	1994
Aggregate error	2.4	2.4	0.8	0.8	0.4
MAPE for race	6.4	6.4	4.3	3.5	1.5
MAPE for sex and race	5.5	5.6	4.2	3.4	1.5
MAPE for sex, race, and age	7.6	7.7	6.3	4.1	3.6

The MAPE for the total is the absolute value of the percent error. The MAPE for men and women considered separately averages to the overall MAPE's, so it is not displayed. When MAPE's are calculated for the three race and one Hispanic origin group, the MAPE's are larger, but the relative standing of the various projections does not change. The errors made when projecting by race offset, giving a more accurate total projection than for any race group. When gender and race are considered, the MAPE's decrease for the first two projections and do not increase for any group. Finally, accounting for age with gender and race results in a larger aggregate error. However, for those who are not using the detailed projection, these projections are better because of the use of age, sex, and race or Hispanic origin. Examination of the detailed projections does not indicate that using more aggregated age groups would have increased the accuracy of the overall projections.

The population of both men and women were under-projected. The difference was greater for men than for women through the projection from 1990. The first two projections had markedly larger projection errors than the last three. That the error was larger for men than for women reflects the greater tendency for men to be undocumented immigrants. Since population projection errors improved as time passed, it is likely that errors because of under-estimates of undocumented workers also decreased. All five projections correctly projected that there would be substantially more women than men in the 16 and older population.

For all five labor force projections, (three population projections), the size of the white population was under projected. As whites comprised 84 percent of the population in 2000, they should also account for most of the error. On the other hand, generally it is easier to measure and project large groups. For all the projections, white's errors were less than 84 percent of the error. Except for the projection from 1990, whites accounted for more than half of the projection error.

Two population groups would be expected to be hard to project: Asians and others and Hispanics. Both groups have high immigration, are fairly heterogeneous, and are relatively small. Asians and others accounted for five percent of the 16 and older population in 2000, but for each

of the projections, the first population projection (first two labor force projections), they accounted for 27 percent of the error. For the next population projection, their numerical error slightly exceeded the error for whites. For the last population projection, the errors were much smaller, accounting for 16 percent of error. However, their projected population was higher than actual, unlike the other three groups.

Hispanics may be of any race, however, more than 90 percent are white. Thus, errors in projecting the numbers of Hispanics carry into the number of whites. Since Hispanics have high immigration rates and it is estimated they are a large component of undocumented immigration, it should be no surprise that the Hispanic population is difficult to project accurately. Hispanics accounted for 11 percent of the 16 and older population in 2000. Errors in their population projection accounted for 38 percent of the error from 1986 and 37 percent of the projections from 1988. For the labor force projection from 1990, which used the same population projection for Hispanics as the previous two, the error was the same size (1.9 million low), but it now exceeded the total population error (1.7 million low). This projection was not based on the 1990 census. The next two projections were the relative size of the projection errors decreased. Even so, the error in the number of Hispanics exceeded that for whites in the projection from 1992. The dynamic changes in the Hispanic population are reflected in the difficulties of projecting this group.

For the first two population (first three labor force projections), the black population had relatively small errors, less than their share of the population, 12 percent. This population group, though growing faster than the overall population, has demonstrated a consistant path of growth. The black population was the most accurately projected group in the projections to 1995. For the projection from 1992 the error was much larger, and accounted for 30 percent of total error. For the projection from 1994, although the size of the numerical error was smallest of the five, because the total error was by far the smallest, black's share of the projection error was larger than their share of the population.

For each of the five projections, there are 108 errors to examine at the level of age, sex, race or Hispanic origin. Summaries are needed. The following tabulation provides summary

information about the depth and dispersion of the errors, in thousands:

Year	1986	1988	1990	1992	1994
Depth					
Lowest	-965	-958	-622	-495	-273
Lowest one-eighth	-167	-167	-114	-106	-53
Lowest quarter	-100	-108	-56	-44	-25
Half (median)	-36	-36	-22	-6	-10
Highest quarters	-1	-1	23	16	8
Highest eighth	41	41	53	33	24
Highest	136	137	226	195	75
Dispersion					
Inner 50 percent	100	107	78	59	33
Inner 75 percent	208	208	167	139	76
Range	1,101	1,095	848	691	348

Which groups had the lowest under-projection? For the projections from 1986 through 1992, it was white men aged 20 to 24. For the projection from 1994, it was Hispanic women 25 to 34. The error for this group of women was always in the lowest one-eighth. For white men 20 to 24 are a large group with a large error. Their relative errors are smaller. The relative errors for Hispanic women 25 to 34 are larger than those for white men ages 20 to 24. Hispanic men 20 to 24 also have large errors, absolute and relative for the first population projections. Errors in projecting the size of the 20 to 24 and 25 to 34 Hispanic population also affected the size of the white population the same age.

What groups were the most over-projected? This varied; for the first projection, it was white men ages 50 to 54, for the next population projection, it was 30 to 34, followed by white men 35 to 39. For the last projection, Asian and other women aged 50 to 54 was the most over-projected group. Again, white men are a large group, the source of a large error. For the last projection, Asian and others were over-projected as a group, so it is not surprising to find an age group from this racial cluster represented. White men's age groups were over- and under-projected, by large amounts. Older white men's population was uniformly over projected.

At this point it is clear that the population projections were too low; given that the aggregate labor force projections were much more accurate, it is easy to infer that the projected labor force participation rates must be too high. It is not clear what effect the errors in the population projections had on users of the labor force projections interested in the distribution of the labor force by race or sex. That question must be answered after examining the labor force participation rate projections.

Labor force participation rates

What the BLS brings to the labor force projection process is its projection of labor force participation rates. Although the population projections currently account for most of projected labor force change, study of the errors made in projecting the labor force participation rates is important since that is the part contributed by BLS. The following tabulation shows the overall labor force participation rate for the five projections with those for men and women.

Projections for 2000 made in	Percent	Error (in percentage points)
1986	67.8	0.6
1988	69.0	1.8
1990	68.7	1.5
1992	68.2	1.0
1994	67.0	-.2
Actual	67.2	

Four of the five projections had the aggregate labor force participation higher than the actual. As the tabulation indicates, the aggregate labor force rate has yet to reach 68 percent, though three of the projections anticipated that this would happen by 2000. Given that 2000 was the last year in a sequence of high economic growth, it is significant that the projected labor force rates were higher than the actual. From the projection made in 1988 on, the error in the aggregate labor force participation rate decreased for each projection. However, the 1986 projection was the second most accurate.

Mean absolute percentage errors may also be calculated for the labor force participation rates. For the aggregate error, they are absolute value of the relative errors. The following tabulation provides MAPE's for various aggregations, in percent:

Projections for 2000 made in:	1986	1988	1990	1992	1994
Aggregate error	0.9	2.7	2.3	1.5	0.3
MAPE for sex	1.1	2.8	2.4	1.6	.8
MAPE for race	1.2	1.9	.9	.8	3.1
MAPE for sex and race	4.4	5.9	2.8	2.1	3.0
MAPE for sex, race, and age	14.5	9.8	5.7	6.2	5.1

Looking at errors by gender provides little additional information beyond that for aggregate error—the greatest difference from the aggregate error occurs with the projection from 1994, which had the rate for women too high and that for men too low—since there is no reward for offsetting errors. The MAPE for race indicates that the worst projection was the one from 1994. Looking at the labor force rates for the four race, Hispanic origin groups shows that the percentage point error for 1994 was zero for whites, their best projection, but that the projection from 1994 was by far the worst for blacks, Asians and others, and Hispanics. The MAPEs were not weighted by size of group. Whites were 83 percent of the 2000 labor force, so that for weighted measures of error, the most accurate year for the overall labor force would be the most accurate year for whites. Turning to the MAPEs by race and gender, the projection from 1988 was least accurate. It was not the case that a good projection for men implied a good projection for women but certainly the converse was not true. (The correlation of men and women's errors is .33.) When the age structure is also considered, then the projection from 1986 had the greatest MAPE. The projection for this year also had the greatest numerical error. Both the population and the labor force participation projections contributed to this error in the 1986 projection, with the population too low and the participation too high.

The labor force participation rate projections from 1994 had the lowest error for whites, but the worst for other race groups. Since whites are the majority of the labor force, the 1994 projection had the lowest error in labor force participation rates. The 1986 projection had large errors in both the population and labor force participation rate projections.

There are 108 labor force participation rate projection errors to examine; the following tabulation summarizes the errors in the participation rates, in percentage points:

	1986	1988	1990	1992	1994
			Depth		
Lowest	-16.9	-9.8	-8.5	-9.0	-10.6
Lowest eighth	-5.8	-3.5	-2.6	-3.4	-4.9
Lowest quarter	-3.1	-2.1	-1.8	-1.3	-3.6
Half (median)	-0.1	0.6	0.6	0.4	-1.6
Highest quarter	3.1	3.5	2.3	1.8	0.3
Highest eight	6.6	6.1	4.3	3.4	1.7
Highest	12.4	11.1	8.3	7.6	7.1
			Dispersion		
Inner 50 percent	6.1	5.6	4.1	3.1	3.9
Inner 75 percent	12.3	9.7	6.9	6.8	6.6
Range	29.3	20.9	16.8	16.6	17.8

The aggregate labor force participation rates were too high in four of the five projections; the median of the errors of the age-sex-race or

Hispanic origin participation rates were closer to zero than the errors of the aggregate, with the exception of the most recent projection. If the thesis is that the labor force participation rates were too high to offset population projections that were too low, then four projections of the five fit that mold. This information is also available in chart 1, which has box-and-whisker plots for the five projections.

One desirable characteristic of the projections as a sequence would be that the dispersion of the errors would be less for the more recent projections. The measures of dispersion and chart 1 indicate that this was taking place until the 1994 projection. That the most recent projection studied for accuracy is not the most recent made seems to be a characteristic of labor force projections, this also happened with the projections to 1990 and 1995.

To examine the question, "Were some age groups harder to project than others?" turn to chart 2, which has box-and-whisker plots of the errors by age-sex-race/Hispanic origin groups. (We have six projection errors for white women aged 20 to 24, six for black women of the same age, and so on.) Although the median of the errors by projection year are near zero, except for the 1994 projection, the data by age indicate that there was significant variation in the errors by age. For the age groups 25 to 54, which exhibit the highest labor force participation rates, the median of the errors were either high or near zero, giving the source of the high aggregate labor force participation rates. For the older ages, the median of the errors were below zero. For these age groups where there is now great interest in their pattern of labor force participation, there was a consistent pattern of too low labor force participation. Labor force participation rates for older men increased from 1985 to 1990, then decreased until 1994 and have increased since then. These changes did not start at the same time for all groups of older men. Starting with the 1996 labor force projections, BLS has projected this change in trend. It was among the first the do so.

According to the box-and-whisker plots of labor force participation rates by age group, chart 2, it is clear that the age groups younger than 60 were over projected. The labor force participation rates for groups older than 60 were uniformly under projected. Some age groups were harder to project than others. The two age groups with the largest boxes were those 18 and 19 and 65 to 69. The latter group had the most extreme errors. However, the extreme errors for those 65 to 69 were high—for Asian and other men in the 1986 projection and Asian and other women in the 1988 projection.

Labor force

At this point, it is clear that the labor force participation rate projections were, as a group, too high. However, the aggregate labor force was fairly accurately projected. As the new labor force projections are reviewed, the reviewers know independently how fast employment was likely to grow. It appears that this review of the labor force projection resulted in an accurate aggregate labor force. In the face of low population projections, labor force participation rates were increased, resulting in an accurate projection of the labor force.

BLS labor force projections have been characterized as having men's labor force too high and women's too low. For three of the five projections the men's labor force lower was than the actual. For women, all were higher than the actual. The traditional view of BLS labor force projection is now wrong. For the two projections with the largest overall error, the male labor force was off by 2.2 million in 1986; women's labor force was off by 1.4 million in 1990. In the projection with least overall error, men and women's errors offset. There does not seem to be a pattern of projecting labor force better for one gender than the other.

The following tabulation shows MAPEs for various aggregations:

Projection for 2000 made in:	1986	1988	1990	1992	1994
Aggregate error	1.5	0.2	1.5	0.7	0.6
MAPE for sex	1.4	1.5	1.5	.7	.8
MAPE for race	6.2	5.9	4.7	3.9	3.4
MAPE for sex and race	6.3	5.6	5.5	4.0	4.1
MAPE for sex, race, and age	15.4	12.3	9.1	7.5	8.5

The first row repeats information from the overview. Once gender is taken into account, the 1988 projection error increases. The 1988 projection had a highly accurate projection of the level, but men's labor force was too low and women's too high. The other four projections did not have large offsetting errors by sex. The accuracy of the overall projection is the result of offsetting errors. The more detailed measures reveal where the errors where made. Thus, taking race and Hispanic origin into account increases the error because less of the offset is concealed. In the 1988 and 1990 projections, the projected

white labor force was too large, while the black and Asian and other labor force was projected too low. The 1994 had an accurate projection of the white labor force, but that for blacks was almost a million low. For all the projections, Hispanics were under projected, by substantial amounts.

Taking race and gender into account, the error in the 1994 projection rises; this is because the accuracy of the white labor force is due to sizable offsetting errors in the men and women's labor force. Once age, sex, race (and Hispanic origin) is taken into account, the errors increase, as offsetting errors of having some ages too high and others too low are taken into account. This shows the pattern of error decreasing from the 1986 to the 1992 projection, then increasing. The accuracy of the overall labor force was obtained through offsetting errors.

The following tabulation summarizes the 108 errors in the components of the labor force in thousands:

	1986	1988	1990	1992	1994
			Depth		
Lowest	-518	-543	-372	-326	-290
Lowest eighth	-244	-140	-118	-97	-100
Lowest quarter	-93	-75	-62	-43	-59
Half (median)	-16	-14	-4	-6	-13
Highest quarter	0	9	21	13	4
Highest eight	39	55	97	61	26
Highest	712	772	563	765	230
			Dispersion		
Inner 50 percent	94	84	84	55	64
Inner 75 percent	282	195	215	158	126
Range	1,230	1,315	935	1,091	520

The median of the individual errors are all small, but negative. The low quartiles or hinges are all negative and the high hinges are all positive—the errors are grouped around zero. The innerquartile range decreaseed from the 1986 projection to the 1992 one, before a slight increase for the 1994 projection. However, the range and the inner 75 percent show a decrease through 1994. The errors for the 1994 projection were systematic, but not large.

The white population and labor force is significantly larger than the black, Hispanic, or Asian and other population and labor forces. Thus, the largest numerical errors are in white groups. For the 1986 through 1992, the group with the largest over projection was white women ages 35 to 39, for the 1994 projection, white women 40 to 44 had the greatest error. For the first four projections, white men 20 to 24 were under projected the most. For the 1994 projection Hispanic men 25 to 34 were the group most under projected.

The older labor force had the greatest relative errors. The labor force for these ages is small, so a modest numerical error yields a large relative error. See chart 3 for relative errors by age group. For those age groups with high labor force participation, the relative errors had a median of zero and the errors were closely grouped around the median. Older ages, which had too-low labor force participation rate projections for men, have negative median errors and wide dispersion around the median. Thus, the greatest errors in the labor force were at ages with modest impact on the size of the labor force. This is confirmed if a box and whisker chart of the errors in thousands is examined. If a user were particularly interested in labor force

participation of older workers or the size of their labor force, this set of projections would have been relatively unhelpful.

For some users, the size and growth rate of the labor force is unimportant; the concern is for the distribution between men and women, among the various race and ethnic groups, or among the various age groups. The following tabulation presents the index of dissimilarity comparing the projections to the 2000 actual, by various levels of aggregation.

Projection for 2000 made in:	*1986*	*1988*	*1990*	*1992*	*1994*
Index of dissimilarity, sex	0.7	0.8	0.3	0.3	0.4
Index of dissimilarity, race	0.6	0.8	0.7	0.5	0.6
Index of dissimilarity, race and sex	1.3	1.5	1.0	0.9	0.9
Index of dissimilarity, race, sex, and age	3.3	2.6	2.0	1.9	1.3

The index of dissimilarity may be interpreted as the amount the one distribution has to change to be like another. In these cases, it records how much the projected distribution has to change to be like the actual 2000 labor force distribution. Thus, the 1986 projection would have had to change by .7 of a percentage point to reflect the actual distribution of the labor force between men and women. The projections were also quite good in reflecting the actual composition of the labor force by race. Taking race and gender into account—was the share of black women correctly projected?—there is a higher index, or greater error. However, in the worst year, 1988, the distribution would have only needed to change by 1.5 percentage points. Once age, race, sex, and age is taken into account, the indexes increase again; however, they improve with time and the error in the worst year is 3.3 percent. Even though the older labor force was under projected, the age composition of the labor force was fairly well projected.

Alternatives and confidence intervals

For each of these labor force projections, BLS prepared three alternatives. Since the presentation of the projections focused on the middle or moderate alternative, this analysis has also. However, a user could reasonably expect the 2000 labor force to be between the low and high alternatives. The following tabulation presents the high and low alternatives for each projection:

Projection for 2000 made in:	*High alternative*		*Low alternative*	
	Labor force (thousands)	*Parti- cipation*	*Labor force (thousands)*	*Parti- cipation*
1986	141.1	68.0	134.5	65.7
1988	146.8	70.7	137.7	67.3
1990	156.2	71.5	141.8	66.1
1992	156.5	70.1	147.3	67.3
1994	153.4	68.7	143.6	65.5
2000	140.9	67.2	140.9	67.2

Unlike projections for earlier years, some of these alternative projections did not cover the actual. Only the projections prepared in 1986 bracketed or covered both the actual 2000 labor force and the participation rate. Given the characteristics of the projections with the labor force more accurately projected than the labor force participation rates, one would expect that the labor force projections would cover the actual while the labor force participation rates would not. For three of the projections the low alternative labor force was higher than the 2000 actual. This happened for only two of the labor force participation rate projections. The 1992 projection was the only one to have neither the labor force or participation rate confidence interval cover the actual. Every possible combination of covering and not covering occurred among the five projections. Evaluations of projections to earlier years indicated that the actual labor force projection was covered by the alternatives.

Concluding thoughts

The review process for preparing labor force projections resulted in a more accurate projection of the size of the labor force. Faced with population projections that were too low, subtle adjustments in the labor force participation rate were made for the work force ages 30 to 64, resulting in somewhat high aggregate labor force

participation rates. For those users of the labor force projections who needed projections of the size of the total labor force or of its growth rate, this projection would have served them well. For those users of projected labor force participation rates, the significant problem was with projections for older workers, whose rates were too low.

Endnotes

---

[1] Howard N Fullerton, Jr., "Labor force projections: 1986 to 2000," *Monthly Labor Review*, September 1987, pp. 19-29. Reprinted with additional detail in *Projections 2000*, BLS Bulletin 2302, March 1988. Howard N Fullerton, Jr., "New labor force projections, spanning 1988 to 2000," *Monthly Labor Review*, November 1989, pp. 3-12. Reprinted with additional detail in *Outlook 2000*, BLS Bulletin 2352, April 1990. Howard N Fullerton, Jr., "Labor force projections: the baby boom moves on," *Monthly Labor Review*, November 1991, pp. 31-44. Reprinted with additional detail in *Outlook 1990-2005*, BLS Bulletin 2402, May 1992. Howard N Fullerton, Jr., "Another look at the labor force," *Monthly Labor Review*, November 1993, pp. 31-40. Reprinted with additional detail in *The American Work Force: 1992-2005*, BLS Bulletin 2452, April 1994. Howard N Fullerton, Jr., "The 2005 labor force: growing, but slowly," *Monthly Labor Review*, November 1995, pp. 29-44.

[2] Fullerton, Howard N, Jr., "Evaluating the 1995 BLS labor force projections," *Proceedings of the Section on Government Statistics and Section on Social Statistics*, (Alexandria, VA, American Statistical Association, 1997), pp.394-399.

[3] Projections of the Population of the United States by Age, Sex and Race: 1987 to 2080*, Current Population Reports*, Series P-25, No. 1018, (Washington, Bureau of the Census, 1989). Projections of the Population of the United States by Age, Sex and Race: 1993 to 2050*, Current Population Reports*, Series P-25, No. 1104, (Washington, Bureau of the Census, 1993). Projections of the Population of the United States by Age, Sex and Race: 1995 to 2050*, Current Population Reports*, Series P-25, No. 1130, (Washington, Bureau of the Census, 1995).

[4] Tammany Mulder, "Accuracy of the U.S. Census Bureau National Population Projections and Their Respective Components of Change," Population Division Working Paper No. 50, (Washington, D.C., Bureau of the Census, forthcoming).

.

## Chart 1. Percentage point error in projected labor force participation rates
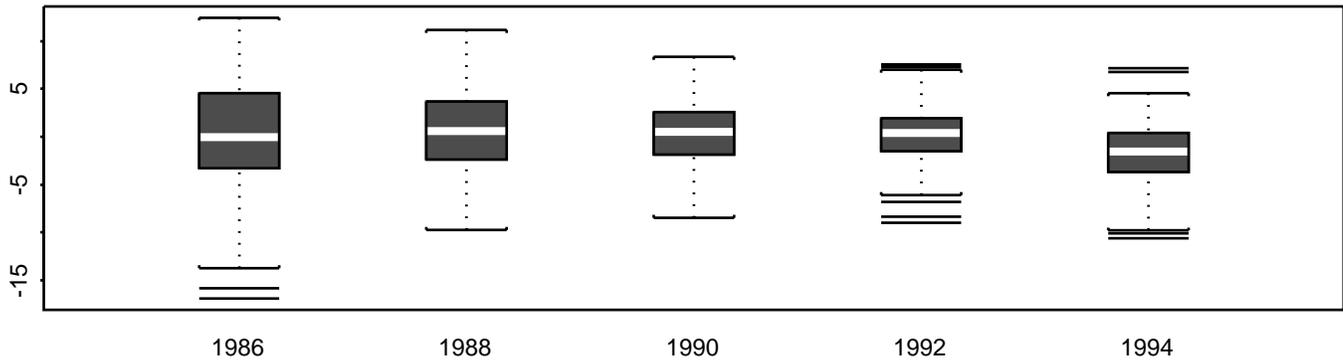
Chart 2. Percentage point error in participation rates, by age, 1986-94
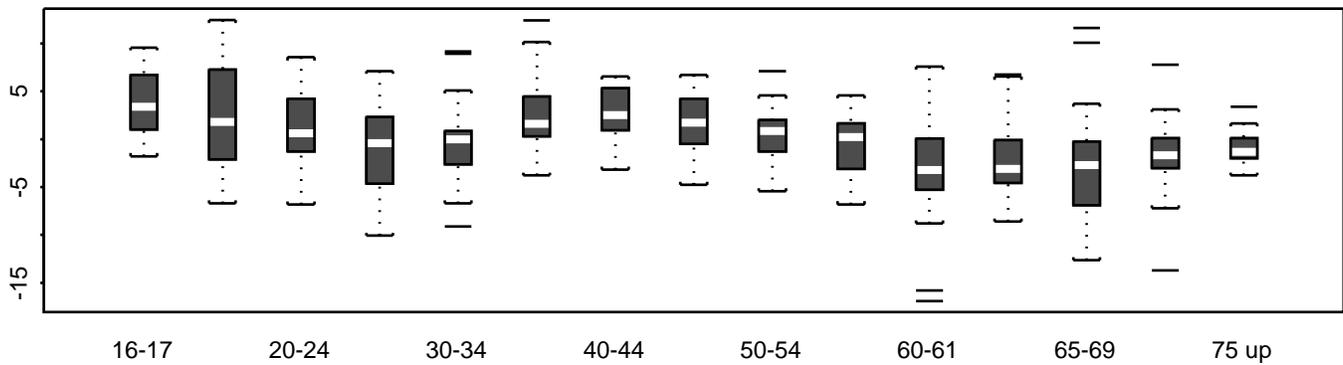

Chart 3. Error in labor force, by age, 1986-94, in percent
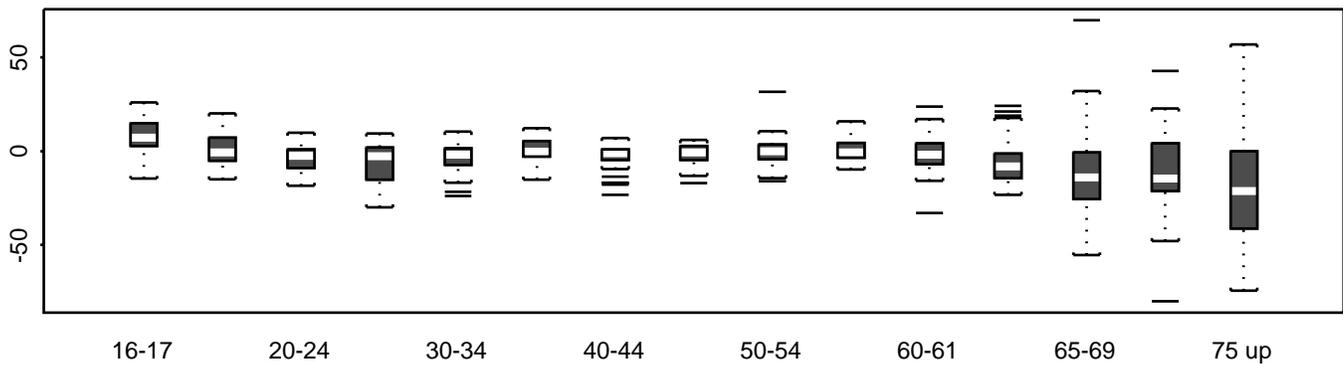
An Evaluation of the 2000 Industry Employment Projections

Arthur Andreassen
Bureau of Labor Statistics
U.S. Department of Labor

Introduction

Biennially the Bureau of Labor Statistics makes projections of employment by industry 10 years into the future. Eventually, when the actual data become available, these projections are evaluated. This paper compares the industry employment projections of 2000 that were published in November 1989 with the actual. Emphasis is placed on how well the process performed while also attempting to trace the sources of discrepancies. Since the projection's procedure is a cohesive web, errors in assumptions at early stages adversely impact what follows. Further, from the evaluation of projected variables, clues can be derived as to the extent the economy has diverged from past trends. From this process not only will the results be evaluated but a better understanding of the evolution of the economic structure over this period will be gleaned.

Starting with estimates of the available labor its intensity of use was calculated. Offsetting errors benefited the conversion of the working age population to household employment such that it was very close to actual, table 1. The combination of a working age population underestimated by 5.1 million with an overestimated participation rate resulted in a labor force slightly overestimated by 200 thousand. Mixing this supply of labor with assumed rates of unemployment and of productivity growth gave an assumed level of GDP. Here a too high unemployment rate co-joined with a too low productivity rate projected a GDP that grew too slowly.

Although projections seemingly present a picture of the future they obviously are based on and thus incorporate much of the ideas of the year in which made. In 1988 BLS projected to 2000 what was considered a non-inflationary, full employment economy that encompassed generally accepted assumptions of labor force participation and productivity growth rates. Despite these very optimistic assumptions baked into the projections, 2000 turned out to be even better as the economy went on an unprecedented tear in the last half of the nineties. In 1988 the economy was already experiencing an upturn that was getting long in tooth, at six years it was already the second longest of the eight post war cycles up to that time. A 5.5% unemployment rate was considered as low as it could reasonably be

expected to go before inflation reared up and imbalances appeared. However, rather than ending, this upturn would continue for another twelve years pausing for a slight 8 month downturn in 1990 which, at that time, was the post war's second shortest Eventually the unemployment rate declined to a monthly low of 3.9% in October of 2000. Jobs would total 143.8 million after increasing at an annual growth rate of 1.7% rather than the 136.2 million jobs at 1.2% that were projected while the consumer price index declined from 4.1% to 3.4%. In short, the economy functioned during the projected period in a way much better than the past would lead one to expect and most of the employment data demonstrate this. After determining this measure of GDP the sources specific to each industry is projected. It is in this area that the BLS method shines in spotlighting industry specific growth.

Demand

Assumed GDP is first decomposed into its major demand components. Getting this allocation correct is vital because each category purchases from a particular set of industries such that relative category growth affects industry employment being the source of much industry variation. Table 2 compares historical growth 1976 to 1988 to both the projected and actual rates from 1988 to 2000. First noticed is the central tendency of the projected sector growth rates as compared to the variation of the real economy. This is a result of the inherent conservative nature of the projections process and the muting of large swings as the review process is carried out. Second to notice is how much closer the 1976 to 1988 period was to the next 12 year period as oppose to the projections. The simple answer to this is the similarity in cyclical paths each period took. In 1976 the unemployment rate was 7.7% as the economy was continuing to recover from the 16 month recession that started in 1973. In 1988 the economy was also in an expansion phase although it was not realized at the time so the similarity in growth reflected cyclical expansions.

Study by sector will highlight the sources of growth over the projected period. Personal consumption expenditures, PCE, is responsible for 2/3 of total demand and is the elephant in the rowboat for many industries. Fortunately this category is very stable, it changes gradually over time and closely mirrors GDP

growth. It continued to do so and we so projected correctly. Moving to the other categories we see numerical substantiation of two themes that resonated through the nineties, globalization and investment in information technology. These are made explicit as investment, exports and imports grew much faster than the total. Again, dependence on past trends was not a preparation for the extraordinary capital investment that defined the last years of the nineties.

Investment benefited from the rapid decline in computer prices concurrent with a need for capital to replace workers becoming scarce as the unemployment rate dropped. Investment spending is mainly in the durable industries portion of manufacturing. Rounding out the investment category is the change in business inventories which is very sensitive to short term economic movements and while increases in the GDP components are usually viewed as positive this is not so for inventories. For the past three decades advances in computers and communications has increased efficiency in the stocking of goods so the inventory to sales ratio has been declining as was so continued through the projections. However in 2000 measures of industrial production, which closely tracks manufacturing, peaked in June and started to decline causing stocks to quickly back up and throwing inventories off its long run trend.

Both exports and imports grew faster than projected as did the deficit in net exports. Both of these sectors have offsetting impacts mainly in manufacturing. Imports satisfy a growing share of apparel, autos and footwear while exports ship from the computer and communications industries. Government, both the Federal and State and local branches complete the demand categories. Only the portion of government expenditures that are purchases of goods and services are included in these categories, not spending on transfer payments and grants in aid, e.g. Federal demand is mainly for defense and, excluding the armed forces and bureaucrats, is spent heavily on manufactured goods. Federal purchases were projected to decline and did so as defense, which is 66% of total Federal purchases. State and local governments specialize in compensation, education and health and grew twice the 1.5% projected rising from half as great to twice the Federal share. Nothing untoward occurred in these areas that threw them off past trends so the projections were pretty good. In fact the projections got the relative growth rates of all but one of the demand components correct.

Employment

Over time, variation occurs not only among the categories but also within each category as to what it

buys. Due to such things as relative prices, the advance of knowledge, new sources of factor supply, demographic changes and changes in taste, the goods supplied to each component will change. Only one half of output and employment is used too satisfy final demand. To measure fully the impacts on industries the other half, which is used as inputs must be accounted for. To do this after the demand components have been allocated to the goods that are singular to them any trends detected in industry production processes are pushed forward. Although changes in the production process occur slowly they also influence employment growth. Many of the same changes that impact final demand do so to the technological inputs. Here also is a shift over time from hard to service goods as information technology and communications increase the speed and efficiency of production and make just in time inventory control and outsourcing endemic.

Final demand is converted to industry output by running it through the production process. Then projected industry productivity measures are applied to the dollar outputs generating employment. Changes in dollar output do not translate directly to a one for one increase in employment since increases in productivity replaces some of the requirement for workers. Derivation of projected GDP was too low because of the assumption of too low productivity growth . Higher projected productivity partially offset the assumed too high unemployment rate giving employment closer to actual, e.g., GDP projected to grow 16% slower than actual gave an employment only off by 9%. Major industry category projected employment, table 3, synopsizes the results. Trends in relative employment growth continued through the projected period. Productivity increases are very healthy in manufacturing, especially in the durable goods area so strong demand growth will not be completely reflected in strong employment growth here. Exceptional growth in technology demand could not even slow the relative decline that was even greater than expected. Services however more than took up the slack. The rest of the categories performed as expected which means they continued their past trends.

Each of the five distinct components that make up the projection process can be isolated and their contribution to the amount that the projections were off can be ascertained, table 4. The two largest contributing factors to errors have previously been discussed and are here quantified. Column 1 is actual employment by major sector. GDP is in column 2 and because of the assumptions of a too high unemployment and a too low productivity contributed to employment projections too low by 22.6 million. On the other hand, as seen in column 6, that very low assumption of productivity led to a 27 million over shoot in employment that more than

canceled out the GDP error. Columns 3 through 5 illustrate the contributions made by: the distribution of GDP to its sectors, the distribution of GDP sectors to their industry demand and, finally, the production process or technological coefficients. The errors of these other three components had little relative impact on the total.

Results

Evaluations can be carried out on either an absolute basis, how close they are to actual values, or on a relative basis, how correct is their direction and the rate of change. Since the BLS projections serves as a guide to choose a field of endeavor it is more important to correctly determine an industry's future growth rate as an indication to entrance or avoidance. When the results are compared by relative growth rates to those of a very simplistic projection method they are not bad. The simplistic method used as a foil replicates the past 12 years' growth rate into the future. This comparison shows the BLS to be better in 114 of 175 industries' employment, table 5 (the bold values in column 5 are those industries which were better projected by the simplistic method). Rather than comparing periods of equal length whose terminal years may be at different

points on the business cycle, such as 1976 and 1988 with unemployment rates of 7.5% to 5.5%, periods with equal unemployment terminal years would exclude some cyclical noise. This rate equality was approached in 1979 and 1988 and the BLS still came out better in 119 industries. Of the 42 industries that grew faster than the average 24 were projected to do so, and, of the 12 fastest growing industries 6 were correctly designated. Further, 53 of the 76 industries with negative growth were projected to do so while of the 10 industries with the largest job declines the BLS projected 3. In terms of absolute job growth, nine of the 10 largest growers were so projected.

In general, even though the economy performed between 1976 to 1988 in way more similar to 1988 to 2000 than was projected when the projections of employment for the individual industries are evaluated the projections are closer to reality than would be an extension of the earlier period growth rates. This illustrates that the extensive study of subcategories of demand and industry trends carried out in the projection process gives much better results.

Table 1.
Sources of Demand Growth
Actual and projected
1988 to 2000

	2000 Actual	2000 Proj.	1988-2000 Actual growth	1988-2000 Proj. growth
Working Age Population (millions)	209.7	204.6	0.9	0.8
Participation Rate (%)	67	69		
Labor Force (millions)	140.9	141.1	1.2	1.2
Unemployment Rate (%)	4.0	*5.5*		
Employment (Households)(millions)	135.2	133.3	1.3	1.2
Nonfarm Business Productivity (%)			1.8	*1.1*
Gross Domestic Product (Real) (%)			3.1	2.3

	1976 to 1988 Actual Growth	1988 to 2000 Projected Growth	1988 to 2000 Actual Growth
Gross Domestic Product	2.9	2.3	3.1
Personal Consumption Expenditures	3.1	2.2	3.2
Gross Private Domestic Investment	4.4	2.4	5.8
Equipment and software	5.7	3.2	9.0
Non-residential Structures	1.5	2.4	1.5
Residential Structures	2.6	2.1	2.1
Inventory Change	5.6	-4.3	8.8
Exports	5.2	4.7	7.5
Imports	6.5	2.7	8.3
Government	2.5	0.8	1.6
Federal	3.2	-0.3	-0.6
State and local	2.0	1.5	3.0

Table 3.
Jobs by Major Industrial Sector: 2000
Actual and projected
(millions)

	Actual 2000 Emp	Proj 2000 Emp	Actual '88'00 Grow	Proj '88'00 Grow	1988 %	Proj 2000 %	Act 2000 %
Total	143,787	136,225	1.7	1.2	100.0	100.0	100.0
Agriculture, forestry, and fisheries	3,526	3,247	0.4	0.0	2.8	2.4	2.5
Mining	559	717	-2.3	-0.4	0.6	0.5	0.4
Construction	8,296	7,388	2.0	1.0	5.6	5.4	5.8
Manufacturing	18,820	19,469	-0.4	-0.1	16.7	14.3	13.1
Durables	11,361	11,450	-0.2	-0.2	9.8	8.4	7.9
Nondurables	7,460	8,019	-0.7	-0.1	6.9	5.9	5.2
Transportation, communications, and utilities	7,422	6,463	2.0	0.8	5.0	4.7	5.2
Wholesale trade	7,305	7,311	1.1	1.1	5.4	5.4	5.1
Retail trade, including eating and drinking places	24,560	24,713	1.5	1.5	17.5	18.1	17.1
Finance, insurance, and real estate	8,257	8,316	1.1	1.1	6.2	6.1	5.7
Services	44,362	39,596	3.3	2.4	25.4	29.1	30.9
Government	20,680	19,005	1.5	0.8	14.8	14.0	14.4

Table 4.
Projected Job Contribution by Individual Factor: 2000
(millions)

	Actual 2000 Emp	Projected GDP	Projected FD Categ	Projected BOG	Projected I/O Rel	Projected Emp/Out	Proj 2000 Emp
Total	143,787	121,179	143,983	148,611	149,364	170,066	136,225
Agriculture, forestry, and fisheries	3,526	2,972	3,679	3,344	3,293	3,943	3,248
Mining	559	471	665	401	587	780	717
Construction	8,296	6,991	8,999	8,403	9,361	7,101	7,368
Manufacturing	18,820	15,861	19,102	19,395	10,129	34,240	19,469
Durables	11,361	9,575	10,936	12,564	10,106	26,561	11,450
Nondurables	7,460	6,287	8,167	6,830	9,023	7,679	8,019
Transportation, communications, and utilities	7,422	6,255	7,261	7,267	7,828	8,228	6,463
Wholesale trade	7,305	6,156	6,585	7,790	7,551	8,767	7,311
Retail trade, including eating and drinking places	24,560	20,699	24,526	25,195	26,608	27,168	24,731
Finance, insurance, and real estate	8,257	6,959	8,267	8,602	7,802	12,468	8,316
Services	44,362	37,387	43,226	47,052	46,475	46,368	39,596
Government	20,680	17,428	21,673	21,162	20,731	21,003	19,005

Table 5.
Jobs by Industry: 2000
Actual and projected
(millions)

	Actual 2000 Emp	Projected 2000 Emp	Actual 1988-2000 Growth	Projected 1988-2000 Growth	1976-1988 Growth
Totals	143,787	136,225	1.7	1.2	2.3
1. Agricultural production	1,979	1,896	-1.2	-1.6	-1.8
2. Ag services, forestry, hunting, & trapping	1,548	1,352	3.2	2.0	6.2
3. Metal mining	43	49	-1.6	-0.4	-5.0
4. Coal mining	77	123	-5.5	-1.7	**-3.3**
5. Crude petroleum, natural gas, and gas liquids	135	177	-3.7	-1.5	1.8
6. Oil and gas field services	189	247	-0.9	1.4	**0.6**
7. Nonmetallic minerals, except fuels	115	120	-0.1	0.3	**-0.2**
8. Construction	8,296	7,388	2.0	1.0	3.2
9. Logging	120	113	-0.4	-0.8	-0.8
10. Sawmills and planing mills	189	174	-0.9	-1.5	**-0.7**
11. Millwork, plywood, and structural members	344	306	1.7	0.7	3.0
12. Wood containers and misc wood products	154	140	0.4	-0.3	**0.7**
13. Wood buildings and mobile homes	92	77	2.4	0.9	-0.3
14. Household furniture	305	337	-0.2	0.6	**0.2**
15. Partitions and fixtures	94	101	1.2	1.8	3.2
16. Office and misc furniture and fixtures	181	177	1.6	1.4	4.5
17. Glass and glass products	150	138	-0.6	-1.4	-1.5
18. Hydraulic cement	18	16	-1.0	-1.8	-3.7
19. Stone, clay, and misc mineral products	186	172	0.0	-0.6	-0.9
20. Concrete, gypsum, & plaster products	248	237	1.0	0.6	**1.0**
21. Blast furnaces and basic steel products	225	243	-1.7	-1.1	-5.5
22. Iron and steel foundries	123	123	-0.9	-0.9	-4.1
23. Primary nonferrous smelting & refining	37	35	-1.6	-1.9	-3.4
24. All other primary metals	45	43	0.0	-0.2	1.1
25. Nonferrous rolling and drawing	176	160	-0.1	-0.9	**-0.7**
26. Nonferrous foundries	94	97	0.6	0.8	**0.5**
27. Metal cans and shipping containers	36	45	-3.3	-1.5	**-3.1**
28. Cutlery, hand tools, and hardware	134	122	-0.3	-1.1	-1.5
29. Plumbing and nonelectric heating equipment	60	60	-0.5	-0.4	-0.3
30. Fabricated structural metal products	498	407	1.2	-0.5	-0.6
31. Screw machine products, bolts, rivets, etc	116	100	1.2	0.0	**0.4**
32. Metal forgings and stampings	255	208	0.9	-0.8	-1.4
33. Metal coating, engraving,and allied services	146	130	1.7	0.7	**2.3**
34. Ordnance and ammunition	39	67	-5.6	-1.3	2.5
35. Miscellaneous fabricated metal products	281	228	1.6	-0.1	-0.2
36. Engines and turbines	85	76	-0.7	-1.6	-2.2
37. Farm and garden machinery and equipment	96	106	-0.7	0.1	-3.7
38. Construction and related machinery	241	220	0.6	-0.1	-3.3
39. Metalworking machinery and equipment	331	318	0.0	-0.4	**-0.1**
40. Special industry machinery	175	155	0.5	-0.5	-0.6

41. General industrial machinery and equipment	256	248	0.5	0.2	-0.9
42. Computer and office equipment	363	483	-1.9	0.4	4.2
43. Refrigeration and service industry machinery	217	196	0.9	0.1	**1.2**
44. Industrial machinery, nec	380	298	1.6	-0.5	**1.4**
45. Electric distribution equipment	85	97	-1.6	-0.5	**-0.6**
46. Electrical industrial apparatus	151	165	-1.4	-0.7	**-1.5**
47. Household appliances	116	116	-1.4	-1.4	-1.6
48. Electric lighting and wiring equipment	184	188	-0.6	-0.5	0.1
49. Household audio and video equipment	80	70	-0.5	-1.6	-2.9
50. Communications equipment	276	266	0.0	-0.3	2.5
51. Electronic components and accessories	685	663	0.8	0.5	4.8
52. Miscellaneous electrical equipment	147	161	-1.3	-0.5	2.1
53. Motor vehicles and equipment	1,018	789	1.4	-0.7	**-0.2**
54. Aerospace	553	944	-3.9	0.3	4.0
55. Ship and boat building and repairing	168	177	-1.5	-1.1	-0.7
56. Railroad equipment	36	31	1.3	0.2	-4.0
57. Miscellaneous transportation equipment	86	52	2.6	-1.7	-2.0
58. Search and navigation equipment	155	418	-5.8	2.4	**2.0**
59. Measuring and controlling devices	302	341	-0.6	0.4	2.3
60. Medical equipment, instruments, and supplies	290	295	1.8	1.9	3.7
61. Ophthalmic goods	33	42	-1.4	0.5	**0.1**
62. Photographic equipment and supplies	70	88	-3.6	-1.8	-1.1
63. Watches, clocks, and parts	5	8	-6.4	-2.9	**-7.7**
64. Jewelry, silverware, and plated ware	62	64	-0.5	-0.3	0.1
65. Toys and sporting goods	109	101	-0.3	-1.0	-1.0
66. Manufactured products, nec	262	217	0.4	-1.2	**-0.4**
67. Meat products	508	417	1.9	0.3	**1.2**
68. Dairy products	146	156	-0.9	-0.3	-1.7
69. Preserved fruits and vegetables	221	240	-0.6	0.1	**-0.2**
70. Grain mill products and fats and oils	152	143	-0.3	-0.8	-1.2
71. Bakery products	209	189	-0.5	-1.3	**-1.0**
72. Sugar and confectionery products	93	84	-0.6	-1.4	**-1.2**
73. Beverages	188	170	-0.5	-1.4	**-0.9**
74. Miscellaneous food and kindred products	180	178	0.4	0.3	0.9
75. Tobacco products	34	39	-3.9	-2.8	-2.8
76. Weaving, finishing, yarn, and thread mills	289	352	-2.8	-1.2	**-2.7**
77. Knitting mills	128	179	-4.3	-1.5	-0.7
78. Carpets and rugs	66	60	0.5	-0.2	**0.5**
79. Miscellaneous textile goods	54	41	0.0	-2.2	**-1.7**
80. Apparel	429	742	-6.0	-1.6	**-2.1**
81. Miscellaneous fabricated textile products	223	188	0.6	-0.9	**1.1**
82. Pulp, paper, and paperboard mills	199	223	-1.7	-0.8	-0.5
83. Paperboard containers and boxes	219	210	0.5	0.1	0.0
84. Converted paper products except containers	241	255	0.0	0.5	1.5
85. Newspapers	461	546	-0.4	1.0	1.7
86. Periodicals	156	154	1.0	0.9	5.2
87. Books	136	137	0.5	0.6	2.0
88. Miscellaneous publishing	100	119	1.3	2.8	5.8
89. Commercial printing and business forms	632	676	-0.1	0.5	3.8
90. Greeting cards	27	27	0.2	0.3	1.3
91. Blankbooks and bookbinding	63	93	-2.1	1.2	2.7
92. Service industries for the printing trade	50	77	-2.4	1.2	3.6

93. Industrial chemicals	218	268	-2.0	-0.4	**-0.9**
94. Plastics materials and synthetics	155	168	-1.1	-0.4	**-1.7**
95. Drugs	315	262	2.7	1.2	**2.4**
96. Soap, cleaners, and toilet goods	156	170	-0.2	0.5	1.9
97. Paints and allied products	52	63	-1.7	0.0	**-0.3**
98. Agricultural chemicals	51	50	-0.1	-0.3	-2.2
99. Miscellaneous chemical products	93	108	-0.6	0.5	1.7
100. Petroleum refining	85	105	-2.9	-1.2	**-2.1**
101. Miscellaneous petroleum and coal products	43	34	0.7	-1.2	**-0.5**
102. Tires and inner tubes	79	65	-0.4	-2.0	**-1.8**
103. Rubber products, plastic hose and footwear	189	158	0.6	-0.9	**-0.7**
104. Miscellaneous plastics products, nec	746	759	1.6	1.8	4.0
105. Footwear, except rubber and plastic	31	70	-8.5	-2.1	**-5.4**
106. Luggage, handbags, and leather products,nec	46	45	-1.9	-2.0	-4.2
107. Railroad transportation	236	230	-1.9	-2.1	-4.8
108. Local and interurban passenger transit	523	377	2.9	0.2	**1.8**
109. Trucking, non-air courier, & warehousing svcs	2,145	1,708	2.5	-0.5	**1.6**
110. Water transportation	207	162	1.3	-0.7	-1.0
111. Air transportation	1,291	1,106	3.5	2.2	7.0
112. Pipelines, except natural gas	14	18	-2.5	0.0	0.4
113. Passenger transportation arrangement	234	274	1.9	3.3	7.9
114. Miscellaneous transportation services	261	194	5.1	2.6	**4.7**
115. Communications exc radio & TV broadcasting	1,394	1,084	2.3	0.2	**0.3**
116. Radio and TV broadcasting	262	263	1.1	1.1	3.4
117. Electric utilities	480	634	-2.0	0.4	2.1
118. Gas utilities	158	206	-2.0	0.1	0.3
119. Water and sanitation	219	207	4.6	4.1	5.5
120. Wholesale trade	7,305	7,311	1.1	1.1	2.4
121. Retail trade exc eating and drinking places	16,214	16,758	1.2	1.5	2.1
122. Eating and drinking places	8,347	7,955	2.0	1.6	4.3
123. Depository institutions	2,032	2,450	-0.9	0.7	3.1
124. Nondepository;holding & investment offices	954	741	4.3	2.1	**6.0**
125. Security and commodity brokers	874	726	4.5	2.9	8.2
126. Insurance carriers	1,589	1,587	0.9	0.8	2.2
127. Insurance agents, brokers, and service	907	996	1.1	1.9	4.6
128. Real estate and royalties	1,901	1,817	1.0	0.5	4.4
129. Owner-occupied dwellings	0	0	0.0	0.0	0.0
130. Hotels and other lodging places	1,968	1,988	1.7	1.8	4.2
131. Laundry, cleaning, and shoe repair	537	607	-0.5	0.5	1.2
132. Personal services, nec	413	379	3.7	2.9	6.9
133. Beauty and barber shops	841	822	1.0	0.8	2.2
134. Funeral service and crematories	107	95	1.5	0.5	**1.4**
135. Advertising	330	362	1.9	2.7	5.5
136. Services to buildings	1,176	1,216	1.4	1.7	6.8
137. Miscellaneous equipment rental and leasing	319	279	3.6	2.4	8.9
138. Personnel supply services	3,918	2,296	9.0	4.2	13.7
139. Computer and data processing services	2,259	1,255	9.8	4.6	**13.2**
140. Miscellaneous business services	2,651	2,310	3.9	2.7	6.0
141. Automotive rentals, without drivers	228	283	0.8	2.7	5.9
142. Automobile parking, repair, and services	1,302	1,286	2.3	2.2	4.5
143. Electrical repair shops	126	186	-1.2	2.1	3.0
144. Watch, jewelry, & furniture repair	64	75	-0.8	3.1	**0.1**

145. Miscellaneous repair services	357	396	-0.1	0.7	2.9
146. Motion pictures and video tape rental	632	415	4.2	0.6	**3.7**
147. Producers, orchestras, and entertainers	312	277	2.8	1.8	5.6
148. Bowling centers	82	89	-1.4	-0.7	-0.7
149. Commercial sports	163	105	4.3	0.6	**2.0**
150. Amusement and recreation services, nec	1,406	954	5.7	2.3	**3.7**
151. Offices of health practitioners	3,356	3,303	3.4	3.3	4.8
152. Nursing and personal care facilities	1,807	1,914	2.6	3.1	4.1
153. Hospitals, private	4,000	4,244	1.6	2.1	2.8
154. Health services, nec	1,315	1,071	6.5	4.7	8.4
155. Legal services	1,199	1,334	0.9	1.8	6.0
156. Educational services	2,441	1,928	3.2	1.2	3.3
157. Individual & miscellaneous social services	1,029	839	4.9	3.1	6.3
158. Job training and related services	381	276	3.8	1.1	6.8
159. Child day care services	1,199	1,317	3.7	4.5	4.6
160. Residential care	827	725	4.9	3.8	8.9
161. Museums and membership organizations	2,584	2,051	2.8	0.9	**1.6**
162. Engineering and architectural services	1,098	1,044	2.7	2.2	5.3
163. Research & test svcs, mgmt & pub relations	1,990	1,624	4.8	3.0	6.3
164. Accounting, auditing, and other services	1,052	1,145	1.9	2.6	5.8
165. Private households	894	1,103	-2.2	-0.4	**-1.5**
166. US Postal Service	860	878	0.3	0.5	1.8
167. Federal electric utilities	27	35	-2.6	-0.7	1.1
168. Federal government enterprises, nec	88	151	-4.7	-0.3	0.0
169. Federal general government	1,802	1,996	-0.6	0.2	0.3
170. Local government passenger transit	223	220	0.9	0.7	4.2
171. State and local electric utilities	90	90	0.8	0.8	2.6
172. State and local government enterprises, nec	557	645	-0.6	0.6	1.6
173. State and local government hospitals	970	1,146	-0.8	0.6	**0.4**
174. State and local government education	9,472	8,289	2.1	1.0	**1.3**
175. State and local general government, nec	6,592	5,555	2.1	0.7	**1.7**
176. Noncomparable imports	0	0	0.0	0.0	0.0
177. Scrap, used and secondhand goods	0	0	0.0	0.0	0.0
178. Rest of the world industry	0	0	0.0	0.0	0.0
179. IVA & government capital services	0	0	0.0	0.0	0.0

This page intentionally left blank.

Evaluating the 2000 Occupational Employment Projections

Andrew D. Alpert
Jill N. Auyer

Bureau of Labor Statistics
Office of Occupational Statistics and Employment Projections

Andrew D. Alpert
Phone: (202) 691-5754
Email: Alpert_A@bls.gov

Jill N. Auyer
Phone: (202) 691-5806
Email: Auyer_J@bls.gov

Background

- Evaluating projections made for the 1988 to 2000 period

    - November, 1989 Monthly Labor Review
    - 1990-91 Occupational Outlook Handbook

- Purpose of evaluation

    - How accurate were we?
    - Improve projections process

The final phase of any projections process is the evaluation of the projections once actual data for the target year become available. The Office of Occupational Statistics and Employment Projections periodically evaluates the results of past projections in order to gauge how accurately the projections tracked actual employment growth. This study examines the accuracy of the 1988-2000 occupational employment projections.

The occupational employment projections for the year 2000 were developed in 1989 using 1988 as the base year. The projections were published in the November, 1989 *Monthly Labor Review* (MLR) and the 1990-91 *Occupational Outlook Handbook* (OOH).

The purpose of this evaluation is not only to verify how accurate our projections turned out to be when compared to the actual target year data, but also to improve the projections process in the development of future editions of the *Handbook.*

It is important to keep in mind that, at this time here the information presented in this slide show is preliminary. The final results of our evaluation will eventually be published in an article in the *Monthly Labor Review.*

What do we still have to do?

Identification of sources of error and evaluating job clusters has just begun. We will continue to analyze occupations, identify sources of error, and evaluate our assumptions and judgements.

We have yet to determine whether occupations or occupational groups are comparable between 1998 and 2000 due to changes in the Standard Occupational Classification system.

We also want to look more closely at how the 1988-2000 projections compare with past projections to see how accurate we have been over time.

Comparability Problems

- The Standard Occupational Classification
  System (SOC) revised for the year 2000

- Classification changes created
  comparability problems between the
  projected 2000 and the actual 2000
  employment data

The first major difficulty we encountered when undertaking this project was that the Standard Occupational Classification System (SOC) underwent a major revision for the year 2000.

The titles and content of the major occupational groups and many detailed occupations in the 2000 SOC are now substantially different. Some major groups have been renamed, combined or reorganized. Some individual occupations were renamed or reclassified into different major groups. Many new occupations were added. Some were aggregated and some were split into more detail.

Because of these changes, the occupations and major groups reflected in our 2000-2010 national employment matrix are not comparable to those reflected in the 1988-2000 employment matrix the 2000 projections.

Note: This same problem will also occur when the projections for 2005 and 2006 are evaluated.

New Matrix

- Created to circumvent SOC comparability problems

- Maintained original 1988 and projected 2000 employment data reconfigured to the 1998 occupational structure

- Actual 2000 data created by applying 1998 staffing patterns to 2000 industry totals

---

An industry-occupation matrix is used to project employment for wage and salary workers. The matrix shows occupational staffing patterns--each occupation as a percent of the work force in every industry. Data for current or actual staffing patterns in the matrix come primarily from the Bureau's Occupational Employment Statistics survey (OES).

The occupational staffing patterns for each industry are projected based on anticipated changes in the way goods and services are produced, then applied to projected industry employment, and the resulting employment summed across industries to get total wage and salary employment by occupation.

Because of revisions to the SOC, we had to find a way to work around the lack of comparability across occupations and major groups between 1988 and 2000. Actual employment data for 2000 was recreated for purposes of this evaluation by applying the 1998 staffing patterns to the 2000 industry totals. The original 1988 and projected 2000 employment data published in 1989 were reconfigured to the 1998 occupational structure.

Occupations Evaluated

- Original 1988-2000 projections were made for 9 major groups and nearly 500 detailed occupations

- For evaluation only 344 detailed occupations were analyzed

- Occupations that were eliminated
  - Residual occupations
  - Occupations where definitions not consistent between 1988 and 1998
  - Occupations with fewer than 25,000 employees in 1988

Employment projections for 1988-2000 were developed for 9 major occupational groups and of nearly 500 detailed occupations. In this evaluation, we are analyzing the projections for all 9 major groups but only 344 detailed occupations. Occupations were eliminated from our evaluation if they were residual occupations, their definitions were not consistent between 1988 and 1998, or they employed fewer than 25,000 workers in 1988. The list of occupations was confined to only those employing more than 25,000 workers in 1988 at this point, because only those were published in the original 1989 *Monthly Labor Review* article.

Historical Context

Projection period	# occupations evaluated	Average percent error	% of occupations with below average error
1964-75	76	20.8	66
1970-80	64	22.4	50
1980-90	132	21.1	60
1984-95	348	24.0	60
1988-2000	344	26.5	69

Source: Bureau of Labor Statistics

In addition to analyzing how the projections compared to the actual target year data, we also looked at how accurate the projections were in comparison to previous sets of projections. The information presented in this table appeared in an article by Neal Rosenthal on the quality of BLS projections over time published in the May 1999 *Monthly Labor Review*.

One traditional measure is to compare projected employment with the actual employment and compute the difference in percentage terms. For the 1988-2000 period, the percent error for all occupations averaged 26.5 percent. This was higher than the average percent error from the previous evaluation, but still consistent with historical errors. However, 69 percent of occupations actually had below average errors, which is higher than other projection periods.

It should be noted that the 1964-75 projections, released in 1966, were the first time an industry-occupation matrix was used. Projections were only developed for 162 occupations. Since 1966, projections have been developed and published every other year. Projections from the 1950s through 1994 were all made to target years ending in 0 or 5. Projections since 1996 have a 10 year span.

Changes in occupational classification and changes in survey methodology that occurred between the time the projections were developed and the target year limited the number of occupations that could be evaluated each time. For example, this evaluation includes 344 of almost 500 occupations, whereas the 1980-90 projections evaluation covered only 132 of 687 occupations.

Major Occupational Groups

Occupational Group	Percent change, projected	Percent change, actual	Numerical error	Absolute percent error
Total, all occupations	15.3	21.7	-7,575	5.3
Executive, administrative, and managerial	22.0	21.9	11	0.1
Professional specialty	24.0	39.8	-2,290	11.2
Technicians and related	31.3	32.7	-56	1.1
Marketing and sales	20.0	27.9	-950	6.1
Administrative support occupations, including clerical	11.9	15.8	-866	3.4
Service	22.6	25.3	-509	2.2
Agriculture, forestry, and fishing, and related	-4.8	14.1	-664	16.6
Precision production, craft, and repair	10.0	11.1	-156	1.0
Operators, fabricators, and laborers	1.1	13.6	-2,096	11.0

Source: Bureau of Labor Statistics   (Numbers in thousands)

This table depicts the preliminary results of our evaluation of the major occupational groups.

Overall employment grew faster than projected by 21.7 percent instead of 15.3 percent. Though the direction of the employment change was anticipated correctly, 8 of 9 groups were under-projected, pointing to the continued conservative nature of our projections.

All but 3 had absolute errors of less than 10 percent Significant errors in the projections for detailed occupations with sizable employment can have substantial impact on the overall projections for their respective groups when aggregated. We are in the process of examining the reasons for these large errors. For example, the error for professional specialty occupations appears to be due mostly to the under-projection of computer-related, engineering, and educational occupations. The error for operators, fabricators, and laborers seems to stem from the under-projection of hand workers, including many highly skilled precision assemblers and helpers, underestimated by about 1 million workers.

Agriculture, forestry, and fishing, and related occupations was the only group that moved in the opposite direction from what was projected. Employment in this group was projected to decline slightly but it actually grew by 14 percent. Though we are still looking into this, the reason may be attributable to changes in the occupational definitions within this group or the addition of new occupations to the Occupational Employment Statistics survey (OES) over the time period.

Detailed Occupations

Summary of findings

- Majority of occupations grew or declined
   in the projected direction

- Majority of projections were
   conservative

- Majority of projection errors
   predominantly due to staffing patterns
   rather than industry projections

A little over two-thirds of the occupations grew or declined in the projected direction, which is consistent with past evaluations.

For those occupations that grew or declined in the projected direction, a little over two-thirds of those occupations were under-projected, suggesting that our projections err on the conservative side. It also suggests that assumptions are correct, just underestimated.

A little over half of all error can be attributed back to incorrectly projected changes in staffing patterns. For about a quarter of the occupations the error can be traced back to projections of industry employment and roughly a quarter of all occupations errors can be attributed equally to both incorrect staffing pattern changes and industry projections.

<div style="border:1px solid black;">

Sources of Error

- Simulated matrices

    – Actual industry totals to projected staffing patterns

    – Actual staffing pattern to projected industry totals

- Job cluster examples

</div>

Errors in the projections for individual occupations can ultimately be traced back to errors in assumptions or judgements, resulting in incorrectly projected changes in staffing patterns, industry projections, or a combination of both.

To determine whether errors are due to staffing patterns or industry projection errors, we created two simulated matrices. The first simulation was generated by multiplying the projected 2000 staffing patterns of industries by the actual 2000 industry employment numbers. This simulation reveals the outcome if our office had projected perfect industry employment. This simulation tests for errors due to staffing patterns.

The second simulation was generated by multiplying the actual 2000 staffing patterns by the 2000 projected industry totals. This simulation reveals the outcome if our office had projected perfect staffing patterns. This simulation tests for errors due to industry projections.

The employment numbers created by each simulated matrix are then each compared with the actual 2000 employment numbers and an absolute percent error is generated for each occupation from both simulated matrices. Whichever absolute percent error is higher indicates whether the projection error was more attributable to staffing pattern or industry error.

Once the source of error has been identified for an occupation or group of occupations, we analyze assumptions and judgements made back when the projections were developed and determine which factors affecting employment that we correctly identified and which assumptions most likely contributed to the error in the projection.

Because it is time consuming to investigate the sources of error for all 344 occupations, we have begun to look at groups of related occupations that had similar sources of error which can highlight the largest factors affecting the projection error.

The next few slides give some examples of the job clusters that we have analyzed to date.

Example of error due to staffing patterns:

Healthcare occupations
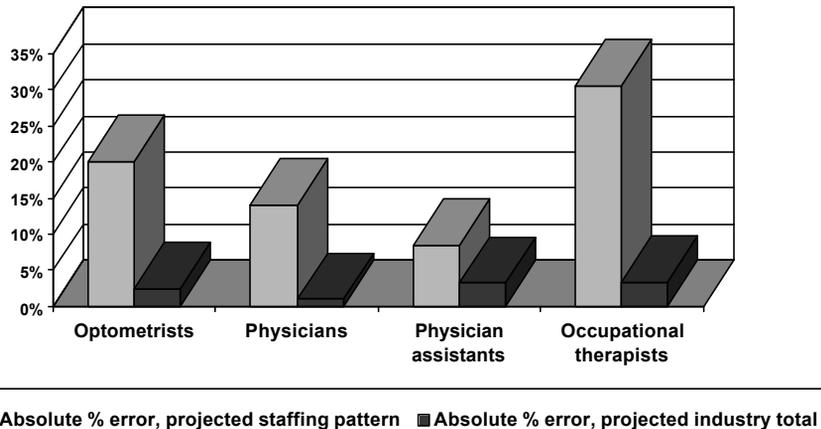
Source: Bureau of Labor Statistics

This is an example where projection error occurs predominantly because of errors in projecting the staffing pattern.

The two occupations on the left, optometrists and physicians, are classified as health diagnosing occupations. All of the health diagnosing occupations had similar patterns of error. All health diagnosing occupations were over-projected. For example, physicians were projected to grow 28 percent and actually grew 10.5 percent. Optometrists were projected to grow 16 percent and actually showed no employment growth.

The two occupations on the right are classified as either health technicians or heath assessment occupations. As with the health diagnosing occupations, there are many more occupations in these categories exhibiting similar error patterns. Most of these occupations were under-projected over the period. For example, physician assistants were projected to grow 30 percent and actually grew 44 percent. Occupational therapists were projected to grow 45 percent and actually grew 118 percent.

Looking at the underlying assumptions or judgements that went into determining changes in staffing patterns for these health care occupations, the health diagnosing occupations were projected to decline in offices of health practitioners because of an increase in large group practices which often require a higher proportion of support staff.

The underlying assumptions or judgements for the health assessment and technician occupations revealed increases in most industries due to an increase in outpatient services and a shifting of responsibilities to lower skilled healthcare workers in an attempt to contain costs.

After examining the errors and the rationales for growth, we concluded that we underestimated the effects that group practices would have on the staffing patterns in doctor's offices, and also underestimated the reliance on lower skilled healthcare workers to deal with the more routine tasks. This seems to explain the over-projection of health diagnosing occupations and the under-projection of health assessment and technicians.

Example of error due to industry projections: Teaching occupations
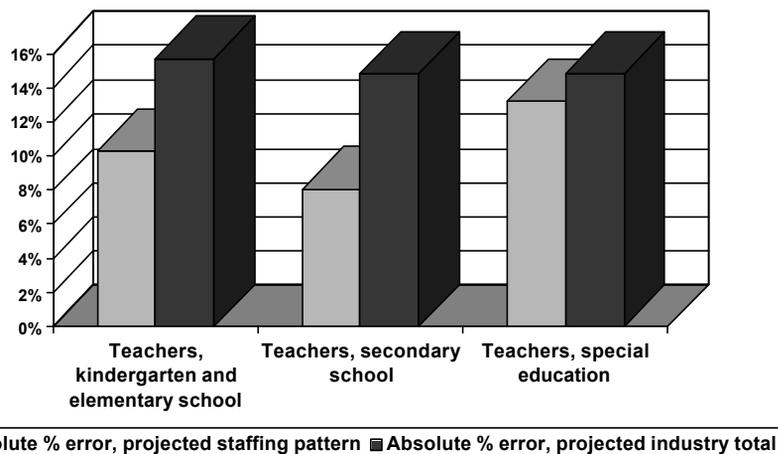
Source: Bureau of Labor Statistics

This is an example of a group of occupations in which the sources of error were predominantly due to errors in industry projections.

All of the teaching occupations were under-projected between 1988 and 2000. One of the largest under-projections for teachers was for kindergarten and elementary school teachers. Employment was projected to grow 15 percent and actually grew about 51 percent.

All of the teaching occupations are concentrated in the educational services industry, which was projected to grow at an average annual rate of 1.2 percent but actually grew at a rate of 3.2 percent. This under-projection of employment in educational services is the main cause of the under-projection in the teaching occupations.

The main assumption contributing to the growth in educational services is an increase in school enrollment, which is a reflection of population growth of youth ages 5 to 17. In 1988, the Census Bureau projected an increase of 2 million in the elementary school population by 2000. It actually increased by approximately 4.4 million. Also in 1988, the Census Bureau projected the secondary school population would increase by 1.3 million by 2000. It actually increased by 1.7 million.

The faster growth in the school-aged population in turn caused enrollment rates to rise and increased demand for teachers.

This page intentionally left blank.

NAICS Conversion Issues in Occupational Statistics

and Employment Projections

Norman C. Saunders

saunders_n@bls.gov

202-691-5707

OOSEP   **April 2002**

• The OOSEP program includes the Occupational Employment Survey (DOES) and the entire BLS projections program (Division of Industry Employment Projections & Division of Occupational Outlook).

• The conversion to a NAICS basis will impact both components of the OOSEP program but the most problematic impacts will be upon the historical time-series underlying the industry and occupational projections models and methods.

• This paper discusses how NAICS came about, defines the data issues faced by BLS forecasters, details the pros and cons of the various options open to OOSEP analysts, and delineates the path chosen for the BLS projections program to make the conversion to a NAICS basis. The assumptions required to make this conversion in a timely manner will be detailed.

• I would add that the decisions laid out here which affect the BLS projections are only one of several possible approaches, and may not be the best approach for other projection programs. A thoroughgoing analysis of the data needs of your particular program, following the approaches noted here, should suggest the best alternative solution.

NAICS -- How it came about

Economic Classification Policy Committee,
formed by OMB in 1992 to construct a new
industry classification paradigm which would:

   1) define industries according to the production processes
      they use

   2) ensure consistent classification among Canada, USA,
      and Mexico (required by NAFTA)

   3) reflect the structure of today's economy

   4) maintain flexibility and currency

OOSEP   **April 2002**

• The need for a completely new, "fresh-slate," classification system for industries in the U.S. economy had become increasingly clear over the latter half of the 1980s:

   • The basis for defining industries varied between production-related and market-oriented.

   • Of our neighboring countries, Mexico had no classification scheme, and Canada had a classification system that was not consistent with the SIC. A common classification scheme is required across all three countries by the North American Free Trade Agreement (NAFTA).

   • The SIC was biased toward the manufacturing sector, with inadequate detail for the burgeoning service-producing sectors.

   • Finally, attempts to revise the SIC were hopelessly mired in past decisions, making it very difficult to react accurately or in a timely manner to newly emerging sectors of the economy.

• Unfortunately, one very important criteria to many users of industry-based data, that of consistency over time, was placed at or near the bottom of considerations taken into account by the ECPC.

•While the NAICS is improved in many, many ways, long revised time series will be, more often than not, the exception to the rule in the conversion process.

OOSEP NAICS Capability Team

- Formed early Fall 2001 with a series of tasks:
  - inventory affected data and determine NAICS status
  - propose alternative proposals for partial & full shifts to NAICS

- Intermediate report on 30 November 2001 & Final Report on 28 February 2002

OOSEP   **April 2002**

• The team, made up of 3 economists from the industry employment side of the Office and another 2 economists from the occupational demand division, first provided an intermediate report, which detailed:

   • An inventory of all industry-based data used by OOSEP in the projections program, including the NAICS status of the data item in question, how the data item was used in the projections process (a measure of criticality), and notes on special issues necessary to consider for that particular item.

   • A proposed NAICS-based industry sectoring plan based on a "best-case" set of available data.

• The team's final report included their recommendations for moving forward with the Office conversion to NAICS, along with a discussion of the impact on our users and some discussion on means for training the staff in NAICS considerations.

The Program

- Labor force by age, sex, race, and ethnicity

- Aggregate economy

- Industry final demand, output, productivity
   and employment

- Occupational demand

OOSEP   April 2002

• The BLS projections program encompasses detailed projections of labor force by age, sex, race, and ethnicity; aggregate economic projections of the economy; industry-based projections of demand, output, employment, hours, and productivity; and detailed occupational demands within detailed industries.

•The projections are produced Biennially, published in the fall of odd-numbered years. The 2003 publication will cover the period 2002-2012 and the 2005 projections publication will cover the period 2004-2014.

• Reasonably long, consistent, time series are required for model formulation and estimation.

• If the OOSEP shifts for the 2012 projections, not all data series on which the projections depend will be available. If we wait until the 2014 projections, our 2012 projections on an SIC basis will be "orphans" for a bit over two years, since BLS industry employment data will be on a NAICS basis prior to the 2003 publication date.

The OOSEP Schedule

- 2002-2012 projections
  - Model & data preparation -- Apr-Nov 2002
  - Projections estimation -- Dec 2002-Jun 2003
  - Clearance -- Jul-Aug 2003
  - Editorial -- Sep-Oct 2003
  - Publication -- Nov 2003

- 2004-2014 projections

OOSEP   **April 2002**

• Here is a typical schedule of the various phases underlying a set of OOSEP projections.

• Even though we publish in the Fall of 2003, we obviously need good, underlying historical time series right now--in the Spring of 2002.

• The same timing issues hold true of any of our projection sets, but the 2002-2012 projections will be the most problematic because of missing (unconverted or no long historical series) data.

Data Issues

- BLS--employment, productivity, prices

- BEA--NIPA's, I-O Accounts, GDO

- Census--Annual Surveys, Economic Census,
  Trade Policy Information System

- Other Federal Agencies--Output extrapolators

OOSEP   **April 2002**

The three agencies providing the most critical data for the BLS projections program are the BLS, the BEA, and the Census Bureau.

The re-coding of the ES-202 establishments (all establishments covered by state unemployment insurance laws) from SIC to NAICS is complete, allowing a full recalculation of CES and OES data, expected in mid-2003. Both the CES and OES surveys draw their samples from the universe defined by the ES-202 establishment list. Prices and productivity by industry depend on the benchmark BEA input-output table for 1997, expected late in 2002 and into 2003.

The NIPA benchmark also depends on the availability of the BEA input-output tables for 1997. Gross duplicated output depends on BLS prices, which in turn depend on the interindustry data.

Census Annual Surveys are being carried out now on a 1997 NAICS basis. The 2002 Economic Census will shift Census to the 2002 definition. TPIS requires bridge tables from the Harmonized accounts to SITC and to NAICS.

Many other federal agencies provide data that is more or less affected by the NAICS conversion: Agriculture, Education, Transportation, Treasury, EIA, the Federal Reserve, USPS, etc.

Three Possible Solutions

- SIC in 2012, NAICS in 2014

- Hybrid SIC/NAICS in 2012, NAICS in 2014

- NAICS in 2012 and following

OOSEP   **April 2002**

• The first option would be to carry out the projections as we have in the past for the 2002-2012 set then make the shift to the NAICS in the 2004 to 2014 set--this is the initial proposal.

• The second possibility would be to carry out the projections as we have in the past, based solely on SIC-based data, then bridge the resulting industry employment data to a NAICS-based definition.

• The third proposal is to make as complete a shift as possible to NAICS for the upcoming 2012 round of projections, cleaning up any remaining problem items in the 2014 round.

The First Solution -- SIC

- Pros:
  - no change to methodology
  - familiar to our current users
  - allows more time for other data conversions to be carried out
- Cons:
  - data released on an SIC basis will become, increasingly, "orphaned" data series
  - new issues raised by NAICS will not be addressable

OOSEP   **April 2002**

• The NAICS team decided that this alternative simply would not work.

• The OOSEP is being pressured by some of our users to make the shift as soon and as completely as possible to NAICS.

• Too many critical data series will be out on a NAICS basis by the time we publish so we will look bad and, at the same time, we would not be addressing our user's needs effectively.

The Second Solution -- Hybrid

- Pros:
  - current methods remain in place, augmented by employment bridge
  - minimizes data demands on our staff
  - allows more time for other data conversions to be carried out
- Cons:
  - analysis of industry-based factors affecting employment change will be difficult, if not impossible

OOSEP   **April 2002**

• The NAICS team considered this alternative very seriously.

• The primary reason for rejecting it is that it would be very difficult to draw logical conclusions about why employment grew as it did in particular industries since the associated production and productivity data wouldn't be on the same basis as the final employment results.

• If there had been a significant subset of our users who preferred that BLS continue to provide projections of industry data on an SIC basis, then this clearly would be a winning choice, but we have been unable to find any of our users who have expressed an interest in our holding the conversion back. While many of the state occupational forecasting groups are viewing the shift to NAICS with some significant concerns about the length and accuracy of state-level NAICS-based industry data, all of the individuals with whom we have spoken claim to want the national-level forecasts switched to NAICS in as timely a manner as possible.

The Third Solution -- NAICS Now

- Pros:
  - in synchrony with majority of statistical community, better able to meet needs of our users
  - a unified analytic system increases utility and explicability of industry projections

- Cons:
  - not all source data is available, necessitating many conversion assumptions
  - what is available--some on 1997 and some on 2002 basis
  - short, if any, historical time series; "competitive" position with BEA

OOSEP   **April 2002**

• Despite the fact that a lot of critical assumptions will need to be made and well-documented, this was determined by the team, and confirmed by senior management, to be the best possible solution.

• Even by the 2004-2014 projections some industry data series will not be available on a NAICS basis. Waiting for that round of projections will just postpone the inevitable so the decision was made to go ahead with this alternative.

• The 1992 BEA Benchmark I-O table will be transformed (albeit roughly) to a NAICS basis using 1997 production rations. At the same time, industry and commodity output series will be transformed to NAICS using the same ratios. The resulting interindustry flows relationships will then be balanced with pre-existing final demand data for the 1992-2002 period, forming the critical basis for our industry employment projections.

NAICS -- Find out more

Check the following websites:

http://www.census.gov/epcd/www/naics.html

http://stats.bls.gov/opub/rtaw/chapter3.htm

OOSEP   **April 2002**

• If you haven't yet begun to look into NAICS, I would recommend the excellent site maintained by the Census Bureau. It has all the working papers of the ECPC, 1997 and 2002 definitions, SIC to NAICS bridges, and many other items pertaining to the NAICS.

• The second reference provides a general overview of the history of both industry and occupational classification methodology in the U.S. This work appeared in the 2001 edition of the Report on the American Workforce, a publication of the Bureau of Labor Statistics.

This page intentionally left blank.

Coping with Issues of Continuity in New Racial Classification

Chairs: Jorge del Pinal and Campbell Gibson, U.S. Census Bureau

Discussant: N. Clyde Tucker, Bureau of Labor Statistics

On October 30, 1997, the Office of Management and Budget (OMB) issued revised standards on race and ethnicity data. Included was the identification of five racial categories–White; Black or African American; American Indian or Alaska Native; Asian; and Native Hawaiian or Other Pacific Islander. One important change was to allow respondents to report one or more races. For the 1980 and 1990 censuses, four racial categories were used–White; Black or African American; American Indian and Alaskan Native; and Asian or Pacific Islander–and respondents were asked to report one race. Now there is a need to understand how the Census 2000 race distributions can be used to compare distributions from previous censuses.

Changing Racial Categorization: Understanding the Past to Explain the Present

Claudette E. Bennett, U.S. Census Bureau, U.S. Department of Commerce

To facilitate comparisons between Census 2000 and other surveys which instruct respondents to mark one race, and with data from the vital records system which uses census data to calculate such indicators as birth and death rates, the Census Bureau conducted a national survey, called the Census Quality Survey, summer 2001. The major objective was to produce a datafile that will improve users' ability to make comparison between Census 2000 data on race that asked for the reporting of one or more races, and data on race from other sources that asked for a single race to be reported. In this paper we provide background information on this survey and preliminary findings from Census 2000 for people who reported two or more races.

A Method to Bridge Multiple-Race Responses to Single-Race Categories for Population Denomintors Vital Events Rates

Jennifer D Parker, National Center for Health Statistics, U.S. Department of Health and Human Services

A particular concern is that data from the decennial census–which provides population denominators–were collected using the new classification; while birth and death records–the numerators for death rates and other vital statistics–will implement the change over several years. OMB suggested several simple bridge methods to assign multiple race responses. This presentation describes the NCHS methodology used to create single race denominators. Using NCHS's National Health Interview Survey (NHIS), which allows reporting of more than one race but follows up to obtain a single race, regression models were developed for each multiple race group using individual and contextual-level predictors available on both the NHIS and the Census file. Methods proposed by OMB will be compared to the NCHS denominators.

Issues and Strategies in Producing Post-2000 Population Estimates with Race Detail

Amy Symens Smith, U.S. Census Bureau, U.S. Department of Commerce

The new racial classifications have introduced several challenges in the production of population estimates. The Census Bureau population estimates are developed using the cohort-component method whereby each component of population change–births, deaths, and net migration–is estimated separately. Starting in 2001, it is necessary to have data for each component that coincide with the new race categories, as well as allowing for multiple race reporting. Before Census 2000, research on the "Two or more races" population was generally limited to estimates of population size. In this paper I report on what we've learned about the "Two or more races" population and juxtapose this with the data requirements for producing post-2000 population estimates.

This page intentionally left blank.

A METHOD TO BRIDGE MULTIPLE- RACE RESPONSES TO SINGLE- RACE CATEGORIES FOR POPULATION DENOMINTORS OF VITAL EVENT RATES

J. Parker, N. Schenker, D. Ingram, J. Weed, E. Arias, B. Hamilton, and J. Madans

National Center for Health Statistics

In 1997, the Office of Management and Budget (OMB) issued revised standards for the collection and reporting of race and ethnicity data within the federal statistical system (OMB, 1997). Among other revisions, the new standard allows individuals to choose one or more race groups when responding to federal surveys and other federal data collections. Although full implementation of the new standard is not mandated until 2003, OMB requires compliance for new data collections.

Each year, the National Center for Health Statistics publishes numerous statistical reports describing births and deaths in the United States (Martin, Hamilton, Ventura et al, 2002; Minino and Smith, 2001; Eberhardt, Ingram, Makuc et al, 2001). Among those statistics are birth and death rates, which rely on birth and death data collected through the state-based vital statistics system, and on population denominators from the Bureau of the Census. Implementation of the new standard for race data took place with the 2000 Census where 2.4% of the respondents, nearly 7 million people, reported more than one race (Census Bureau, 2001); however, states will not uniformly implement the new standards for vital records for multiple-race groups until 2003 or later. This lag in implementation means that, for several years at least, numerator and denominator data will not be compatible. This will result in potentially biased race-specific rates for the single-race groups; the extent of the bias due to differential implementation would depend on both the size of the single-race group, the size of the related multiple-race groups, and the underlying rates.

To produce rates based on more compatible numerators and denominators, the National Center for Health Statistics will create an approach specifically developed for population data provided by Census 2000. In 2000, the OMB issued Provisional Guidance on the Implementation of the 1997 Standards for Federal Data on Race and Ethnicity (OMB, 2000), available on the Internet. This Guidance document contains a detailed discussion of what the OMB calls "bridge" methods, methods to compare or combine data collected under the previous OMB-Directive 15 with data collected under the new standard. Schenker and Parker [forthcoming], in subsequent work, demonstrated the benefits of including additional covariate information when using regression models in bridging decisions.

This paper describes some preliminary models used to estimate race-specific population counts that will be comparable with vital records. These population counts will be used for birth and death rates produced by NCHS for 2000 and later years, updating rates derived from the 1990 census as appropriate. The method described herein builds on the previous work of Schenker and Parker, tailored toward the specific task of 2000 denominator estimates. Because the models are still being developed, details are not presented.

METHODS

Data source. The National Health Interview Survey (NHIS) is a continuous household survey designed to measure the health status of residents of the United States (Botman, Moore, Moriarty, and Parsons, 2000). Data from the 1997-2000 surveys were used for this work. Each year about 40,000 households are included in the sample, covering about 100,000 respondents.

The NHIS provides a unique opportunity to investigate multiple-race groups. Since 1976, the NHIS has allowed respondents to choose more than one race (OMB, 2000). As the respondent is handed a card with numbered race categories, the interviewer asks: "What is the number of the group or groups that represent your race". If a respondent selects more than one category, the interviewer then asks a follow-up question: "Which of those groups would you say best describes your race?" For bridging purposes, consistent with the 1977 OMB directive, race was collapsed to four single-race categories: white; black; American Indian,

including Alaska Natives (AIAN); and, Asian or Pacific Islander (API). Although the 1997 OMB standard separated the Asians from the Native Hawaiian and Other Pacific Islanders (NHOPI), the combined group, API, was used for the bridging models. The NHIS includes an additional category, Other Race, for respondents who mentioned a race group off the standard list.

For this study, multiple-race respondents were identified from responses to the first question, which allowed respondents to choose more than one race group. The detailed multiple race responses are not included on public use data files of the NHIS and is not used directly for national estimates from the survey. Responses to the follow-up question were used to create Primary Race, that is, the "best" single-race category for a multiple-race respondent.

For this analysis, if a multiple-race response included Other Race, the Other Race response was dropped. For example, respondents who reported black and Other Race were included in the single-race black group; respondents who reported AIAN, API, and Other Race were included in the AIAN and API group. Multiple-race respondents who did not report a single-race group in the follow-up question were not included in the regression models; however, since these respondents represent a real and measurable part of each multiple-race group that would be bridged to a single-race using this approach, they were included in some examinations of the model results.

In these four combined years of the NHIS from 1997 to 2000, 4,898 respondents reported more than one race. Corresponding to the four single-race groups to which we are bridging, there are eleven multiple-race groups: AIAN/API; AIAN/black; AIAN/white; API/black; API/white; black/white; AIAN/API/black; AIAN/API/white; AIAN/black/white; API/black/white; and, AIAN/API/black/white. Both Primary race identification and the likelihood of providing a primary race differed between race groups (Table 1).

A logistic regression models was developed for each two-race multiple-race group with more than 100 respondents: black/white, AIAN/white, API/white, black/AIAN, and, black/API. For the AIAN/black/white group, a multi-logit model, which allows more than two

responses, was fitted. To illustrate the method, the models for the three largest multiple-race groups are presented (Table 2).

While the other multiple-race groups had too few respondents to support fitting separate models to the NHIS data, they are represented in the census population. Estimates for these groups were derived from a composite multi-logit model fitted using all multiple-race respondents. The idea is that the associations between Primary Race and the covariates for the smaller race groups can be inferred approximately using the associations for the larger groups. Although the evidence suggests that a separate model would be preferable for each multiple-race group, this approach was considered reasonable, given the data constraints. With the goal of balancing race detail with an estimable model, several forms of representing the multiple-race groups in the multi-logit model were considered. The model shown here includes three indicator variables to describe the multiple race groups: not black, not AIAN, and, not API. For the multi-logit model, the coefficients for the indicator variables were constrained to zero for the corresponding Primary race outcomes (Table 3); for example, the parameter estimate for the variable "not black" was constrained to zero for the Primary race outcome level black.

Demographic factors available from county-level Census population files, as well as the NHIS, were included in the regression models: age, Hispanic origin, and sex. After considering a handful of forms for the age variable, including transformed and categorical variables, we decided to add age to each model as a continuous variable.

County of residence is available on in-house versions of NHIS data files. For each respondent, we added the region of the country (Northeast, Midwest, South, or West), a county-specific index of urbanicity (Eberhardt, Ingram, Makuc, et al 2001), and the county-specific percent race distribution, for example, the percentage of county residents who reported AIAN and the percentage who reported more than one race.

The demographic covariates were included in all models to make the models comparable, both for drawing inferences and making population predictions for specific

subgroups. Single-race population percentage variables were included in models when appropriate. Percent single-race black was included in the model for black/white respondents, for example, but not in the model for AIAN/white respondents. Percent multiple-race was included in each model; percent single-race white was not included in any model. Forms of these variables differed across models. For black/white respondents, the square of percent single-race black improved the model fit, indicating that the probability of a respondent's Primary race being black increased relatively rapidly as the percentage of county residents who reported black increased. The logarithm of percent single-race AIAN improved the fit for the AIAN/white respondents, indicating that reporting AIAN as a primary race increases slowly as the percentage of single-race AIAN in the county increases.

NHIS survey weights were used to fit models, however, no attempt has been made, so far, to control for aspects of the survey design other than the weights.

To help us examine the applicability of the regression models, the Census Bureau provided us with a Census Research File containing preliminary 2000 population counts by county, age in five-year groupings, sex, Hispanic origin, and race, including five single-race groups (AIAN, Asian, black, Native Hawaiian and other Pacific Islanders, and white) and all multiple-race combinations. For this study, we combined the counts for the Asians with those for the Native Hawaiian and other Pacific Islanders to produce the prior OMB category, API.

Using the NHIS models, we created a file of predicted probabilities. Within each county, multiple race group, age group, Hispanic origin category, and gender a separate probability was estimated. We then applied these probabilities to the multiple-race population counts in each county and demographic subgroup, apportioning the multiple-race counts to the corresponding single-race groups. Both the model predictions and the resulting single-race population distributions are being examined for consistency. Additional work to determine the effects of rounding the resulting counts at different levels of geography on resulting race-specific population estimates at each level of geography is ongoing.

RESULTS

Both the strength of the associations and their directions differed between race groups (Table 2). For example, increasing age was associated with a higher likelihood of choosing API as Primary race among the API/white respondents and associated with a decreased likelihood of choosing AIAN among the AIAN/white respondents. Similarly, living in an area with a relatively high proportion of multiple-race persons increased the likelihood of Primary Race reported as black or API among the black/white and API/white, respectively; a high proportion of multiple-race persons decreased the likelihood of AIAN as a Primary Race response among the AIAN/white.

Given that the largest multiple-race groups have the greatest influence on the estimates from the combined model, it is not surprising that the combined model shows many similar trends between Primary race and the demographic covariates (Table 3). However, there are some regional differences, perhaps attributable to clustering of smaller multiple-race groups.

Application of these models to the 175,896 county-age-Hispanic-sex-race cells in the Census Research File led to a large variation in predicted probabilities (Table 4). Although these values are still being examined for consistency and local validity, these results support the idea that bridge methods that incorporate covariate information may be better able to capture demographic variation between counties which, in turn, may influence Primary race identification.

DISCUSSION

There are likely many individual level factors that influence whether or not a single Primary race is identified among muliple-race respondents, and if so, which race is mentioned. On a population level, however, only a handful of these factors are available for prediction purposes. However, even using these basic demograpic and geographic factors, our predictions showed great variation by county. Although additional individual-level variables will likely be unavailable for estimating population denominators, additional contextual information, perhaps income levels (used in

Schenker and Parker (forthcoming)) or political factors, may improve future predictions.

REFERENCES

Botman SL, Moore TF, Moriarty CL, and Parsons VL. Design and estimation for the National Health Interview Survey, 1995-2004. National Center for Health Statistics. Vital Health Stat 2(130). 2000

Bureau of the Census. 2001. "Overview of Race and Hispanic Origin". Census 2000 Brief, Issued March 2001.

Eberhardt MS, Ingram DD, Makuc DM, et al. Urban and Rural Health Chartbook. Health United States, 2001. Hyattsville, Maryland: National Center for Health Statistics, 2001

Martin JA, Hamilton BE, Ventura SJ, Menacker F, Park MM. Births: Final data for 2000. National vital statistics reports, vol 50 no 5. Hyattsville, Maryland. National Center for Health Statistics 2002.

Minino AM, Smith BL. Deaths: Preliminary data for 2000. National vital statistics reports, vol 49 no 12, Hyattsville, Maryland. National Center for Health Statistics 2001.

Office of Management and Budget. 1997. "Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity", Federal Register 62FR58781-58790, October 30, 1997. http://www.whitehouse.gov/wh/eop/omb/html/fe dreg.html .

Office of Management and Budget. 2000. Provisional Guidance on the Implementation of the 1997 Standards for the Collection of Federal Data on Race and Ethnicity. December 15, 2000. http://www.whitehouse.gov/omb/inforeg/r&e_gu idance2000update.pdf

Schenker N, Parker J. From Single-Race Reporting to Multiple -Race Reporting: Using Imputation Methods to Bridge the Transition. Statistics in Medicine (forthcoming).

Table 1. Sample sizes and percent distribution* of primary race. Multiple-race survey respondents, NHIS 1997-2000.

		Primary Race				
Multiple-race	Sample size	AIAN	API	Black	White	No Primary Race
AIAN/API	27					
AIAN/Black	393	13.3		78.7		8.0
AIAN/White	1593	21.2			74.0	4.8
API/Black	130		33.8	51.0		15.2
API/White	1147		39.6		41.2	19.2
Black/White	1138			45.4	26.9	27.7
AIAN/AP I/Black	12					
AIAN/API/White	70	1.4	54.5		35.0	9.1
AIAN/Black/White	346	6.9		27.6	8.5	57.0
API/Black/White	38					
AIAN/API/Black/White	4					

AIAN = American Indian and Alaska Native; API = Asian and Pacific Islander
* Primary race distribution not provided for multiple-race groups with fewer than 50 respondents.

Table 2. Direction of model coefficients logistic regression models of primary race for selected multiple-race groups, NHIS 1997-2000.

	Multiple-race group		
	AIAN/white	API/white	Black/white
Primary response predicted	AIAN	API	Black
Age, continuous years	-	+	+
Hispanic origin, yes	+	+	-
Sex, male	+	+	+
Region			
Northeast	+	-	-
Midwest	+	-	+
South	-	+	-
West	ref	ref	ref
Urban to rural			
Most urban	ref	ref	ref
Mostly urban	-	+	-
Somewhat rural	-	+	+
Rural	-	-	+
% AIAN population in county, 2000 (log)	+	N/A	N/A
% API population in county, 2000	N/A	+	N/A
% Black population in county, 2000 (squared)	N/A	N/A	+
% Multiple-race in county, 2000	-	+	+

AIAN = American Indian and Alaska Native; API = Asian and Pacific Islander

Table 3. Direction of model coefficients for combined multi-logit model for primary race, NHIS 1997-2000.

	Primary race response		
Covariates	AIAN	API	Black
Race			
Not AIAN	#	+	+
Not API	+	#	+
Not black	+	+	#
Age, continuous years	-	+	-
Hispanic origin, yes	+	+	-
Sex, male	+	+	-
Region			
Northeast	-	-	+
Midwest	-	-	+
South	-	-	-
West	ref	ref	ref
Urban to rural			
Most urban	ref	ref	ref
Mostly urban	-	+	+
Somewhat rural	-	-	-
Rural	+	-	-
% AIAN population in county, 2000	+	+	-
% API population in county, 2000	+	+	+
% Black population in county, 2000	+	+	+
% Multiple-race in county, 2000	-	+	-

coefficient constrained to zero.

AIAN = American Indian and Alaska Native; API = Asian and Pacific Islander

Table 4: Distributional characteristics of predicted probabilities from NHIS models applied to Census Research File of population counts by county, age, sex, and Hispanic origin, for selected multiple-race groups.

Race group	Predicted race	Mean (SD)	Median	Interquartile range	Minimum - Maximum
AIAN/White	AIAN	.25 (.19)	.22	.14-.32	.00-.92
AIAN/Black	Black	.65 (.24)	.69	.46-.86	.01-1.00
AIAN/Black/White	AIAN	.35 (.26)	.34	.11-.56	.00-.98
AIAN/Black/White	Black	.45 (.27)	.41	.21-.68	.01-.99
API/Black	API	.54 (.31)	.51	.24-.92	.01-1.00
API/White	API	.26 (.10)	.23	.20-.31	.13-.87
Black/White	Black	.62 (.14)	.62	.53-.72	.21-1.00

AIAN = American Indian and Alaska Native; API = Asian and Pacific Islander

ISSUES AND STRATEGIES IN PRODUCING
POST-2000 POPULATION ESTIMATES WITH RACE DETAIL
Amy Symens Smith, U.S. Census Bureau

Outline of the Issues

•Federal agencies are not mandated to adopt the 1997 Office of Management and Budget race and ethnicity standards until January 2003.

•Population estimates depend on administrative records sources supplied by Federal agencies (National Center for Health Statistics and Immigration and Naturalization Service).

•The 4/1/2000 base population has new OMB race categories.

National Estimates Challenge

Produce Population Estimates:

•consistent with the new race standards

–White; Black; American Indian and Alaska Native; Asian; Native Hawaiian and Other Pacific Islander
–Two or more races

•with administrative records data by old race categories

–Mark one…White; Black; American Indian, Eskimo and Aleut; Asian and Pacific Islander

National Estimates Methodology

•Cohort Component

$P_1 = P_0 + B - D + NIM + NCM$

where:

P_1 = Population at time 1
P_0 = Population at time 2
B = Births
D = Deaths
NIM = Net International Migration
NCM = Net Civilian Migration

•Little is known about the fertility, mortality, and migration patterns of the Two or more races population.

•Thus, several strategies have to be used to model race.

Relevant Issues

•The Two or more races population varied in important ways from their single race counterparts.

•The Two or more races population is not homogeneous.

National Estimates Strategies

Base Population

Modification was necessary to reconcile the Census 2000 race data that include the "Some other race" category with the racial categories that appear in administrative records (i.e. birth certificate data and immigration data).

Modified Race Data

•No modification necessary for OMB race alone or in combination.

   –White; Black; American Indian and Alaska Native; Asian; Native Hawaiian and Other Pacific Islander

•"Some other race" alone response blanked and race imputed.

•Response of both OMB race and "Some other race", "Some other race" blanked and OMB race(s) maintained.

Birth Certificate Data

•Use modeling to estimate full race distribution.

•Race modeling based on the proportion of the age zero population in a specific race/Hispanic origin group in Census 2000.

Death Certificate Data

•Use modeling to estimate full race distribution.

•Modeling using death rates
       –Rates varied for the race alone groups.
       –Constant rate for the multiple race groups.

International Migration Data

•Use modeling to estimate full race distribution.

•Modeling based on the proportion of the population in a specific race/Hispanic origin group by single years of age, sex, and country of birth in Census 2000.

Future Challenges

•Likely that not all Federal agencies will transition to the new race and ethnicity standards by 2003.

•Research currently underway to explore new modeling techniques for the future.

•Research looking at race reporting in households.

Table 1. Total Population by Number of Races Reported: 2000

Number of races	Number	Percent of total population	Percent of total Two or more races population
Total population	**281,421,906**	**100.0**	**(X)**
One race	274,595,678	97.6	(X)
Two or more races	6,826,228	2.4	100.0
Two races	6,368,075	2.3	93.3
Three races	410,285	0.1	6.0
Four races	38,408	-	0.6
Five races	8,637	-	0.1
Six races	823	-	-

 - Percentage rounds to 0.0.

(X) Not applicable

Source: The Two or More races Population: 2000. Census 2000 Brief. Nicholas A. Jones and Amy Symens Smith.

Table 2. Total Population by Number of Races Reported and Hispanic Origin: 2000

| Hispanic Origin	Total Number	Per-cent	One race Number	Per-cent	Percent of one race	Two or more races Number	Per-cent	Percent of Two or more races
Total Population	281,421,906	100	274,595,678	97.6	100.0	6,826,228	2.4	100.0
Hispanic Origin	35,305,818	100	33,081,736	93.7	12.0	2,224,082	6.3	32.6
Not Hispanic	246,116,088	100	241,513,942	98.1	88.0	4,602,146	1.9	67.4

Source: The Two or More Races Population: 2000. Census 2000 Brief. Nicholas A. Jones and Amy Symens Smith.

Table 3. Comparison of Census 2000 Race Data and Race Data After Modification

Race	Census 2000 Number	Percent	Modified Race Number	Percent
Total	281,421,906	100.0	281,421,906	100.0
Single race	274,595,678	97.6	277,524,226	98.6
White	211,460,626	75.1	228,104,485	81.1
Black or African American	34,658,190	12.3	35,704,124	12.7
American Indian and Alaska Native	2,475,956	0.9	2,663,818	0.9
Asian	10,242,998	3.6	10,589,265	3.8
Native Hawaiian and Other Pacific Islander	398,835	0.1	462,534	0.2
Some other race	15,359,073	5.5	(X)	(X)
Two or more races	6,826,228	2.4	3,897,680	1.4

(X) Not applicable.

Figure 1. Percent Under Age 18 by Number of Races Reported and Hispanic Origin
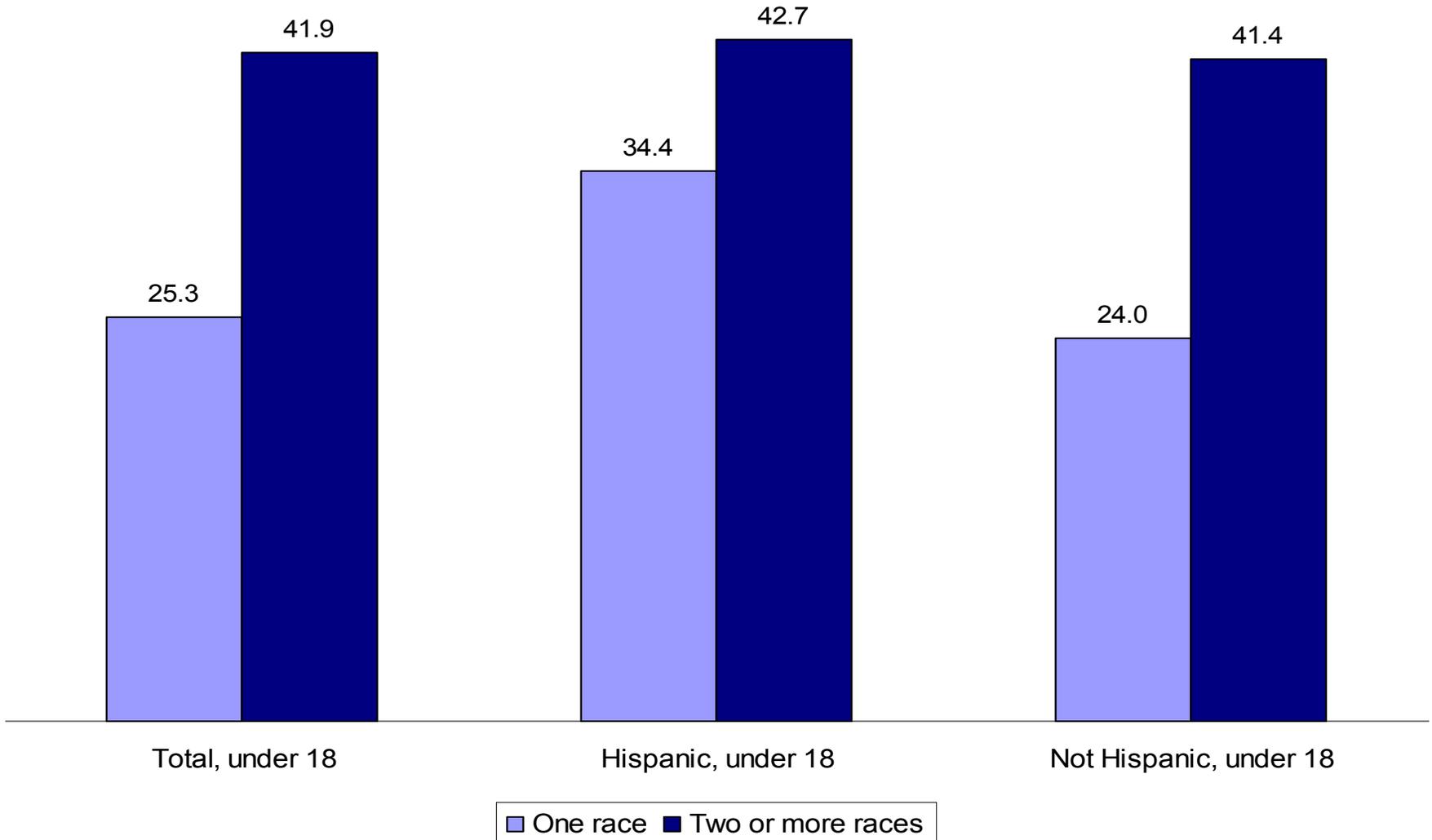
Total, under 18 — One race: 25.3, Two or more races: 41.9

Hispanic, under 18 — One race: 34.4, Two or more races: 42.7

Not Hispanic, under 18 — One race: 24.0, Two or more races: 41.4

☐ One race   ■ Two or more races

Source: The Two or More Races Population: 2000. Census 2000 Brief. Nicholas A. Jones and Amy Symens Smith

This page intentionally left blank.

Forecasting in Transportation

Chair: Peg Young, U.S. Department of Transportation

Models and Methodology of FAA Domestic Air Carrier Forecasts

Roger Schaufele, Federal Aviation Administration, U.S. Department of Transportation

The FAA uses forecasts of domestic revenue passenger miles (RPMs) and domestic passenger enplanements to provide the basis for forecasts of aviation activity. The forecasts of aviation activity are used to determine staffing levels and capital expenditures necessary to accommodate the growth of aviation activity while maintaining a safe, secure, and efficient environment. Historically, FAA forecasts of domestic air carrier traffic and revenues have been based on results of econometric models. In light of the events of September 11th, traditional methodology for developing the forecasts was not viewed as useful for developing forecasts for FY 2002 and FY 2003 and new methods were utilized. Forecasts for FY 2004 - 2013 were based on results of econometric models. This paper presents an overview of the methodology and models that were used by the FAA to forecast U.S. air carrier domestic (RPMs), domestic passenger enplanements, and domestic revenues for FY 2002 – 2013.

The Impact of Terrorism on Tourism by Use of Time Series Methods

Brian Sloboda, Bureau of Transportation Statistics, U.S. Department of Transportation

Terrorists use extra-normal violence or threaten to engage in violent acts to gain a political objective through intimidation or fear. They often unleash their attacks at targets which are not directly involve in the decision-making process that the terrorist seek to influence, i.e., harm people in a crowded street or passengers waiting at the airport. These acts have primarily occurred overseas, but the threat of terrorism finally hit the United States in 1993 with the bombing of the World Trade Center by Islamic fundamentalists. More recently, simultaneous attacks on the World Trade Center and the Pentagon has increased fears of subsequent attacks and these fears started an effect into the US economy. The purpose of this analysis is to determine if the degree of impact of these recent attacks and to quantify these impacts in terms of losses in tourism revenue. The empirical analysis will entail the use of an ARIMA model with a transfer function for the United States.

The Impact of September 11, 2001 on Transportation Indicators

Peg Young, Bureau of Transportation Statistics, U.S. Department of Transportation
Keith Ord, The McDonough School of Business, Georgetown University

The Bureau of Transportation Statistics produces a monthly report, called Transportation Indicators, which reports on key measures related to the transportation enterprise. The co-authors of this paper have created a procedure, using STAMP, to decompose the time series of interest and to create monthly forecasts of these indicators. In addition, the procedure compares the new actual values of these measures to the one-step-ahead forecasts in order to provide alerts for those measures that deviated more than expected every month. This presentation will show the results of this forecast and statistical process control procedure, particularly in light of the events on September 11, 2001 on transportation data.

This page intentionally left blank.

MODELS AND METHODOLOGY OF FAA DOMESTIC AIR CARRIER FORECASTS

Roger Schaufele Jr, Federal Aviation Administration

Background

This paper presents an overview of the methodology and models used by the FAA to forecast U.S. air carrier domestic revenue passenger miles (RPMs), domestic passenger enplanements, and domestic revenues. The FAA uses the forecasts of RPMs and enplanements to provide the basis for forecasts of aviation activity which are in turn, used to determine staffing levels and capital expenditures required to accommodate the growth of aviation activity while maintaining a safe, secure, and efficient environment.

Aviation forecasters have known for years that demand for aviation services, typically measured by Revenue Passenger Miles, or RPMs (one revenue passenger flying one mile) or enplanements, is influenced by a number of factors. In particular, demand is positively related to income and negatively related to price, typically measured by yield – passenger revenue divided by RPMs. Additional structural changes to the industry such as the introduction of jet aircraft in the late 1950's or deregulation of fares and routes (October 1978) have over time altered the relationships between demand and income and price. In addition, some unique events (such as when U.S. carriers engaged in destructive fare wars in 1986 and 1992) have temporarily altered the relationship between demand and the economic variables mentioned above. Despite these short-term alterations, aviation forecasters have been able to accurately predict the level of demand and thus enhance the planning process. FAA forecasts of demand have been quite accurate. During the seven year period 1995-2001, the average 1-year RPM forecast error was –0.2 percent (-1.0 percent for the period 1995 – 2000), while the average 5-year RPM forecast error was –1.5 percent. Events such as those that occurred on September 11th may so significantly alter the structure of the industry that models based on the historic relationships are inadequate for projecting demand in the future. In such circumstances, forecasters have to adopt different methods for projecting demand.

Overview of Models

In general, the models used for developing the FAA domestic air carrier forecast of traffic and yield rely upon a system of statistical and deterministic equations. The pivotal equations of the system relate RPMs and enplanements to two primary variables— Real U.S. Gross Domestic Product (GDP) and yield -- both adjusted for inflation. This analytical framework for forecasting enplanements ties the domestic forecast model closer to projected changes in economic activity and reduces the number of subjective inputs. To adjust for the jointly dependent variables in the demand and supply equations, two-stage least squares is used to estimate the system of equations. The primary independent variable used to estimate real yield is real unit costs, which is an identity relating Operating Expenses and ASMs (available seat miles). ASMs are a standard measure of capacity in the industry and are defined as one seat traveling one mile. ASMs are typically either an identity relating RPMs and load factor (RPMs/ASMs – a measure of the percentage of capacity that is utilized) or a function of GDP. Operating Expenses are in turn a function of three variables: ASMs, the Bank Prime Rate, and Jet Kerosene Prices. The inclusion of ASMs and Jet Kerosene Prices in the Operating Expenses equation is obvious as it follows that operating expenses are directly impacted by the amount of capacity offered and the price of jet kerosene. The inclusion of the prime rate is less clear. In airline accounting, interest expenses are not considered operating expenses. Thus, there is no a priori reason for including the prime rate in an operating expense equation. However the prime rate may serve as a proxy for the operating expenses of a general nature (passenger service, advertising, general and administrative, transport related) as well as depreciation and amortization expenses. The general functional form of the equation systems is as follows:

RPMs = f (GDP, Yield)
Yield = f (RPMs, Unit Costs)
Unit Costs = f (ASMs, Operating Expenses)
Operating Expenses = f (ASMs, Prime Rate, Jet Kerosene Price)
ASMs = f (RPMs, Load Factor) or f (GDP)

In the equation systems there are a number of exogenous shift variables. The majority of these dummy variables are temporary in nature, attempting to account for short run disruptions to the long run relationships. One of these variables accounts for the impact to yields resulting from the fare war in 1992 (referred to in the industry as the "summer sale"). A second accounts for the impact to traffic and yields of

Continental's low fare pricing experiment in East Coast markets during the 1993-1995 period. Dummy variables are also used to account for the impact on yields of the absence of the passenger excise tax between January and August 1996, and for the impacts resulting from the shutdown of Valujet following an accident in May 1996. In addition there are dummy variables to account for the structural changes resulting from Southwest's expansion into East Coast markets and the introduction of Regional Jets into service.

Description of Data

The data for RPMs, ASMs, enplanements, revenues and expenses is compiled and published by the Bureau of Transportation Statistics (BTS). It is available in the P (financial) and T (traffic) schedules of the Form 41 database. The RPM, ASM, and enplanement data is compiled and published monthly while the revenue and expense data are published on a quarterly basis. The revenue and expense data that were used do not include the revenue and expenses for all-cargo carriers such as Federal Express, UPS, etc. The jet kerosene price data is also compiled by BTS on a monthly basis. It does not include fuel taxes but does include consumption and expenditures of the all-cargo carriers. GDP and Personal Consumption Expenditure (PCE) data is from the Bureau of Economic Analysis (BEA) while annual values for the bank prime rate were obtained from the Federal Reserve Bank. Historic values of real yield were computed by dividing passenger yield by the Consumer Price Index (CPI-U).

Methodology

The FAA's forecasting process is a continuous and interactive one that involves the FAA Statistics and Forecast Branch, as well as other FAA offices, government agencies, and aviation industry groups. The forecast process has been referred to as "decision-theoretic" in nature. The approach is generally accomplished in two stages. Initially, projections are made with the use of the econometric models described later in the paper. The model results are then adjusted based upon "expert industry opinion" to arrive at the posterior forecasts used in the decision-making process. The industry is segmented into three classes: Network Majors [1], Low

Cost/Low Fare [2], and Other carriers. The rationale for this segmentation is that the response of travelers to changes in independent variables will be different in the three classes. For example, one would expect that the Low Fare/Low Cost carriers would have a higher price elasticity than the Network Majors whose passenger mix is typically more business oriented and less price sensitive. The Network Majors are those carriers who operate for the most part a traditional "hub and spoke" network. The Low Cost/Low Fare carriers are a select set of carriers who are or have been most recognized for their low fares. Other carriers is simply the industry total minus the sum of the Network Majors and Low Cost/Low Fare carriers, made up mostly of small carriers whose route networks are regional in nature. Each of these classes of carriers has their own system of equations that are used to project traffic and revenue. The total domestic air carrier forecast of traffic and revenue is simply the sum of the results of the forecasts for the three classes of carriers.

In normal circumstances, the process described above has served the FAA very well. However, as a result of the events of September 11th, the forecast process this year was somewhat different. The forecast was developed for 3 distinct periods: FY2002, FY2003, and FY 2004 -13. Very shortly after the events of September 11th it was decided that the current econometric models could not be used to predict aviation demand for 2002 and 2003. Instead, the process was modified to focus on generating a forecast of capacity (ASMs) for a smaller subset of the industry (namely, the carriers who make up the data published on a monthly basis in the Air Transport Association (ATA) traffic statistics). Once the capacity forecast was generated, the traffic forecast was obtained by combining a load factor forecast with the capacity forecast. Dividing forecasted RPMs by forecasted trip length generated the passenger forecast. Finally, the results of the monthly forecasts for the ATA sample were then used to generate forecasts for total industry capacity and traffic.

The forecasts of ASMs were based upon future schedules published in the Official Airline Guide (OAG). FAA forecasters had access to four updates of 12-month schedules—beginning with October

---

Northwest, TWA, United, US Airways and others (most notably Eastern and Pan Am (old))

[2] The low cost/low fare carriers are Southwest, Jet Blue, AirTran, Frontier, Vanguard, Spirit, Pro Air, Valujet, Morris Air, Kiwi, Carnival, TranStar, National, New York Air, Legend, Pan Am (new), People Express, Sun Country, American Trans Air, Western Pacific, Eastwind, and Air South.

---

[1] DOT defines major carriers as those with annual revenues in excess of $1 billion. The passenger carriers making up this class are Alaska, America West, American, Continental, Delta,

2001 and ending with January 2002. Each update was compared to previously published schedules and actual results to determine the accuracy of the published schedules. As a general rule, the FAA assumed that the 3 to 6 month forward schedules were fairly accurate. Beyond the 3 to 6 month period, adjustments were made based on discussions with airline planning staff and historic month to month patterns. Once the capacity forecast was created, traffic forecasts were developed based on assumptions about load factor. The assumptions about load factor for FY2002 were based on examining load factor behavior in the months following previous terror incidents – the bombing of Libya in April 1986, the bombing of Pan Am 103 in December 1988, and the start of the Gulf War in January 1991. Load factor forecasts for FY2003 were based upon historic month over month changes. As actual data became available, load factor and traffic forecasts were modified. Fortunately, the ATA shared with the FAA traffic and capacity data on a daily basis from its member carriers. This allowed the FAA to constantly assess the post-September 11th trends in traffic and capacity. Once finalized and reviewed internally, the forecasts and assumptions were presented to airline industry staff and aviation associations who were asked to comment on the reasonableness of the forecasts. Their comments were then incorporated into the final forecasts.

Forecasts of aviation demand for 2004 – 2013 were based on results of the models described in this paper. A key assumption of the FAA forecast is that the events of September 11th did not change the long-term relationships inherent in the forecast models. The December 2001 OMB forecast was the basis for the assumptions regarding GDP, PCE, CPI-U, oil prices, and the prime rate. Load factor assumptions were based on expert judgment. In particular for the Network Majors, it was assumed that load factor would decrease from the levels of FY 2003 as these levels were temporarily elevated due to supply constraints. Load factor was assumed to decline from 73.2% in FY 2003 to 72.5% in FY 2006 and remains at that level throughout the balance of the forecast period. Jet kerosene price forecasts were estimated outside the system of equations utilizing a simple model relating jet kerosene prices to U.S. refiner acquisition cost of oil and the prior year price of jet kerosene.

Discussion of Results

The equation systems for all three classes of carriers were estimated using two-stage least squares regression. Estimation results are found in Table 1 at the end of the paper. For the Network Majors, the estimation period was 1988-2000. While most of the variables were statistically significant (see Table 1), some variables remained in the models on theoretical grounds despite test statistics that suggest omission. Key among these variables is the inclusion of the real yield variable in the RPM equation. In addition some variables remained in the system for stability purposes (The CALITE variable in the RPM equation and the WN variable in the enplaned passenger equation). Forecast values of ASMs were determined by dividing forecast values of RPMs by forecast values of load factor.

The Low Cost/Low Fare carrier equation system was estimated for the period 1992 to 2000. This is shorter estimation period than the estimation period for the network major equation system. When the low cost/low fare equation system was estimated for the same period as the network majors equation system, the income elasticities were significantly higher than in the final low cost/low fare equation system, and generated unreasonable demand forecasts for this class of carriers. Reducing the period for which the low cost/low fare equation system was estimated, allowed the FAA to have a system of equations that has attractive goodness of fit properties and produces forecasts of demand which appear to be reasonable in light of projected levels of economic activity.

As in the case of the Network Major equation system, some variables that were not statistically significant (see Table 1) remained in the models on theoretical grounds as well as variables that provided stability to the system. (LRYLD in the RPM equation and CALITE in the ASM equation) Unlike the network majors the key income variable is PCE, not GDP. This is consistent with the premise that the passenger mix for the low fare carriers is much more leisure oriented than for the network majors, and that changes in leisure travel are more closely linked to changes in spending (PCE) rather than changes in output (GDP). When GDP was substituted for PCE in the RPM and passenger equations, the coefficient estimate for GDP was not statistically significant from zero in either equation. These results provided additional evidence to support the choice of PCE as the appropriate income variable in the Low Cost/Low Fare equation system.

The Other carriers equation system was estimated for the period 1990 to 2000. Similar to the Network Major and Low Cost/Low Fare systems, certain variables remained in the system on theoretical grounds despite test statistics that suggest omission (The GDP variable in both the RPM and ASM equation and the CALITE and VALUJET variables in the ASM equation). These variables have the intuitively correct signs and appear to provide stability to the system. Unlike the Network Major and Low Cost/Low Fare systems, the Other carrier system contains no yield equation. A number of specifications for a yield equation were tried but none produced satisfactory results and it was assumed that future rates of growth in yield for this class of carriers would be the same as that of the Network Majors.

Demand Forecasts

In the aftermath of the September 11th terrorist attack, U.S. air carriers immediately reduced domestic capacity by approximately 20 percent across the board. The FAA forecast assumes that domestic capacity will gradually return to the pre-September 11th capacity levels over a 3-year period. Domestic capacity is forecast to decline by 10.1 percent in 2002, then increase by 7.3 percent in 2003 and 4.6 percent in 2004. Thereafter, capacity is expected to increase at an average annual rate of 4.1 percent over the final 9 years of the forecast period. Capacity for the Network Majors is projected to decline 12.1 percent in 2002, then increase 7.2 percent in 2003, and 2.6 percent per year thereafter. The Low Cost/Low Fare carriers are projected to experience a slow down in capacity growth to 2.9 percent in 2002, then increase to 9.5 percent in 2003 and average 8.7 percent per year from 2004 through 2013.

Domestic air carrier RPMs and passenger enplanements are forecast to increase at average annual rates of 3.5 and 3.1 percent, respectively, over the 12-year forecast period. Domestic RPMs and enplanements are forecast to decline by 12.0 and 13.4 percent, respectively, in 2002, then grow by 14.0 and 14.8 percent, respectively, in 2003. For the Network Major carriers, domestic RPMs and enplanements are projected to decline 13.2 and 14.4 percent, respectively in 2002. Growth resumes in 2003 with RPMs up 13.5 percent while enplanements increase 15.2 percent. Domestic RPMs and enplanements for the Low Cost/Low Fare carriers are projected to fall 3.4 and 4.3 percent, respectively, in 2002. In 2003, domestic RPMs and enplanements for this class of carriers increase 17.3 and 14.5 percent, respectively.

U.S. carriers are expected to return to normal growth trends beginning in 2004, with RPMs and enplanements averaging 4.2 and 3.8 percent, respectively, over the remainder of the forecast period. During the same period, domestic RPMs and enplanements for the Network Majors average 2.5 and 1.6 percent per year, respectively, while averaging 9.5 and 7.9 percent per year for the Low Cost/Low Fare carriers.

After declining for 2 consecutive years (to 68.2 percent in 2002), domestic load factors are expected to increase to 72.5 percent in both 2003 and 2004. Load factors are then expected to increase gradually over the remainder of the forecast period, averaging 73.2 percent in 2013.

Domestic passenger yields, which declined by 3.5 percent in 2001, are expected to decline an additional 3.4 percent (down 4.1 percent for the Network Majors but up 0.6 percent for the Low Cost/Low Fare carriers) in 2002. Yields are forecast to increase by 7.9 percent in 2003 (Network Majors up 9.7 percent; Low Cost/Low Fare carriers up 4.3 percent) and then grow at an average annual rate of 1.2 percent (down 1.2 percent in real terms) over the remaining 10 years of the forecast period. Both the Network Majors and the Low Cost/Low Fare carriers are forecast to have real yields decline an average of 1.3 and 0.5 percent per year, respectively, during this period. The relatively large increase in 2003 is due, in large part, to anticipated strong demand from both leisure and business travelers such that the resultant traffic mix more closely approximates the levels achieved prior to the start of the 2001 recession.

The decline in real yields over the latter years of the forecast is based on the assumption that competitive pressures will continue to exert pressure on carriers to hold the line on fare increases. Competition in domestic markets will come from established low-fare carriers such as Southwest, as well as from smaller low-cost carriers such as AirTran, Frontier, and JetBlue.

In addition to forecasting U.S. air carrier domestic traffic and revenues utilizing the models and methods described previously, the FAA forecasts demand for U.S. air carriers in international markets, the regional/commuter industry, and the general aviation industry. These forecasts employ a variety of methods including the use of econometric models as well as delphi forecasts. The demand forecasts for these segments of the aviation industry combined with the forecasts of U.S. air carrier domestic

demand are used to develop forecasts for overall activity forecasts at FAA air traffic facilities.

Areas of Further Research

Although it is aggregate demand that the FAA forecasts, it would be preferable to use different models to estimate the two distinct components of each market--business and personal travel. A further refinement would distinguish the long-haul from the short-haul market. This approach would provide important information for developing public policy and would most likely improve the accuracy of the forecasts. Clearly, these markets are affected by different sets of variables, and adjust at different rates to them.

For example, most experts in the industry would agree that the price elasticity of demand for business travel differs from the price elasticity of demand for pleasure travel. Furthermore, theory would suggest that business profits are a factor in determining business travel, and that some measure of personal or family income is an important variable affecting pleasure travel.

At this time, however, the lack of an adequate database subdivided into these four components precludes the development of forecasts for each market at the national level. Additional research and data collection is necessary to advance this approach.

A second area of further research involves the measure of the price influencing aviation demand. In recent years the amount of excise taxes and fees added on to the base price of a ticket have increased greatly and may now be large enough to influence the modal choice of travelers. In addition, as more and more consumers have access to low base fares, the percentage of the average ticket price that taxes and fees run is increasing, distorting the information that the consumer receives when he may make his decision to fly. For example, the $200 round trip ticket to Florida may actually cost the customer $250-$260 after all the taxes and fees are levied. If airline demand is becoming increasingly leisure oriented and more prices sensitive, ignoring the tax impacts on behavior may lead us to overestimate the level of demand in the future. The definition of yield that is currently used does not include the amount of taxes that the consumer pays and may represent a misspecification of the price variable that should be used in models estimating aviation demand.

Finally, further work needs to be done on the impact of consolidation on demand and fares. In general we know that if the supply is reduced and competition lessened, prices will go up or at the very least not decline as rapidly. Given the precarious financial state of the industry, it may be a worthwhile exercise to examine what would happen to demand should one or two of the major network carriers cease to operate. Much of the impact would depend upon the nature of the consolidation and the carriers involved. It is obvious that the impacts on demand and prices of a failure of a network major carrier with significant operations in all regions would be very different than the failure of a network major carrier with significant operations in few regions. Such forecasts would involve doing scenario analyses but would probably improve the usefulness of the FAA forecasts.

Independent Variables	Network Majors				Low Cost/Low Fare					Other Carriers			
	LRPM	LRYLD	LENP	LOPEXP	LRPM	LRYLD	LENP	LRCASM	LASM	LRPM	LENP	LOPEXP	LASM
C	**4.96**	**4.81**	**8.00**	0.39	(13.12)	**(1.04)**	**(9.72)**	(0.53)	**(13.10)**	2.53	**(9.09)**	**(1.70)**	3.94
LGDP	**0.86**	---	**0.55**	---	---	---	---	---	2.63	0.74	**2.18**	---	0.64
LRYLD	(0.07)	---	---	---	(0.87)	---	---	---	---	---	---	---	---
CALITE	(0.02)	---	---	---	---	(0.03)	---	---	(0.12)	**(0.40)**	(0.08)	---	(0.33)
LRPM	---	**(0.38)**	---	---	---	**(0.09)**	---	---	---	---	---	---	---
LRCASM	---	**0.83**	---	---	---	---	---	---	---	---	---	---	---
LRCASM(-1)	---	---	---	---	---	---	---	**0.78**	---	---	---	---	---
SUMSALE	---	**(0.05)**	---	---	---	---	---	---	---	---	---	---	---
WN	---	---	0.03	**(0.05)**	---	---	**0.32**	---	**0.50**	---	---	---	---
LPRIME	---	---	---	**0.10**	---	---	---	---	---	---	---	---	---
LASM	---	---	---	**0.80**	---	---	---	---	---	---	---	**0.65**	---
LKERO	---	---	---	**0.12**	---	---	---	**0.10**	---	---	---	---	---
TKTTAX	---	---	---	**(0.03)**	---	---	---	---	---	---	---	---	---
LPCE	---	---	---	---	**2.57**	---	**2.39**	---	---	---	---	---	---
VALUJET	---	---	---	---	---	---	---	**0.03**	---	**(0.22)**	---	---	**(0.17)**
RJETS	---	---	---	---	---	---	---	---	---	**0.72**	**0.22**	---	**0.60**
LOPEXP(-1)	---	---	---	---	---	---	---	---	---	---	---	**0.44**	---
AR(1)	---	---	**0.52**	**(0.60)**	**0.71**	**(0.69)**	---	---	---	**0.44**	---	---	0.44
R^2	0.978	0.929	0.940	0.845	0.977	0.949	0.955	0.937	0.947	0.937	0.979	0.960	0.895

Coefficient values in bold have t-statistics significant at 10% level

AR(1) = first order autoregressive term

CALITE = 0-1 variable to account for impact of Continental's low fare pricing in East Coast markets that was in effect between Oct 1993 through Mar 1995. Variable has value of 1 for FY 1994 and 1995 and 0 in all other years.

LASM = Log of Domestic ASMs

LENP = Log of Domestic Enplanements

LGDP = Log of U.S. Real GDP (in 96$)

LKERO = Log of Jet Kerosene Price per Gallon in 2001 cents

LOPEXP = Log of Domestic Operating Expenses in 2001$

LOPEXP(-1) = One period lag of LOPEXP

LPCE = Log of U.S. Real PCE (in 96$)

LPRIME = Log of U.S. Bank Prime Loan Rate

LRCASM = Log of Domestic Operating Cost Per ASM in 2001 cents

LRCASM(-1) = One period lag of LRCASM

LRPM = Log of Domestic RPMs

LRYLD = Log of Domestic Real Yields in 2001 cents

RJETS = 0-1 variable to account for structural change to industry resulting from introduction of regional jets into service. Variable has value of 0 through FY 1993 and 1 thereafter.

SUMSALE = 0-1 variable to account for impact to yields from "summer sale" of 1992. Variable has value of 1 for FY 1992 and 0 in all other years.

TKTTAX = 0-1 variable to account for impact of lack of passenger excise tax between January 1996 and August 1996. Variable has a value of 1 in FY 1996 and 0 for all other years.

VALUJET = 0-1 variable to account for impacts resulting from shutdown of Valujet following accident in May 1996. Variable has a value of 1 in FY 1996 and 0 for all other years.

WN = 0-1 variable to account for structural change to industry resulting from expansion of Southwest into East Coast destinations. Variable has value of 0 through FY 1993 and 1 thereafter.

The Impact of Terrorism on Tourism by Use of Time Series Methods
Brian W. Sloboda

U.S. Department of Transportation
Bureau of Transportation Statistics

Tourism is one of the largest industries and many open economies especially small nations rely heavily on tourism as a major revenue source. Business and public-policy officials are often interested in the impact of tourism at all levels. Thus, the terrorism can hinder the tourist sector by keeping tourists away after major terrorist attacks, and in the long-term the indirect costs of terrorism include greater expenditures for advertising in order to attract tourists, the rebuilding of tourist facilities, and providing greater security measures to lessen terrorist activities.

The escalation of terrorism in recent years may have caused some impact on tourism in the United States, which became more evident after the September 11 attacks on the World Trade Center and the Pentagon. These effects rippled throughout the economy, and this negative ripple proved more damaging in the travel and tourism sectors. Bellhops from Las Vegas to baggage handlers at airports such as O' Hare and Los Angeles to the housekeepers at hotels throughout the nation experienced unemployment. In the longer term, resources will shift from the production of other goods and services to security based issues. In the travel and tourism industry, the onset of new terrorism[1] security will become a primary

concern in the tourism industry which means greater costs to be borne by the travelers. As a result of these higher costs, travelers will reduce their demand for travel and tourism

The purpose of this paper is to determine whether terrorism has impact on tourism to the United States by use of the ARMAX models. ARMAX models which allow for an assessment of impacts of terrorism on tourism. The balance of this paper is as follows. Section 2 provides a review of the literature in this area. Section 3 presents an empirical methodology for this analysis and a discussion of the data sources. Section 4 provides the empirical results of the analysis and Section 5 concludes the paper.

Section 2: Review of the Literature[3]

In the consumer maximization model, consumers are faced with an optimization

---

[1] In recent years, a new breed of terrorism has been emerging which may be attributed to our success in controlling state-sponsored terrorism. Today' s threat stems from non-state sponsored terrorism such as the Al-Quaida network, Aum Shimrikyo in Japan, and FARC in Columbia. In addition to non state sponsored terrorism, more terrorists are acting on their own to portray their cause in religious and cultural terms. In fact, the tactic used is to conceal their actual political goals, generate popular support, and silence the opposition. Yet the generation of public support stems from the resentment and suffering of people who feel marginalized in our global economy. Furthermore, when the state government is weak in providing stability and providing basic services to the people, terrorist groups construct parallel institutions in providing these services, i.e., madrasses or religious schools in Pakistan and other Islamic nations.

[2] This is a shortened version of the paper presented. If interested in the complete paper, contact the author at brian.sloboda@bts.gov.

[3] In addition to time series analysis of terrorism, the effects of terrorism on tourism can also be analyzed by use of economic base models. However, this will not be delved in this paper because of space constraints. The basic questions often addressed by this type of analysis are:

1. How much do tourists spend in the area?
2. What portion of the sales by local businesses is due to tourism?
3. How much income does tourism generate for households and businesses in the area?
4. How many jobs in the area does tourism support?
5. How much tax revenue is generated from tourism?

In addition, there are other types of economic impact studies that assess the effects of tourism on a regional economy. Frechtling (1994) provided for an analysis via input-output analysis which traces the flows of spending associated with tourism activity in a region by identifying the changes in sales, tax revenues, income, and jobs creation attributed to tourism. The reader should consult [Walsh 1986, Johnson and Thomas 1992; Stokey and Zeckhauers 1978; Sudgen and Williams 1978; and Warnell 1986] for various analyses using these models.

problem given many constraints or the consumer needs to allocate scarce resources among many choices. Thus, consumers want to maximize their utility subject to a budget and time constraints. As for terrorists, they are also behave in a rational way since terrorists want to maximize their goals subject to constraints that include resources and risks imposed by authorities[4] [Sandler, Tschirhart, and Cauley (1983), Lapan and Sandler (1987), Sandler and Scott (1987), Atkinson, Sandler, and Tschirhart (1987), and Im, Cauley, and Sandler (1987)]. By having changes in their constraints, these changes lead to predictable adjustments. A application of the economic choice model to terrorism was done by Landes (1978). Landes examined hijackings from the period 1961-1976 in the United States. His empirical results revealed that the use of sky marshals and metal detectors had a positive effect on the probability of preventing hijackings and a negative influence on the number of hijackings. Thus, the analysis implies that the terrorists were acting rationally.

Enders and Sandler (1991) and Enders, Sandler, and Praise (1992) provide an empirical framework concerning the link between terrorism and the tourism industry for a sample of European nations. Enders and Sandler (1991) find a significant negative impact of terrorism on tourism in Spain. Enders, Sandler, and Praise (1992) used the same sample of European nations for the period 1974-1988 as in Enders and Sandler (1991) analysis. Their later analysis used an autoregressive integrated moving average (ARIMA) with a transfer function, and they modeled the share of tourism in these European nations using quarterly data for number of terrorists incidents and tourist receipts. From their analysis, they concluded

that terrorist incidents have an adverse effect on tourism revenues in Europe; in addition, tourists often substitute from some countries to others to minimize their risk of being involved in a terrorist incident.

Drakos and Kutan (2001) extend the analysis of Enders, Sandler, and Praise (1992). First, their analysis tests the cross-country effects of terrorism on tourism in the Mediterranean region. That is, they test to see if terrorism affects regional competition for tourism. Second, their analysis incorporates countries from Mediterranean nations while earlier studies focused on European nations. In addition, they incorporate the countries of Israel and Turkey since they have been subject to increasing terrorist incidents in recent years and neglected in prior analyses. More importantly, their analysis incorporates Italy as a control country which represents the Mediterranean region to estimate the effects of terrorism on other market shares in the region. For the empirical methodology, they use an autoregressive model and estimate it by use of the seemingly unrelated regressions (SUR) model. Their empirical results reveal that the tourism industry in Israel and Turkey are more sensitive to terrorism than in Greece.

Aly and Strazicich (2000) examined the effects of terrorism on the tourism sector for Egypt and Israel. The examined to see if the shocks to the time paths are permanent or transitory since no prior analyses have directly examined the time path of tourist visits and its subsequent effects. To conduct the analysis, they utilized a two-break minimum Lagrange Multiplier (LM) unit root test as developed by Lee and Strazicich (1999)[5].

Section 3:
Empirical Methodology and Data Sources

3.1 Empirical Methodology

A common approach to model the effects of terrorism on tourism is through a transfer function or dynamic regression. In a dynamic regression or transfer function, the output time series is influenced by the input time series. In other words, the input time series will exert influence over the output series over several future time periods. The generalized model can be written as

$$(1)\ A_i(L)y_i(t) = a_i + B_i(L)x_i(t) + C_i(L)e_i(t)$$

where $y_i(t)$ is the logarithm of the tourist receipts; $x_i(t)$ is the number of terrorist incidents occurring in the time period, a_i is the intercept or constant term; and e_i is the error term in which $e_i \sim WN(0,s^2)$. $A_i(L)$, $B_i(L)$, and $C_i(L)$ are polynomials with a lag operator L. The interpretation of equation (1) is rather straightforward. The coefficients of $A_i(L)$ provide the autoregressive (AR) components while the coefficients of $C_i(L)$ provide the moving average (MA) components of the ARIMA model. The coefficients shown by $B_i(L)$ reveal the immediate impact of the terrorism into tourism receipts or is called the transfer function because it shows the movement of the exogenous variable (terrorist acts in this case) affects the time path of the endogenous variable of tourist receipts. In addition, the values of the coefficients as given by $B_i(L)$ are called the impulse response weights and these coefficients assess how tourist receipts respond to a change in terrorist acts. Given equation (1), the effects of terrorism can be estimated. Then, equation (2) will be estimated by the methods of Box Jenkins (1976).

In a transfer function approach, the specification of the error term follows an AR process. It is possible to describe the error term as a MA process, and the transfer function, then, becomes an ARMAX or (Autoregressive Moving Average with Explanatory Variables) models. From these ARMAX models, impacts can be assessed by the general shape of the lag distributions of the impacts of the explanatory variables can be assesses by taking the ratio of the lag polynomial for the dependent and independent variables. The general ARMAX model for this analysis can be given as

$$(2)\ \begin{aligned} y_t &= m + g_1 y_{t-1} + g_2 y_{t-2} + \dots + g_p y_{t-p} \\ &+ b_0 x_1 + b_2 x_{t-2} + e_t - q e_{t-1} - \dots - q e_{t-q} \end{aligned}$$

where p denotes the lag length for the autoregressive terms and q represents the lag of the moving average error terms. Also $e_t \sim WN(0,s^2)$. Now it can be shown how the ratio of the lag polynomials can be derived to show impacts of terrorism on tourism. To keep the discussion simple, only a two period case will be considered, and the total impact of a change of terrorism on tourism can be measured[6].

$$(3)$$
$$y_t = (1 - g_1 L + g_2 L^2) = (b_0 + b_1 L + b_2 L^2) + x_t$$
and A(L)= $(1 - g_1 L + g_2 L^2)$ and B(L)=
$$(1 - g_1 L + g_2 L^2)$$

then written in more compact form, it becomes A(L)y=B(L)x. Now the impacts for terrorism on tourism can be assessed by the ratio of two polynomials of A(L) and B(L) which is given as W(L)=B(L)/A(L) This can be shown by (4) as

$$(4)$$
$$y = W(L)x = \frac{B(L)}{A(L)} x = (w_0 + w_1 L + w_2 L^2 + w_3 L^3 + \dots)$$

Equation (4) shows an infinite series lag distribution. Now the computation of the lag

---

[6] The lag impact of terrorism or x is independent from the constant and the disturbance terms.

weights in W(L) is shown from the expanded form or W(L)A(L)=B(L) or

(5)
$$(w_0 + w_1L + w_2L^2 + z_3L^3 + ...)(1 - g_1L + g_2L^2)$$
$$= (b_0 + b_1L + b_2L^2)$$

Now expand the product of W(L) A(L) by multiplying the two terms in (5) and obtain

$$w_0 + w_1L + w_2L^2 + w_3L^3 + ... - w_0g_1L - w_1g_2L^2 - w_2g_1L^3$$
$$- ..... - w_0g_2L^2 - w_1g_2L^3 - w_2g_2L^4 - ...$$
$$= (b_0 + b_1L + b_2L)$$

After combining the L terms, proceed to solve for the w terms. The significance of the w terms is that the summation of these terms provide the total impact of terrorism on tourism

3.2 Data Sources:

The data on terrorist incidents were taken from the US Department of State publication, *Patterns of Global Terrorism* for the years 1987-2000. The data for terrorist incidents represent the x variable and are assembled on a quarterly basis. At the present, there is no one definition of terrorism which is used universally. The terrorist incident data used in this analysis uses the data from the *Patterns of Global Terrorism* that follows the definition as contained in Title 22 of the United States Code, Section 2656f(d). This statute contains the following definitions in regards to terrorism[7]:

- The term "terrorism" means premeditated, politically motivated violence perpetrated against

noncombatant targets[8] by sub-national groups or clandestine agents, usually intended to influence an audience.
- The term "international terrorism" means terrorism involving citizens or the territory of more than one country.
- The term "terrorist group" means any group practicing, or that has significant subgroups that practice, international terrorism.

The US Government has employed this definition of terrorism for statistical and analytical purposes since 1983. However, these data do not provide statistics on domestic terrorism.

As for the y variable, the data on tourist receipts are taken from the Bureau of Economic Analysis' publication *Survey of Current Business* data on Balance of Payment of Data. More specifically, the quarterly values on travel (line 6) and passenger services (line 7) were summed to provide an approximate measure of total receipts of a nation from tourism.

Section 4: Conclusions and Suggestions for Future Work

The analysis demonstrated here did not reveal substantial losses to the tourism industry for the United States for the indicated time period. Instead, the tourist receipts have been increasing steadily over time. As a remedy, this analysis needs to incorporate other countries into the analysis in order to fully

---

[7] There are chronologies concerning terrorist incidents such as [(Mickolus 1980); Mickolus, Sandler, and Murdock 1989); (Mickolus 1993); and (Mickolus and Simmons 1997)]. These chronologies are the only one publically available and relies heavily on articles available in the press and the key sources include the Associated Press, United Press International, *the Washington Post, the New York Times, the Washington Times*, the Foreign Broadcast Information Service Daily Reports, and the network news.

[8] For purposes of this definition, the term "noncombatant" is interpreted to include, in addition to civilians, military personnel who at the time of the incident are unarmed or not on duty. For example, in past reports we have listed as terrorist incidents the murders of the following US military personnel: Col. James Rowe, killed in Manila in April 1989; Capt. William Nordeen, US defense attache killed in Athens in June 1988; the two servicemen killed in the Labelle Discotheque bombing in West Berlin in April 1986; and the four off-duty US Embassy Marine guards killed in a cafe in El Salvador in June 1985. These data also consider as acts of terrorism attacks on military installations or on armed military personnel when a state of military hostilities does not exist at the site, such as bombings against US bases in Europe, the Philippines, or elsewhere.

realize the impacts of terrorism. More specifically, the incorporation of countries from Europe and the Middle East into the analysis would provide a better assessment of the impacts of terrorism.

Also after incorporating other nations into the analysis, the analysis can incorporate structural breaks since the failure to incorporate structural breaks leads to the loss of power of the unit root tests such as the Dickey Fuller and Phillips-Perron tests. There are more recent theoretical models as indicated in the review of the literature section of the paper that employ structural breaks through the LM unit root test that is not subject to spurious regressions. Based on the structural breaks, then proceed to estimate the losses for tourism.

References

Aly, H.Y and M.C Strazicich. (2000). "Terrorism and Tourism: Is the Impact Permanent or Transitory? Time Series Evidence from some MEAN Countries." Department of Economics, The Ohio State University Working Paper Series.

Atkinson, S.E, T. Sandler, and J. Tschirhart. (1987). "Terrorism in a Bargaining Framework." *Journal of Law and Economics*, 30: 1-21.

Dickey, D. and W. Fuller. (1979). "Distribution of the Estimates for Autoregressive Time Series with a Unit Root.*" Journal of the American Statistical Association*, 74, 427-431.

Drakos, K. and A.M Kutan. (2001). "Regional Effects of Terrorism on Tourism: Evidence from Three Mediterranean Countries." A Working Paper B 26, Center for European Integration Studies, Rheinische Friedrich-Wilhelms-Universitat Bonn.

Enders, W and T. Sandler. (1991). "Causality Between Transnational Terrorism and Tourism: The Case of Spain." *Terrorism*, 14, 49-58.

Enders, W and T Sandler, and G.F. Parise.(1992). "An Econometric Analysis of the Impact of Terrorism on Tourism." *Kyklos*, 45, 531-554.

Frechtling, D.C. (1994). "Assessing the Economic Impacts of Travel and Tourism—Measuring Economic Benefits." In *Travel, Tourism , and Hospitality Research*, 2nd edition. J.R. Brent Ritchie and C.R. Goeldner (editors) .New York: Jon Wiley and Sons.

Harvey, A.C. (1990). *The Econometric Analysis of Time Series*. Cambridge, Massschusetts: The MIT Press.

Im, E.I, J. Cauley, and T. Sandler. (1987). "Cycles and Substitutions in Terrorist Activities: A Spectral Approach." *Kyklos* 40: 238-255.

Johnson, P. and B. Thomas. (1992). *Choice and Demand in Tourism*. London: Mansell.

Landes, W.M.(1978). "An Economic Study of U.S. Aircraft Hijackings, 1961-1976." *Journal of Law and Economics,* 21, 1-31.

Lapan, H.E. and T. Sandler. (1987). "To Bargain or Not to Bargain: That is the Question." Unpublished manuscript.

Lee, J. and M. Strazicich. (1999). "Minimum LM Unit Root Test with Two Structural Breaks." Department of Economics, University of Central Florida Working Paper Series.

Lumsdaine, R. and D. Papell. (1997). "Multiple Trend Breaks and the Unit Root Hypothesis." *Review of Economics and Statistics*, 212-218.

Mickolus, E.F. (1980). *Transnational Terrorism: A Chronology of Events 1968-1979,* Westport, CT: Greenwood Press.

Mickolus, E.F., T. Sandler, and J.M Murdock. (1989*). International Terrorism in the 1980's: A Chronology of Events*, Vol 1 (1980-1983) and Vol 2 (1984-1987), Ames, Iowa: Iowa State University Press.

Mickolus, E. (1993). *Terrorism, 1988-1991: A Chronology of Events,* Westport, CT: Greenwood Press.

Mickolus, E. and S.L Simmons. (1997). *Terrorism, 1992-1995: A Chronology of Events and A Selectively Annotated Bibliography,* Westport, CT: Greenwood Press.

Nunes, L. P. Newbold, and C. Kuan. (1997). "Testing for Unit Roots with Breaks: Evidence on the Great Crash and the Unit Root Hypothesis Reconsidered." *Oxford Bulletin of Economics and Statistics*, 59, 435-448.

Perron, P. (1989). "The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis." *Econometrica*, 57, 1361-1401.

Perron, P. (1997). "Further Evidence on Breaking Trend Functions in Macroeconomic Variables." *Journal of Econometrics*, 80, 355-385.

Sandler, T., J. Tschirhart, and J.Cauley. (1983). "A Theoretical Analysis of Transnational Terrorism." *American Political Science Review*, 77: 36-54.

Sandler, T. and J.L. Scott. (1987). "Terrorist Success in Hostage-taking Incidents. *Journal of Conflict Resolution*, 31: 35-53.

Stynes, D. (1998). "Understanding Multipliers and How to Interpret Them." Michigan State University, available at http://www.msu.edu/course/prr/840/econimpact/index.htm

Sudgen, R. and A. Williams. (1978). *The Principles of Practical Cost-Benefit Analysis*. Oxford: Oxford University Press.

Walsh, R.G. (1986). *Recreation Economic Decisions-Comparing Costs and Benefits*. State College, PA: Venture Publishing Co.

Warnell, G. (1986). *Feasibility Analysis in Tourism*. Extension Bulletin E-1992. East Lansing, MI: Michigan State University Cooperative Extension Service.

THE IMPACT OF SEPTEMBER 11, 2001 ON TRANSPORTATION INDICATORS

Peg Young, Bureau of Transportation Statistics, U.S. Department of Transportation
Keith Ord, The McDonough School of Business, Georgetown University

ABSTRACT:

The Bureau of Transportation Statistics has created a monthly report, called Transportation Indicators, which reports on key measures related to the transportation enterprise. The co-authors of this paper have created a procedure, using STAMP, to decompose the time series of interest and to create monthly forecasts of these indicators. In addition, the procedure compares the new actual values of these measures to the one-step-ahead forecasts in order to provide alerts for those measures that deviated more than expected every month. This presentation will show the results of this forecast and statistical process control procedure, particularly in light of the events on September 11, 2001 on transportation data.

1. BTS Transportation Indicators Report

In November 1999, the Bureau of Transportation Statistics (BTS) undertook a project to create a monthly report on key measures, or indicators, of the transportation system in the United States. The general spectrum of these measures would cover the strategic goals of the Department: Safety, Mobility, Economic Growth, Human and Natural Environment, and National Security. Since it would be a monthly report, the data series selected to be included in this report were of a frequent nature: weekly, monthly and quarterly. (Yearly data sets would be included only in those instances where more frequent data did not exist.) It was understood by BTS management that this report would be a continually changing document, with new variables being introduced and old variables that proved to be of little value being removed. The first issue of the Transportation Indicators (TI) report, which came out in May 2000, contained over 70 indicators and encompassed over 120 data series; currently, 117 indicators are represented by over 320 time series. An example of an indicator page in the report is provided in Figure 1.

2. Goals of Monitoring System

The report was designed to serve as a resource of up-to-date information on the transportation enterprise that transportation executives could not obtain, in a single source, anywhere else inside or outside the Department. The indicators would be offered in a simple form, incorporating a short paragraph describing the data set, a graph of each series over the past 10 years or so, and a table comparing the most recent values of the data series. For highly seasonal data, data comparisons would be provided for the same period in the previous year.

ENPLANEMENTS ON DOMESTIC FLIGHTS

Domestic Airline Revenue Enplanements
(monthly data, not seasonally adjusted)
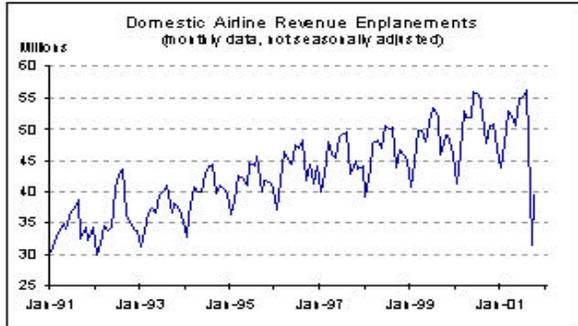
Domestic Passenger Aviation	Oct-00	Oct-01
Revenue aircraft enplanements (millions)	50.53	39.80
Percent change from same month previous year	2.88	-21.23

NOTES: The current value is compared to the value from the same period in the previous year to account for seasonality.

The data have been adjusted to have a standard 30-day month by multiplying the data for each month by the ratio: 30/actual days in month.

The data do not include international flights by U.S. domestic carriers or domestic flights by foreign carriers.

The dramatic changes in the September 2001 data reflect the impact of the terrorist attacks on September 11, 2001, on aviation, including several days in which commercial air operations were suspended.

SOURCE: U.S. Department of Transportation, Bureau of Transportation Statistics, *Air Carrier Traffic Statistics Monthly*, October 2001.

Revenue enplanements, the number of passengers boarding aircraft, indicate the demand for gate and luggage services. Enplanements differ from the number of trips because passengers may board more than one flight between their origination point and ultimate destination.

FIGURE 1. Example of page from Transportation Indicators report (February 2002 issue).

In addition to providing the time series data within the report, the project also had as an agenda item the task of superimposing a monitoring system for each time series within the report. Such a tracking system would provide a monthly alert system for senior management to advise them that the new values of certain indictors had behaved in an 'unexpected' manner. In order to state that behavior was unexpected, the BTS TI team was tasked to devise a method to describe expected behavior. That is, we needed to forecast each series at least one period ahead, to create a procedure to compare the forecasts to the new observations, and to be able to declare the new data as either expected or unexpected. The terminology "expected/unexpected" may be viewed as broadly equivalent to the more common SPC concept of being in or out of control.

This task would require a forecasting mechanism that would provide a forecasting model that could be updated easily and quickly, and would also permit comparison with recent data. Since many of the data series exhibit strong seasonality, these forecasting models would also need to allow for deseasonalization (or for decomposition of the seasonal component); in this way the readers of the report would be able to see the underlying trend along with the actual series. Finally, the forecasting process would also need to be able to handle interventions in the data. The intervention of interest for this paper reflects the impact of September 11, 2001.

3. Description of the Data Series

The indicators can be classified by eight different criteria. The first criterion is simply a count of the number of time series within that particular indicator; this ranges from one to seven individual time series. The indicator is also classified by the strategic goal it represents (Safety, Mobility, Economic Growth, Human and Natural Environment, or National Security). The data in the report are generated in one of three possible ways: by sample surveys, by enumeration of the whole population, and by model-based analysis of empirical data. Some series are drawn from BTS-controlled sources, such as the data from the Office of Airline Information; the other datasets are outside the control of BTS. The recording frequency of the series can be weekly, monthly, quarterly, and annual. Some of the series are unusually short, so they were also classified by start date. Several of the series are pulled from sources that only provided the data in a seasonally adjusted format; since this would affect the forecasting models selected, the data series are classified as to whether or not they had already been seasonally adjusted. Finally, the data series are reviewed to determine if they exhibited any additional characteristics not yet captured that would affect the forecast model selected (e.g., apparent interventions).

4. Monitoring Using Time Series

One purpose for using time series analysis would be to break a series down in to its core components (trend, seasonal, and irregular) so that we may examine each one separately. The basic ideas for monitoring flow from statistical process control (SPC). The use of time series modeling in SPC follows from the seminal work of Alwan and Roberts (1988, 1995).

In SPC, we conventionally distinguish two sources of variation (c.f. Alwan, 2000, pp. 217-220):

- *Common cause variation:* reflects the natural variation inherent in the process, and
- *Special (or assignable) cause variation:* any variation in the process introduced by a recognizable factor [e.g. a worn tool or a poorly trained operative].

In the present context, we are interested in monitoring changes in a phenomenon over time, and the possible types of assignable cause need to be identified more clearly.

Thus, it is useful to divide assignable cause variation into four categories, which we may examine by different means:

- *Temporary:* a factor has a short-term impact on the series, which returns to its previous level fairly rapidly. For example, severe winter weather may temporarily reduce employment, but the economy would recover in the next month or two.
- *Level-shift:* a factor causes the series to shift to a new level, and it stays at that new level. For example, a change in reporting requirements might change the level of a series, but not otherwise affect the nature of the phenomenon.

- *Seasonal:* the seasonal pattern in the series may change over time. For example, airlines may change their seasonal pricing strategies, which would lead to a shift in travel patterns.
- *Long-term*: over a period of time, changing conditions lead to fundamental changes in the series of interest. For example, improved engine design might produce improved fuel efficiency ratings for automobiles, but such an effect would be seen only very gradually in an aggregated series on average miles per gallon.

The components approach to time series enables us to search for each of these assignable causes, while making due allowance for common cause variation. We may use both graphical and numerical procedures to identify problems, as summarized in Table 1.

Assignable cause	Graphical procedure	Numerical procedure
Temporary	Plot recursive [one-step-ahead] residuals	Shewhart chart
Level-shift	Plot recursive [one-step-ahead] residuals	Shewhart chart or Cusum charts
Seasonal	Plot seasonal component	Check variance of seasonal component
Long-Term	Plot trend component	Cusum chart

TABLE 1. Graphical and numerical procedures for the identification of assignable causes in time series.

5. The structural Time Series Model

Although terms such as ' trend' and ' seasonal' are intuitively appealing, they are mental constructs; we cannot observe them directly. Therefore, we use a structural modeling approach that treats them as *unobserved components* (Harvey, 1989; Harvey and Shephard, 1993). We used the STAMP software in conjunction with GiveWin; for details, see Koopman et al., (2000).

We define the components at time *t* as follows: trend = μ_t; slope = β_t; seasonal component = γ_t; and irregular component = ε_t. We assume that the process is observed at unit time intervals (*t, t+1,…*) and that there are *s* such intervals in a year. We then allow each component to evolve over time according to the specifications:

$$m_t = m_{t-1} + b_{t-1} + h_t \tag{1}$$

$$b_t = b_{t-1} + V_t \tag{2}$$

and

$$g_t + g_{t-1} + \dots + g_{t-s+1} = w_t \tag{3}$$

The quantities η_t, ζ_t, and ω_t represent zero mean, random shifts in the corresponding component. We assume such shifts to be independent of one another and uncorrelated over time; we also assume that they are independent of the

'irregular' component defined below. Equations (1)-(3) are known as the *state* or *transition* equations since they describe the underlying state of the process, or the transition of the components from one time period to the next.

Expressions (1) and (2) provide a very general framework for describing the evolution of the trend. If the process being modeled does not require all these components they can be dropped from the specification. The components are tested in sequential fashion as follows (Harvey, 1989, pp. 248-56):

    1T. Does the slope disturbance term have positive variance? [Zero variance corresponds to removing that term.]

    2T. If the slope disturbance is dropped, does the level disturbance have positive variance?

    3T. If the slope disturbance is dropped, does the slope differ from zero?

If all three tests produced negative outcomes the trend term would be reduced to a constant.

When the time series is seasonal, we check:

    1S. Does the seasonal disturbance term have positive variance?

    2S. If the seasonal disturbance is dropped, are the seasonal components significantly different from zero? [Is there a seasonal pattern?]

If we drop the disturbance term we are left with a "classical" model with fixed seasonals. If the seasonal pattern is rejected completely, we reduce the model purely to its trend components.

The state of the system is related to the observed series by the *observation equation*:

$$y_t = \boldsymbol{m}_t + \boldsymbol{g}_t + \boldsymbol{e}_t \tag{4}$$

where ε_t denotes the 'irregular' component. The irregular component has zero mean and is assumed to be unrelated to its own past (i.e. not predictable) and independent of the disturbances in the state equations.

Estimation proceeds by maximum likelihood (Harvey, 1989, pp. 125-128). Operational details are provided in Koopman et al. (2000, section 8.3). The key parameters are the four variances corresponding to the disturbance terms $[\sigma_\varepsilon 2, \sigma_\eta 2, \sigma_\zeta 2$ and $\sigma_\omega 2]$. Note that we assume these variances are constant over time; the time series may need to be transformed to justify this assumption, at least to a reasonable degree of approximation. The four variance terms control the form of the model, allowing each of level, slope and seasonal to be stochastic or fixed; slope and seasonal may be present or absent. Table 2 illustrates the principal variations. If fixed components are included in a model, the corresponding terms appear in the state equations (e.g. fixed seasonal coefficients) but the variance term is zero. If the components are stochastic, the same terms appear in the model, but the variance is strictly positive. The most general form is the Basic Structural Model (BSM), in which all components are stochastic. The BSM forms the starting point for the model development process, and is the standard form employed in STAMP. We then 'tested down' to eliminate any components that were not required for a particular series. An initial set of interventions (prior to September 2001) was identified using the notes provided in the original Transportation Indicators documentation, combined with an initial analysis using AUTOBOX.

6. Analysis of Airline Delays

To illustrate the proposed forecasting and monitoring techniques with respect to the impact of September 2001, we consider an example that has gained considerable publicity of late – airline delays. Figure 2, a page drawn from the February 2002 issue of the TI report, provides a summary of the on-time performance measures for the major US carriers for the past 10 years.

MAJOR U.S. AIR CARRIER ON-TIME PERFORMANCE

Major U.S. Air Carrier On-Time Performance
(monthly data, not seasonally adjusted)
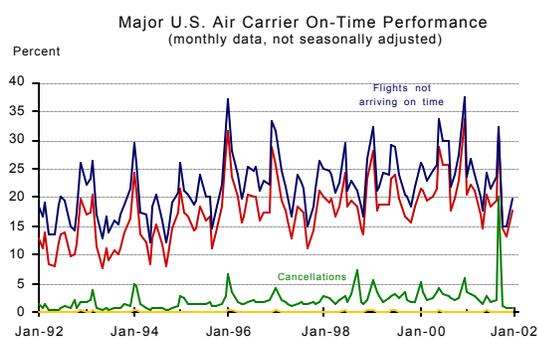
The number of flights not departing or arriving on time, cancellations, and diversions are measures of service quality.

These indicators are strongly seasonal and are affected by weather and heavy demand in winter and summer months, respectively.

On-Time Performance	Dec-00	Dec-01
Number of scheduled flights	475,398	394,787
Percent change from same month previous year	*1.16*	*-16.96*
Percent of flights not arriving on time	37.59	19.77
Change from same month previous year	*15.59*	*-17.83*
Percent of flights not departing on time	33.80	17.68
Change from same month previous year	*15.38*	*-16.12*
Percent of cancelled flights*	6.00	0.86
Change from same month previous year	*4.26*	*-5.13*
Percent of diverted flights**	0.31	0.14
Change from same month previous year	*0.12*	*-0.17*

* Also counted in flights not arriving or departing on time.
** Also counted in flights not arriving on time.

NOTES: The current value is compared to the value from the same period in the previous year to account for seasonality.

The data cover the 10 largest U.S. air carriers. A scheduled operation consists of any nonstop segment of a flight. The term "late" is defined as 15 minutes after the scheduled departure or arrival time. A cancelled flight is one that was not operated but was listed in a carrier's computer reservation system within seven calendar days of the scheduled departure. A diverted flight is one that left from the scheduled departure airport but flew to a destination point other than the scheduled destination point.

Aloha Airlines began reporting in October 2000 and is included here starting in October 2001. For comparability, the year-ago changes and growth rates are based on data that excludes Aloha. American Eagle began reporting in January 2001, will be excluded here until one year's data is available to retain comparability with previous years.

The dramatic changes in the September 2001 data reflect the impact of the terrorist attacks on September 11, 2001, on aviation, including several days in which commercial air operations were suspended.

SOURCE: U.S. Department of Transportation, Bureau of Transportation Statistics, Airline Service Quality Performance data.

FIGURE 2. Major US air carrier on-time performance (February 2002 issue).

Included in this set of measures are

- Flights not arriving on time,
- Flights not departing on time,
- Cancellations, and
- Diversions

For our analysis, we selected the percent of flights not arriving on time as the variable of interest, which we refer to as ' late arrivals.'   A graph of this single data series, starting in September 1987, is provided in Figure 3.

Prior to modeling the data in STAMP, the late arrival data were analyzed in AUTOBOX to find an initial set of interventions. Three significant pulses within the time period of September 1987 through August 2001 were found:

January 1996, and December 2000. These two interventions were incorporated into the STAMP modeling process. Our analysis of the series using STAMP revealed that the most appropriate model was one with stochastic level, no slope and fixed seasonals.
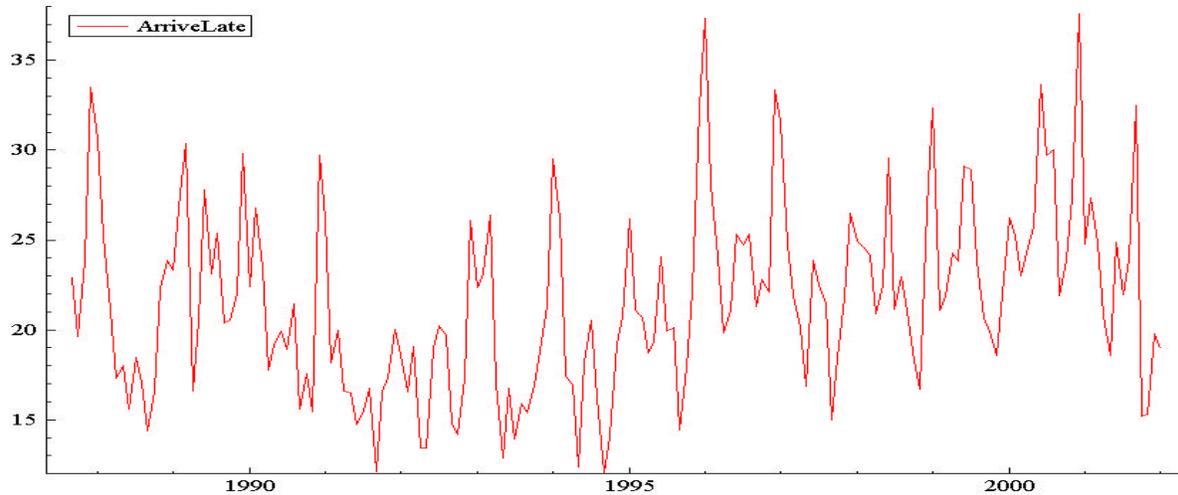
FIGURE 3. Late arrivals as a percent of total operations for major US air carriers.

This model yields the outputs shown in Figures 4-8. Figure 4 shows the smoothed trend and seasonal components; the *smoothed* versions are the better choice for gaining a perspective on the evolution of the series as the estimates use observations both before and after the time period in question.
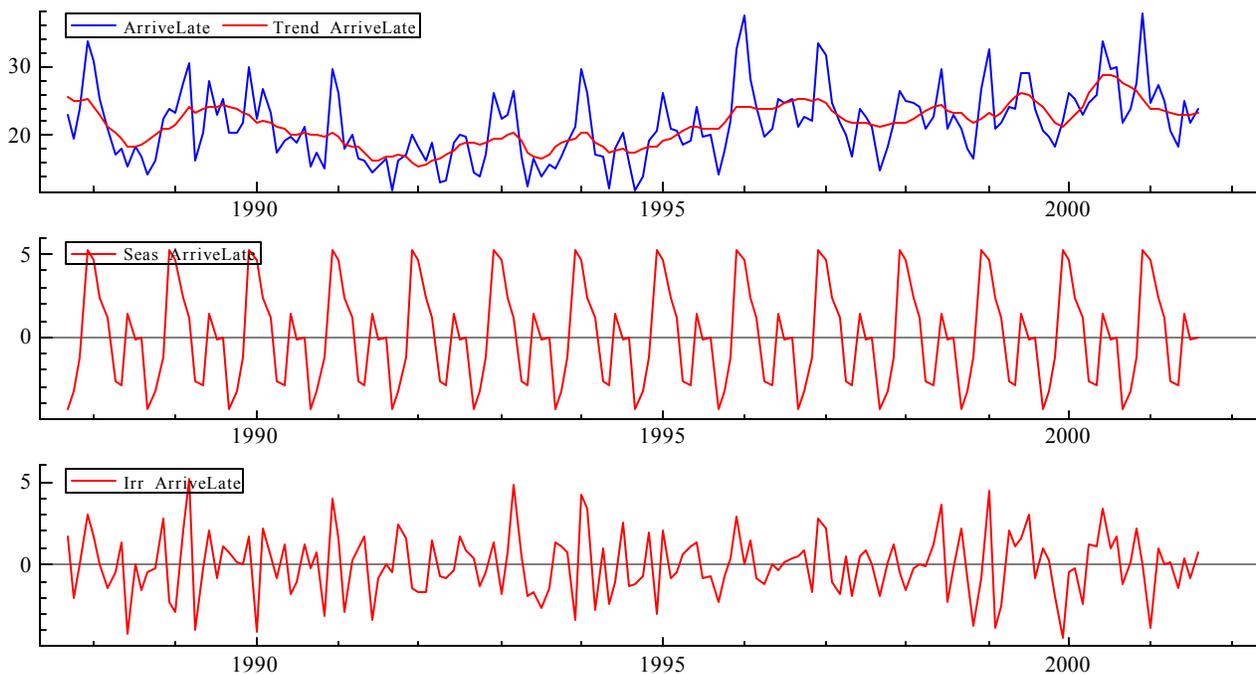
FIGURE 4. Smoothed components of the airline delays series generated by STAMP.

When these plots are compared with the *filtered* components in Figure 5, the increased roughness of the latter set becomes evident. However, the filtered components use only the observations up to the time period under investigation and are therefore more useful for monitoring purposes.
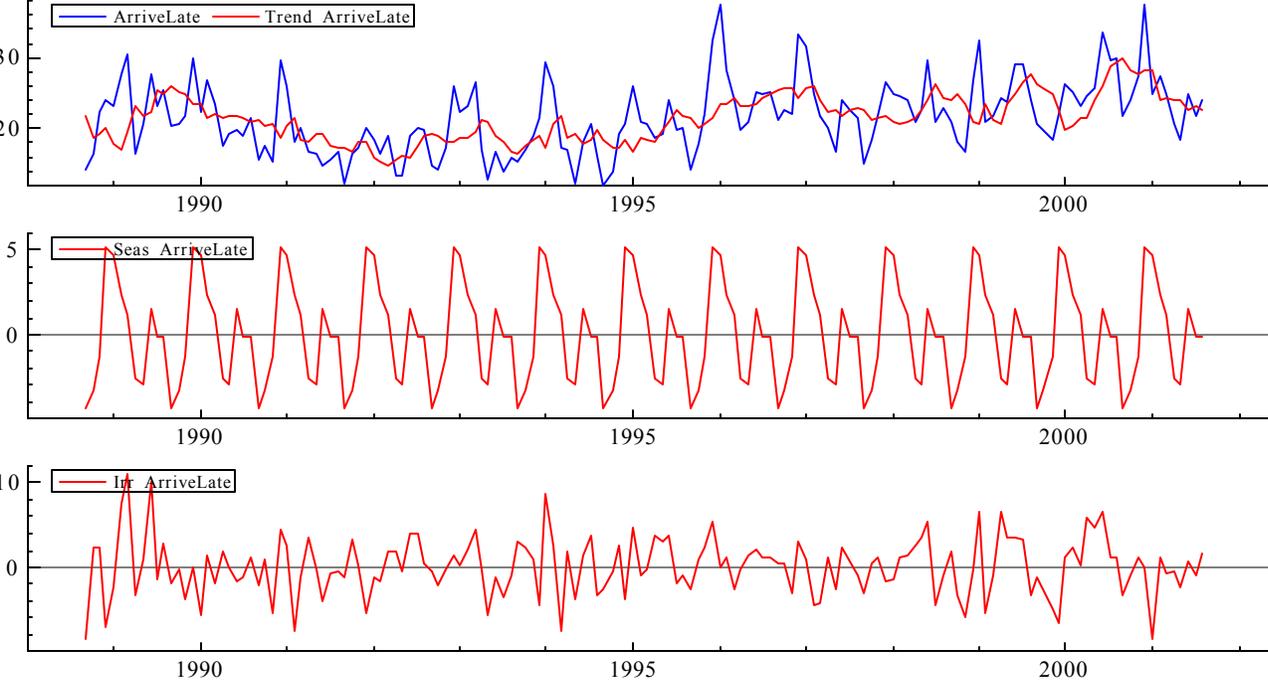
FIGURE 5. Filtered components of the airline delays series generated by STAMP.

Note that the optimal model has been based upon the data from September 1987 through August 2001. Since the model has been specified, the hold-out sample of data from September 2001 through January 2002 is placed back into the data set, and the full set of data is run through the optimal model specifications in STAMP. We can now analyze the resultant residuals with the Shewart and Cusum charts to study the impact of the newest set of data on the STAMP model.

Figure 6 shows the standardized residuals for the full fitted series and highlights the impact of post-September 11, 2001. The residuals analysis in the Shewart chart indicates a sharp rise in late arrivals in September, followed by a severe decline in late arrivals in October 2001.
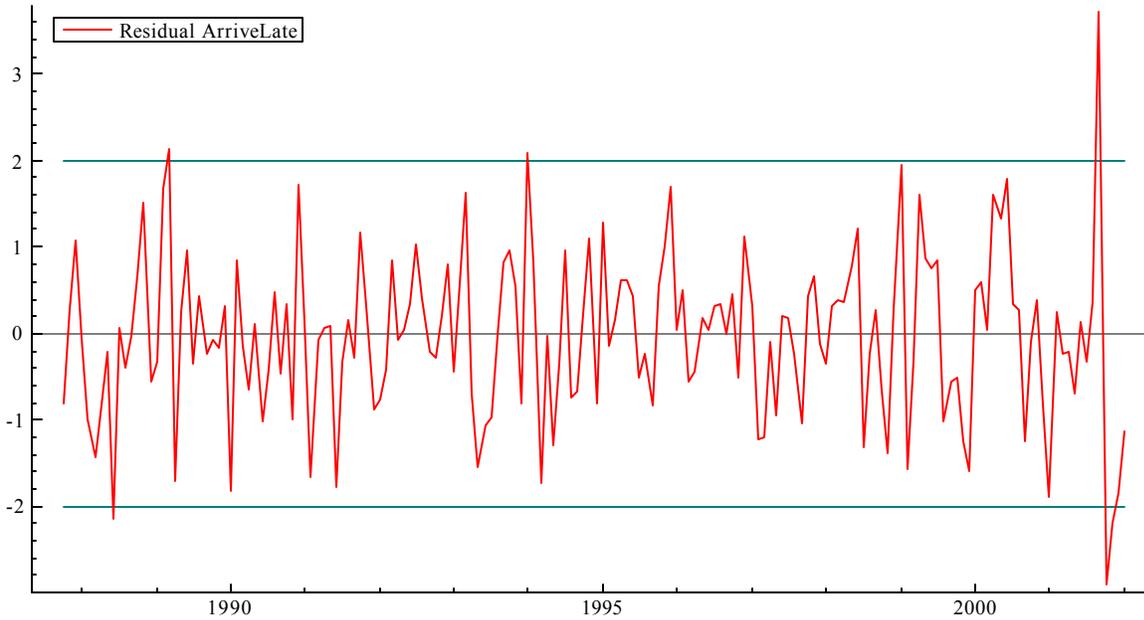
FIGURE 6. Airline delay series: Shewart chart of standardized residuals.

In order to test for long-term assignable causes, we ran a Cusum test on the residuals (see Figure 7).


FIGURE 7. Cusum chart for Percent late arrivals, resetting after each alert.

We see an alert in October 2001, which indicates a level shift at this point in time. This concurs with the October drop in the Shewart analysis. The two charts together highlight a pulse in September 2001 and a level shift in October 2001. These interventions are also noted in the STAMP results for the full data set. The final fitted underlying trend for the full set of data is shown in Figure 8.
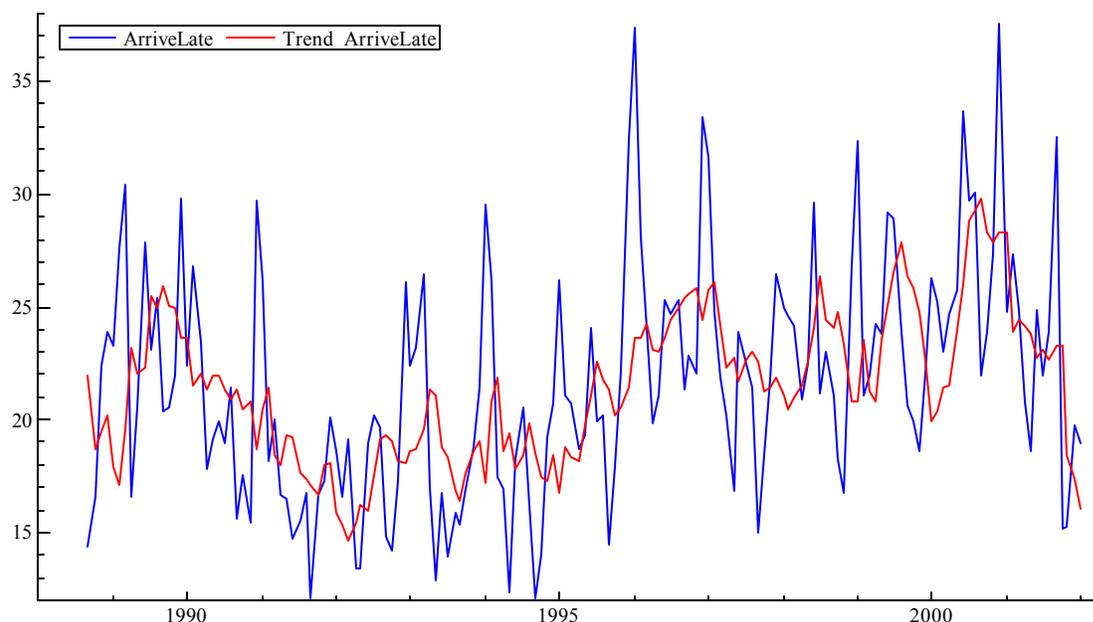
FIGURE 8. Final trend for late arrivals.

6. Final Comments

Additional datasets that felt the impact of September 2001 could not be shown in this paper, due to space limitations. But several of these indicators were shown in the FFC presentation on April 18, 2002. For a copy of the full PowerPoint presentation, please contact the lead author of this paper, Peg Young.

References

Alwan, L. C. (2000). *Statistical Process Analysis*. Irwin McGraw-Hill: Boston.

Alwan, L.C. and Roberts, H.V. (1988). Time-series modeling for statistical process control. *Journal of Business and Economic Statistics,* 16, 83-95.

Alwan, L.C. and Roberts, H.V. (1995). The problem of misplaced control limits. *Applied Statistics,* 44, 269-278.

Harvey, A.C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter.* Cambridge University Press: Cambridge, UK.

Koopman, S. J., Harvey, A.C., Doornik, J.A. and Shephard, N. (2000). *Stamp: Structural Time Series Analyser, Modeller and Predictor.* Timberlake Consultants: London, UK.

Projecting Taxpayer Behavior

Chair: Bonnie Nichols, National Endowment for the Arts

Discussant: Jeff Butler, Bureau of Transportation Statistics, U.S. Department of Transportation

Projections of Individual Income Tax Returns and the Shifts Among Forms 1040, Form 1040A, and Form 1040EZ

Andre Palmer, Internal Revenue Service, U.S. Department of Treasury

Projections of individual income tax returns to be filed by the major return types—Form 1040EZ, Form 1040A and Form 1040—are important to IRS resource planning efforts. The shorter Forms 1040EZ and 1040A require fewer resources to process than does the longer Form 1040 return. However, forecasting the respective numbers of Forms 1040EZ, 1040A and 1040 to be filed is challenging because of their erratic historical trend lines which contain numerous shifts and other irregularities. This paper examines the causes for these erratic trends, and summarizes the projection methodologies IRS staff use to deal with them.

Dealing with Uncertainty In Projections of Electronic Filing of Individual Income Tax Returns

Javier Framinan, Internal Revenue Service, U.S. Department of Treasury

This paper focuses on the methodology IRS staff use to project the number of individual income tax returns to be filed electronically. The 15-year trend in electronically filed (e-file) individual returns tends to follow the "S-shape" pattern of the new product diffusion curve. However, developments in private-sector tax preparation services coupled with changes in IRS administrative practices have repeatedly altered the underlying growth trend and have made it difficult to offer precise e-file projections. To deal with this uncertainty, IRS staff employ alternative e-file projection scenarios to provide decision makers with a fuller range of potential outcomes.

Accounts Receivable Resolution and the Impact of Lien Filing Policy on Sole Proprietor Businesses

Terry Ashley and Alex Turk
Internal Revenue Service, U.S. Department of Treasury

The federal tax lien is an important collection tool in resolving delinquent tax accounts. The lien helps to secure the government's right to the value embodied in the taxpayer's assets and provides creditors with important information on the credit worthiness of taxpayers. However, U.S. legislation, IRS policy, and IRS budgetary shocks can alter the degree to which this collection tool is used. This paper develops a model of accounts receivable resolution and uses that model to forecast the impact of changes in the number and the timing of federal tax liens filed against sole proprietor businesses with delinquent tax accounts.

This page intentionally left blank.

PROJECTIONS OF INDIVIDUAL INCOME TAX RETURNS AND THE SHIFTS AMONG FORM 1040, FORM 1040A AND FORM 1040EZ

By Andre Federal Palmer, Internal Revenue Service

Introduction

Most American adults are required to file an annual federal tax return with the Internal Revenue Service (IRS) to report their income and pay their proper amount of tax. In calendar year (CY) 2001, taxpayers filed over 129.4 million individual income tax returns. This total includes 62.4 million Form 1040 returns, 15.2 million Form 1040A returns, and 11.6 million Form 1040EZ returns, plus an additional 40.2 million returns filed via "alternative" means (i.e., electronically). On a regular basis, IRS staff from the National Headquarters Office of Research prepare updated projections of these individual income tax returns series for use in budget submissions and for other resource allocation and planning efforts within IRS.

Individual tax return projections by type of filing medium (i.e., paper versus electronic) and type of form are an important budgetary matter since IRS processing costs vary by type. In terms of traditional paper forms, the more lines of data on the return that must be transcribed by IRS employees, the more it costs to process. For example, data developed for IRS by a consulting firm for fiscal year 1999 indicated that the average direct labor cost to process a Form 1040 return filed on paper was $1.93, compared to $1.50 for a paper Form 1040A, and $1.01 for a paper Form 1040EZ.

The following paper summarizes the basic methodologies we in the IRS use to generate US forecasts of the number of individual income tax returns to be filed by type. We focus particularly on the methods for forecasting the respective volumes of paper Forms 1040, versus 1040A, versus 1040EZ, since these particular trend lines have repeatedly experienced significant "interventions" (i.e., major disruptions to their underlying historical trend patterns) over the years. We highlight tax law changes, form changes, administrative developments and other unique factors that have impacted the Form 1040 family of returns over the past several decades. We also explain the basic statistical approaches we employ to deal with these disruptions in the trends. Also, readers should note that the while the projections cited in this article are indicative of those prepared by IRS staff, the forecasts presented here are only for illustrative purposes and do not reflect official IRS/Office of Research projections-- which go through a more formal management review process and which are updated at a different time of year.

Setting the Total Individual Series Tax Return Forecasts With a 1993 Intervention

In general, we look to use econometric-based models (i.e., regression models based on economic or demographic information) when there exists a reasonably logical causal relationship between independent and dependent variables, and when long term projections are needed. Time series (extrapolation) models are preferred when the estimation trend horizons are short or when historical values contain all the information. We also turn to extrapolation models when limited data or other factors prevent us from establishing a credible regression model.

In terms of projecting individual income tax returns, our first step is to set (forecast) the total individual Form 1040 series—defined as the sum of paper Forms 1040, 1040A, and 1040EZ, plus electronically filed (e-file) returns. (E-file returns include those filed electronically through an authorized third party tax professional, on-line filed returns using commercial tax preparation software, and "TeleFile" returns.) In the case of the total individual return series, prior IRS modeling has established a clear relationship between it and various economic indicators of the U.S. economy such as total employment and personal income (lagged one year). Personal income is a logical predictor variable since the returns filed are in fact income tax returns. Total employment is an equally logical predictor variable, not only because employment generates income potentially subject to tax, but also because "employment" generally entails a situation where the individual's wages are subject to withholding. And in an employment situation involving withholding, one might need to file a tax return to claim a refund even in instances where there is no income tax liability, per se.

The historical time series data for the total individual return series is presented in Table 1, along with the corresponding year-to-year percentage change. Overall, this time series reveals a fairly "smooth" trend line that tracks the overall performance of the U.S. economy (e.g., a decline or anemic growth in return filings during periods of recession). There are few signs of interventions. This is generally to be expected since the individual return series reflects an aggregate total that is unaffected by shifts among the subordinate pieces such as medium of filing or particular paper form type.

However, there are in fact a couple interventions in the total individual series worth note. One intervention

occurred in 1988 and 1989. Filings in these two years were somewhat higher than expected (as indicated by the relatively high recorded growth rates of 3.7 percent and 2.7 percent, respectively) as a result of the impact of the Tax Reform Act of 1986 (TRA86).   While certain provisions of TRA86 eliminated filing requirements, other provisions (particularly the repeal of the personal exemption for those who could be claimed as a dependent on another's return) actually lead to an upward spike in total individual return filings. A second intervention in the total individual return series occurred in 1993. Total individual return filings in that year actually dropped by nearly one percent. While we suspect that this drop somewhat reflects a delayed effect from the recession in the early 1990's (and associated developments such as an unprecedented drop in interest rates which could have reduced income earned from savings), there were also other key factors at work. One of those factors was the culmination in 1993 of a major IRS initiative to reduce the number of individuals filing returns unnecessarily.

In terms of the specific forecasting model we pursue for the total individual return series, we test various econometric models incorporating economic variables like personal income, employment and gross domestic product, along with a "dummy variable" (i.e. step function) for the 1993 intervention. (In more recent years we have ignored the TRA86 impact since the intervention appears to follow a decayed response where the impact has slowly faded away.) We consider various model combinations, along with their comparative statistics (such as coefficients of determinations, F-tests , T-tests, P-values, Durbin-Watson values, etc.) and other characteristics such as the out-of-sample trend "nowcasts" (estimated values at the origin of the forecasts). We also examine the forecast results obtained from averaging the projections from two or more models.

An illustration of our approach is a recent effort where the methodology for individual return series trend entailed an average of two time series multiple regression models, both with the base period 1973-2001.

The first Ordinary Least Squares (OLS) model was estimated as follows:

Individual Return Series = 35,462,022 + 514,721(x1) − 1,190,671(x2) + 847,492(x3)

where

x1 = total employment in the previous year (measured in millions)

x2 = a dummy (indicator) step variable to adjust for the effects of the 1993 drop in the total return series

x3 = a time trend

Model Statistics:
  Adjusted R squared = 0.994   Significance F = 0.0001

  Parameter t-statistic P-value for x1 = 0.0019
                                     x2 = 0.1748
                                     x3 = 0.0083

  Durbin Watson = 0.85   Mean Absolute Percentage
                                     Error (MAPE) = 0.73%

The second Ordinary Least Squares (OLS) model was estimated as follows:

Individual Return Series = 67,093,726 + 3,826(x1) − 3,105,395(x2) + 1,304,056(x3)

where

x1 = chained (inflation adjusted) 1996 personal income during the prior year (measured in millions of dollars)

x2 = a dummy (indicator) step variable to adjust for the effects of the 1993 drop in the total return

x3 = a time trend

Model Statistics:
  Adjusted R squared = 0.994   Significance F = 0.0001

  Parameter t-statistic P-value for x1 = 0.0042
                                     x2 = 0.0003
                                     x3 = 0.0001

  Durbin Watson = 0.74   Mean Absolute Percentage
                                     Error (MAPE) = 0.81%

Both of the multiple regression models contain variables that have significant T statistics, i.e., P-values of less than 0.05 with confidence intervals of 95% and F-test less than 0.005. The one exception in the first model is the P-value of 0.17 for the dummy variable.   However, we were comfortable with this relaxation of the 0.05 rule of thumb because of the intuitive logic of the "intervention" this variable represented, and because of the proper (i.e., negative) sign on the coefficient parameter (indicative of the observed drop in the series). In addition, we observed that the residual values fell within their horizontal bands on their relative correlogram (pass white noise/autocorrelation test) and the both models had mean absolute percentage errors (MAPE) less than 1%.

Both models also recorded adjusted R squares above 0.99, although we note that these values are based on nominal data that were not detrended. However, we also note that

the time variable in both models served as a de facto method for detrending the data. In fact, in our experimentation we regressed the annual percentage changes in the return series data against the percentage changes in personal income, and in total employment, and got very comparable results/forecasts (albeit with lower Adjusted R Squares of approximately 0.7 and 0.4, respectively). However, we preferred the initial OLS models (with the time variable) since their resulting forecasts tended to be a slightly more conservative (lower) than the detrended models based on annual percentage change. We also elected to use an average of the two OLS models since it was a simple approach that gave us a set of forecasts that seemed intuitively sensible, and that also got us around the problem of multicollinearity—given that personal income and total employment are so highly correlated.

In terminology we use at IRS, the above-described model forecasts are considered our "baseline trends"; baseline trends that may then require further "off-model" (subjective) adjustments to account for future intervention(s) such as tax law changes that are not captured in the historical data. As a general rule, our forecasts only incorporate the effects of enacted legislation and confirmed future developments. In the case of the total individual return series, there are presently no future interventions that we are aware of that would significantly increase or decrease total filings. Hence our total individual return series is set after averaging the output of our two OLS models, and now serves as an overall "control" on all the other subordinate forecasts by filing medium and by form type--to be described in more detail below. These total individual series forecast controls are presented in Table 1. For example, for filing year 2002 and 2003, the total number of US individual returns is estimated to be 131,270,800 and 132,465,600, respectively (reflecting annual growth rates of 1.41percent and 0.91 percent). They are projected to reach nearly 143 million returns by CY 2008.

Forms 1040, 1040A and 1040EZ, and the Impact of Alternative Ways of Filing

Table 1 also presents historical time series data on paper Forms 1040, 1040A, 1040EZ and "Alternative Ways of Filing." In calendar year (CY) 1973, 78.2 million individual tax returns were filed using the paper Form 1040. In the following year, with the introduction of a simpler Form 1040A, Form 1040 filings dropped around 24 percent to nearly 59.2 million. Similarly, in CY 1983, when the Form 1040EZ became a filing option, Form 1040A filings dropped from 37.6 in 1982 to barely 21 million volumes. As stated earlier, such "form changes" reflect obvious interventions that dramatically impact the time series data by form type.

Another major development "disrupting" the trends by the three major form types (i.e., by Form 1040 versus Form 1040A versus 1040EZ) is the introduction and growth of "Alternative Ways of Filing" (AWF). For the most part, AWF reflect the introduction and growth of the various methods for filing returns electronically (e-file). However, from the early 1990's through 2000, AWF also included a highly condensed paper return, produced by special IRS-approved software, called the Form 1040PC. Still, whether an e-file return or a Form 1040PC, the effects of these AWF options were the same, viz., to reduce and otherwise alter the underlying trends in the volumes of paper Forms 1040, 1040A and 1040EZ—as taxpayers elected to use these more modern approaches to filing.

Signs of the effects of AWF options on the paper return volumes by type are revealed in a cursory review of the data in Table 1, as well as in Figure 1. For example, in 1991, alternative filings made up barely 7% of overall 1040 filings with volumes of about 7.6 million. However, by 2001, they made up over 31% or roughly 40 million tax returns. Growth in electronic filing is expected to continue, with strong Congressional encouragement such as provisions embedded in the IRS Restructuring and Reform Act of 1998. As alternative ways of filing increase (e-file is expected to reach at least 46 million in 2002), they serve to mask the underlying trends in the paper forms by type. This fact is particularly important when selecting modeling methodology because in a pure extrapolation time series, all information is assumed to be embedded in the historical values (and requiring stabilized variance and a constant mean). Also, the principal tool of econometrics, regression analysis, is hard to apply in scenarios involving the erosion of paper return filings arising from the electronic filing alternatives.

"Adjusted" Levels Form 1040, Form 1040A and Form 1040EZ

To help get at the dynamics involved in this interplay between form types and AWF options, we transform the data involved to an "adjusted level" format. This format uses analyses of the AWF returns to determine the simplest form that could have been used, had the taxpayer filed on paper. We then add those AWF returns by type to the corresponding paper counts to derive "adjusted level" figures. In effect, adjusted level data serves to negate the impacts of AWF options and helps better reveal the true historical trends in the Form 1040 type returns, versus Form 1040A type returns, versus Form 1040EZ type returns. The adjusted level data by form type is presented in Table 2. Also Figures 2, 3 and 4 contrast the "paper only" trend line versus the "adjusted level" trend for the Forms 1040, 1040A and

1040EZ, respectively. While there are still shifts and other discontinuities in the adjusted series, these adjusted data contain fewer interventions and provide a clearer picture of the underlying direction of the trends involved.

In the projection approach we use at IRS, the next components to be forecasted, after the total individual return series control is set, are the three adjusted levels (i.e., adjusted Form 1040 type, adjusted Form 1040A type and adjusted Form 1040EZ type). Later we complete estimates of e-file, by form type, and then subtract these e-file components from the adjusted levels to arrive at the final paper only forecasts. The adjusted level trends are used in modeling the form types because the transformations have the affect of helping stabilize the variances in the trends. Unfortunately, other legislative and administrative interventions are still embedded in the adjusted data, so some problems with non-stationary remain, and force us to employed other strategies to work with a more limited set of data points.

Legislative Tax Law and Administrative "Interventions"

Presented below is a summary of the major operational and legislative interventions that affect the adjusted level mix of individual returns by form type. As readers will find, there are in fact many. The statistical nature of these impacts vary, but most effects are instantaneous and step-based with the trends continuing at the new level, similar to the case of the 1993 drop in total individual return filings which was noted earlier (and associated with the IRS' "Reduce Unnecessary Filing" program). However, some of the interventions had temporary effects on the data series, which then tended to return to "steady state" such as the impact of the 1995 Revenue Protection Strategy" listed below.

The adjusted level data by form type and the associated year-to-year percentage changes are presented in Table 2. Embedded in the adjusted level data by form type for 1983 on (the year Form 1040EZ was first introduced) are the following major interventions.

- Data for 1984 through 1986 reflect the adjustment period for the introduction of the simple Form 1040EZ. In effect, taxpayers shifted from using the Form 1040A to using the Form 1040EZ—as they became aware of, and comfortable with, the simpler Form 1040EZ, and as IRS operational programs became better at identifying and encouraging those able to use it.

- The 1988 through 1990 filing volumes for all three form types were significantly impacted (changed) by the sweeping provisions of the Tax Reform Act of 1986. Provisions such as the repeal of the personal exemption for those taxpayers (primarily young people) who could be claimed as a dependent on another's (primarily parent's) return initially increased filings of the shorter Forms 1040A and 1040EZ. In addition, major changes to the rules on itemized deductions, including the gradual phase-out of the deduction for state sales taxes paid, further added to a shift from Form 1040 filings to the simpler Forms 1040A and 1040EZ. Later, the Technical and Miscellaneous Revenue Act of 1988 allowed parents to claim the unearned income of certain children on their return, starting with 1990 filings, distorting yet again the nature of the underlying trends in the Forms 1040, 1040A and 1040EZ in the wake of TRA86.

- Data for 1991 reflect a major shift from Form 1040 filings to 1040A, as a result of a form change to the latter which enabled it to accept the reporting of pension income and estimated tax payments.

- The filing experience in 1993 reflects the highly unusual drop in total individual return filings and tends to throw suspicion on the observed change for that year for all three form types.

- Data for 1994 reflect a major shift from Forms 1040 and 1040A filings, to Form 1040EZ, as a result of a form change to the 1040EZ enabling it to accept the "married, filing joint" filing status.

- IRS's 1995 "Revenue Protection Strategy" instituted a series of measures to combat refund fraud, particularly with respect to electronic filings, and contributed to a dramatic drop in the volume of e-file returns—particularly among those submitted through tax preparation professionals. As these former e-file returns were switched to paper, however, many were submitted on Form 1040 (the default paper return type for most tax practitioners)—even though a large share of these returns had characteristics of the simpler Forms 1040A or 1040EZ. This, in turn, tended to distort the recorded adjusted level results for all three forms in that year.

- The filings for 1996 again contain disruptions to the underlying trends of all three types, as the tax practitioner community adjusted to the practices instituted with the 1995 Revenue Protection Strategy and recouped a major portion of the returns filed electronically. There was also an added shift of returns from Form 1040A to 1040EZ, as a result of a form change to the latter to accept the reporting of income from unemployment compensation.

- Finally, in the more recent years, 1999 showed a shift from Form 1040EZ filings to Form 1040 and 1040A, as a result of tax law changes that introduced new education credits, and ability to deduct interest paid on certain student loans. And these law changes continued to contribute to a shift from Form 1040EZ to 1040A in the year 2000.

To illustrate the shear volume of significant interventions impacting the adjusted level data by form type, we have placed a "#" sign next to each corresponding year-to-year percentage change figure in Table 2 where such an effect occurred. As is apparent from the many "#" signs in Table 2, most of the recorded historical data at the adjusted level from 1983 to 2001 reflect interventions. Not surprisingly, such a situation significantly limits the number of traditional statistical forecasting methodologies we can apply, and requires us to employ more judgment in those processes we ultimately select.

Forecasting "Adjusted" Levels Form 1040, Form 1040A and Form 1040EZ

The adjusted level Forms 1040, 1040A and 1040EZ projection approach we currently favor uses moving average models of order 3 (3 MA)—starting with the three most recent historical years without an intervention. The methodology is summarized as follows:

1. The historical time series for each of the three adjusted level data categories were detrended by computing their year-to-year percent change to obtain stationary or a constant mean.

2. Next the models were trended using a 3 period weighted moving average (3 MA) to produce the underlying existing patterns in each data series. In other words, the trend-cycles were estimated by smoothing (averaging) the trends to reduce the random variation.

3. The 3 MA models were computed with the following weights 0.7, 0.2 and 0.1, with the most recent data

point attached the heaviest weight (inferred to have the most explanatory value). The average was computed over the projection horizon by dropping the oldest observation and including the next. The first 3 data points used to start the forecasting were the three most recent historical years without an intervention (e.g., for the Form 1040EZ, these were 2001, 1998 and 1997). The averaging moves through the data until the trend-cycles are computed.

4. The projected year-to-year growth rates were then converted to projected tax return volumes for adjusted Forms 1040, 1040A, and 1040EZ by using the respective CY 2001 actual volumes as the starting point.

5. Finally, the resulting projected trends for adjusted Form 1040, adjusted Form 1040A and adjusted Form 1040EZ returns were "forced" (further adjusted) on a year by year basis to ensure their sum equaled the results from the total individual return series projection model. This "force" was handled by leaving the Form 1040EZ forecasted trend exactly as projected by the 3 MA model, and proportionately adjusting the results from the Form 1040 and Form 1040A 3 MA models to absorb the difference needed to match the control figure for the total individual return series.

The final resulting baseline projections for the adjusted level Forms 1040, 1040A and 1040EZ are presented in Table 2. Forecasts for filing year 2002 were estimated at 77.5 million, 31.1 million and 22.6 million, respectively. They reflect a change rate of 1.79 percent, 1.59 percent and -0.09 percent, respectively, relative to last season filing.

Deriving the Volumes of Paper Forms 1040, 1040A and Form 1040EZ

The baseline forecasts of paper Forms 1040, 1040A and 1040EZ are derived by simply subtracting expected e-file return volumes by form type, by year, from the corresponding adjusted level volumes. An explanation of how e-file returns are forecasted is beyond the scope of this paper (although it is the topic of another IRS paper to be presented at the 2002 Federal Forecasters Conference). However, we note that, in general, IRS e-file forecasts are developed from models that apply the classic structure of the innovation diffusion ("S") curve. The IRS e-file methodology must also deal with interventions issues, both in past historical data and in future developments. Again, generally speaking, total e-file volumes are typically set and then used as "controls" in attempting their breakouts by form type. The breakouts of e-file by form type are derived for the AWF data by the means

noted earlier which slots the e-file returns by the simplest form the taxpayers could have filed had they filed on paper. Typically the e-file returns by form type are projected by first transforming these data into ratios, such as by shares of total e-file, or as shares the corresponding adjusted level volumes. These shares are then projecting by using simple extrapolation techniques not unlike those summarized above for adjusted level data by form type.

A representative example of an IRS forecast of total e-file is presented in Table 1 under the projected figures for Alternative Ways of Filing. The corresponding projections of e-file by form type are similarly presented in Table 2 under the AWF volumes. Subtracting the e-file components in Table 2 from their corresponding adjusted level volumes by form type in Table 2, yields our resulting baseline projections of paper Forms 1040, 1040A and 1040EZ. These projected baseline paper volumes by form type derived by subtraction are, in turn, the corresponding forecasts presented in Table 1. If there were known future interventions to impact this mix of paper Forms 1040, 1040A and 1040EZ, such as another planned change in the line items reflected on a particular form, or fall out from an enacted piece of legislation, it would be against these baseline trends that we would apply any additional needed off-model adjustments.

Summary

This paper examined in some detail the various methodologies used by the IRS staff to generate forecasts for the total individual return series, and the associated detail by Forms 1040 versus 1040A versus 1040EZ. Forecasts in this area are particularly important to IRS and to Congress since the associated costs of processing these returns do vary by type, and require a substantial allocation of federal funds. However, preparing projections of individual returns by form type is a complicated matter since, historically, this individual return series and its components have experienced numerous interventions stemming from a diverse array of administrative developments and tax law changes.

As a result, we employ a combination of forecasting techniques to merge irregularities into the expected future trends. Principally, ordinary least squares regression models, based primarily on total employment and personal income, are used to set the overall total volume of individual income tax returns to be filed. Then, a series of data transformations are used on the various subordinate pieces of individual returns to help negate the intervention impact of alternative ways of filing (i.e., primarily that of electronic filing) and to more clearly reveal the remaining series of interventions—which still remain substantial. Finally,

an inter-connected web of trend extrapolation models are built for the various components of the individual tax returns involved, by filing medium and by major form type, and combined in a manner that ensures consistency with larger aggregate components and ultimately with the overall projected total individual returns series. The extrapolation models for the component pieces we select are those we feel best capture the trends reflected in the most recent years, but which also isolate the core underlying growth patterns from those of nominal changes merely arising from the many unique interventions.

References

Armstrong, J. Scott, *Principles of Forecasting: A Handbook for Researchers and Practitioners*, Kluwer Academic Publishers, Norwell, MA, 2001

Makridakis, Spyros, Steven C. Wheelwright, and Rob J. Hyndman, *Forecasting: Methods and Applications*, Third Edition, John Wiley & Sons, Inc. New York, New York, 1998

United States Department of Treasury, Internal Revenue Service, *The IRS Research Bulletin*: 1990 Update, IRS Publication 1500 (revised 8-90).

United States General Accounting Office, *Tax Administration: Electronic Filing's Past and Future Impact on Processing Costs Dependent on Several Factors*, GAO-02-205, (January 2002).

Note

The views expressed in this article represent the opinions and conclusion of the author. They do not necessarily represent the opinion of the Internal Revenue Service.

Table 1. Illustration of Historical and Projected Individual Tax Return Filings by Processing Form Type *

	Year	Total Individual 1040 Series Returns	Yr-to-Yr % change in Individual Series	Paper Only Form 1040	Paper Only Form 1040A	Paper Only Form 1040EZ	Alternative Ways of Filing **
Actual	1973	78,216,100		78,216,100			
	1974	81,381,000	4.05%	59,176,100	22,204,900		
	1975	83,905,000	3.10%	61,484,000	22,421,000		
	1976	86,307,000	2.86%	57,314,000	28,993,000		
	1977	85,643,000	-0.77%	57,774,000	27,869,000		
	1978	87,363,100	2.01%	58,375,100	28,988,000		
	1979	90,727,100	3.85%	56,559,100	34,168,000		
	1980	93,052,300	2.56%	54,885,000	37,571,000		
	1981	94,014,000	1.03%	57,089,000	36,925,000		
	1982	95,420,000	1.50%	57,801,000	37,619,000	-	
	1983	95,541,300	0.13%	59,580,100	21,045,300	14,915,900	
	1984	96,496,900	1.00%	61,185,900	19,590,100	15,720,900	
	1985	99,528,900	3.14%	64,010,100	18,779,100	16,739,800	
	1986	101,750,900	2.23%	66,341,700	18,421,600	16,987,500	
	1987	103,250,700	1.47%	68,490,400	17,911,800	16,848,500	-
	1988	107,028,900	3.66%	70,471,800	17,645,100	18,378,600	533,500
	1989	109,868,300	2.65%	70,511,100	18,713,100	19,483,600	1,160,500
	1990	112,305,000	2.22%	70,284,400	18,380,100	19,436,400	4,204,200
	1991	113,829,200	1.36%	67,118,700	21,784,900	17,358,400	7,567,100
	1992	114,718,800	0.78%	64,608,400	20,398,200	17,185,500	12,526,700
	1993	113,754,400	-0.84%	59,893,000	20,084,600	16,520,400	17,256,400
	1994	114,683,400	0.82%	60,420,100	18,902,300	17,161,900	18,199,200
	1995	116,059,700	1.20%	62,797,100	20,460,800	18,088,000	14,713,900
	1996	118,362,500	1.98%	60,274,300	19,003,500	17,099,600	21,985,100
	1997	120,342,500	1.67%	61,019,800	17,336,200	14,452,300	27,534,300
	1998	122,546,900	1.83%	61,346,900	16,361,900	12,755,200	32,083,000
	1999	124,887,100	1.91%	61,622,900	15,589,100	11,833,500	35,841,600
	2000	127,097,200	1.77%	60,669,900	14,743,800	11,145,600	40,537,800
	2001	129,444,900	1.85%	62,449,800	15,212,600	11,575,800	40,206,800
Projected	2002	131,270,800	1.41%	60,688,400	13,828,900	10,753,500	46,000,000
	2003	132,465,600	0.91%	57,796,200	11,858,800	9,910,600	52,900,000
	2004	134,565,600	1.59%	55,666,400	10,475,300	9,040,900	59,383,000
	2005	136,913,100	1.74%	53,327,000	10,001,400	8,185,700	65,399,000
	2006	138,937,400	1.48%	50,596,300	9,940,200	7,550,100	70,850,800
	2007	140,862,400	1.39%	48,169,600	9,812,800	7,257,600	75,622,400
	2008	142,775,600	1.36%	46,247,700	9,796,900	7,019,900	79,711,200

 * Projections are for the illustrative purposes of this article only; they should not be interpreted as official IRS forecasts.
** Alternative Ways of Filing include electronically filed returns, Telefile returns and Form 1040PC, in applicalble years.

	Year	Adjusted Level Individual Returns by Form Type **						Alternative Ways of Filing (AWF) by Form Type ***		
		Adjusted Form 1040		Adjusted Form 1040A		Adjusted Form 1040EZ		Form 1040	Form 1040A	Form 1040EZ
		Number	Yr-to-Yr % Change	Number	Yr-to-Yr % Change	Number	Yr-to-Yr % Change	Type	Type	Type
Actual	1983	59,580,100		21,045,300		14,915,900				
	1984	61,185,900	2.70%	19,590,100	-6.91% #	15,720,900	5.40% #			
	1985	64,010,100	4.62%	18,779,100	-4.14% #	16,739,800	6.48% #			
	1986	66,341,700	3.64%	18,421,600	-1.90% #	16,987,500	1.48% #			
	1987	68,490,400	3.24%	17,911,800	-2.77%	16,848,500	-0.82%			
	1988	70,639,800	3.14% #	17,965,900	0.30% #	18,423,200	9.35% #	168,000	320,800	44,600
	1989	70,846,800	0.29% #	19,342,600	7.66% #	19,678,900	6.82% #	335,700	629,500	195,300
	1990	71,372,000	0.74% #	20,717,100	7.11% #	20,216,000	2.73% #	1,087,600	2,337,000	779,600
	1991	68,682,000	-3.77% #	26,309,300	26.99% #	18,837,900	-6.82% #	1,563,200	4,524,400	1,479,500
	1992	68,373,800	-0.45% #	27,003,200	2.64%	19,341,900	2.68%	3,765,300	6,605,000	2,156,400
	1993	65,011,900	-4.92% #	29,236,400	8.27% #	19,506,200	0.85% #	5,118,900	9,151,700	2,985,800
	1994	65,186,200	0.27%	28,586,700	-2.22% #	20,910,500	7.20% #	4,766,100	9,684,400	3,748,600
	1995	67,263,400	3.19% #	27,666,000	-3.22% #	21,130,300	1.05% #	4,466,300	7,205,200	3,042,300
	1996	66,849,900	-0.61% #	28,029,000	1.31% #	23,483,700	11.14% #	6,575,600	9,025,500	6,384,000
	1997	69,210,400	3.53%	27,609,300	-1.50%	23,522,800	0.17%	8,190,600	10,273,100	9,070,500
	1998	70,774,200	2.26%	28,025,400	1.51%	23,747,300	0.95%	9,427,300	11,663,500	10,992,100
	1999	73,095,700	3.28% #	28,349,200	1.16% #	23,442,300	-1.28% #	11,472,800	12,760,100	11,608,800
	2000	74,501,600	1.92%	29,837,200	5.25% #	22,758,400	-2.92% #	13,831,700	15,093,400	11,612,700
	2001	76,169,600	2.24%	30,614,900	2.61%	22,660,500	-0.43%	13,831,100	15,375,700	11,000,100
Projected	2002	77,530,900	1.79%	31,100,500	1.59%	22,639,300	-0.09%	16,842,600	17,271,700	11,885,800
	2003	78,416,800	1.14%	31,422,200	1.03%	22,626,700	-0.06%	20,620,500	19,563,400	12,716,100
	2004	79,940,200	1.94%	32,007,600	1.86%	22,617,800	-0.04%	24,273,800	21,532,200	13,576,900
	2005	81,646,700	2.13%	32,659,400	2.04%	22,606,900	-0.05%	28,319,700	22,658,100	14,421,200
	2006	83,122,500	1.81%	33,218,600	1.71%	22,596,300	-0.05%	32,526,200	23,278,400	15,046,200
	2007	84,527,600	1.69%	33,748,900	1.60%	22,585,800	-0.05%	36,358,000	23,936,100	15,328,300
	2008	85,925,300	1.65%	34,275,000	1.56%	22,575,300	-0.05%	39,677,600	24,478,200	15,555,400

Table 2.   Underlying Composition of Individual Returns and Alternative Ways of Filings by Form Type *

* Projections are for the illustrative purposes of this article only; they should not be interpreted as official IRS forecasts.
** "Adjusted Level" counts reflect total individual returns by approximate form type had Alternative Ways of Filing not existed.
*** Alternative Ways of Filing includes all electronically filed returns, TeleFile returns and Form 1040PC volumes.
 # indicates years impacted by interventions such as law changes, changes to tax forms, or other unique administrative developments.

Figure 1. Alternative Ways of Filing


Form 1040 Individual Returns

Percent of Total Returns

■ Alternative Ways of Filing
■ Paper O/A/EZ

Calendar Year
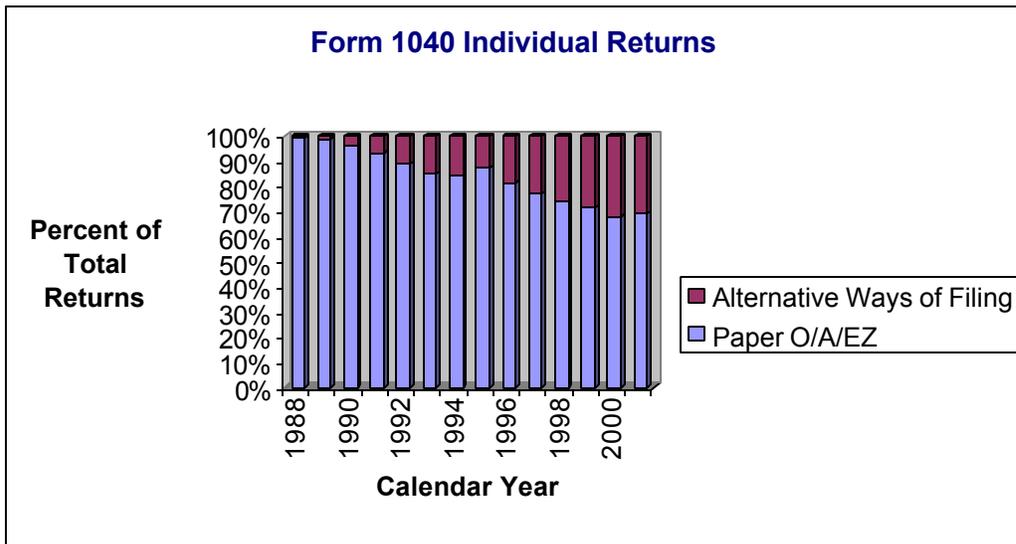
Figure 2. Form 1040 Paper Only and Adjusted Level Returns
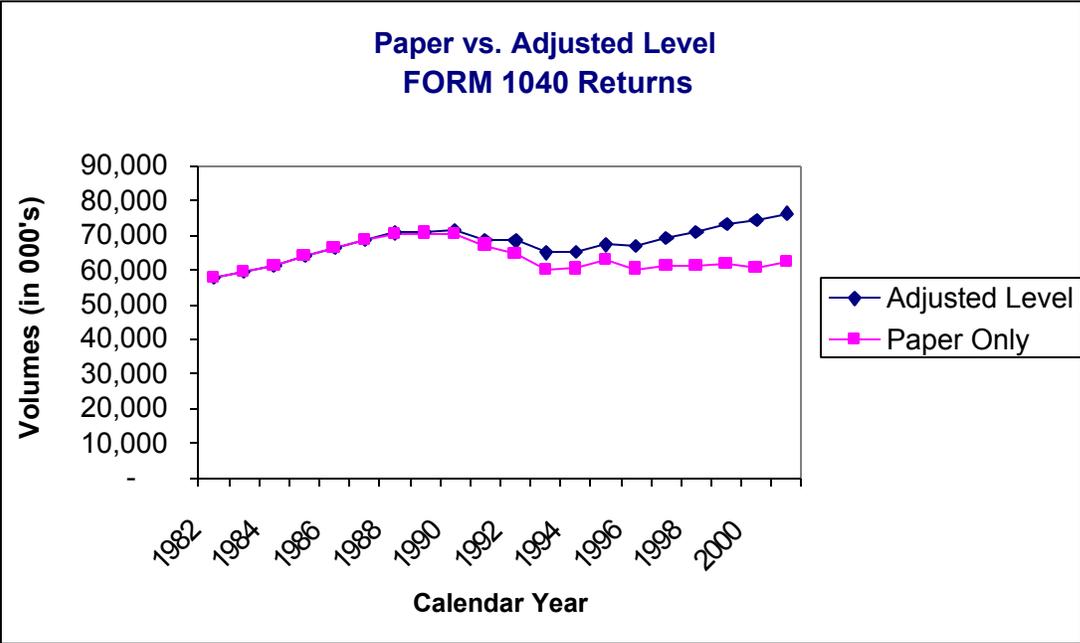
Figure 3. Form 1040A Paper Only and Adjusted Level Returns
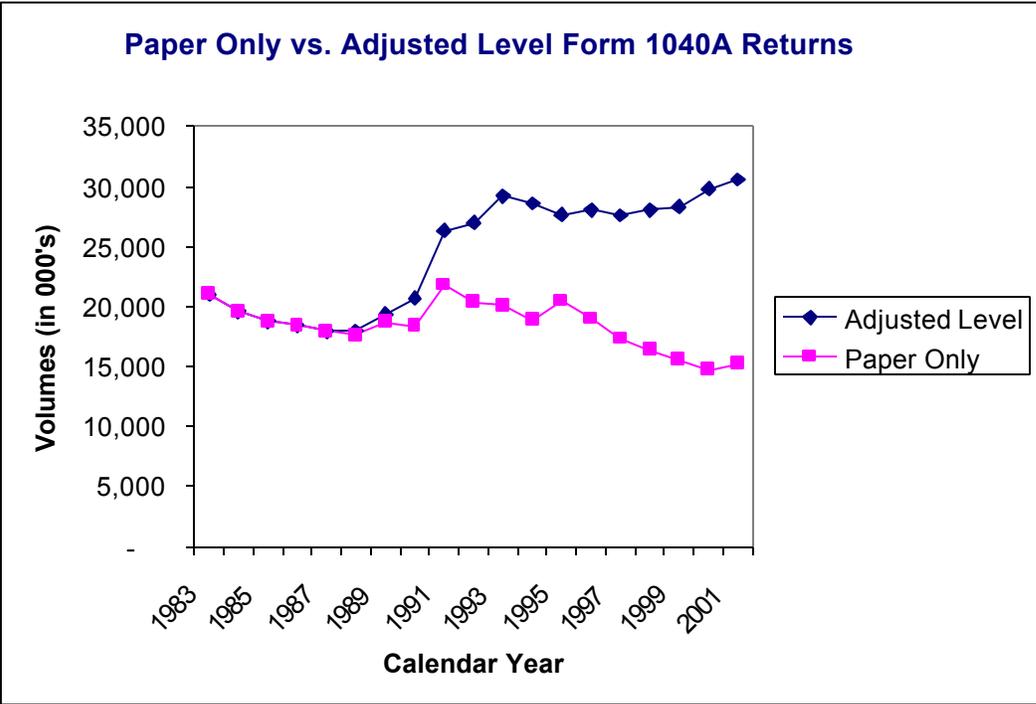
**Figure 4. Form 1040EZ Paper Only and Adjusted Level Returns**
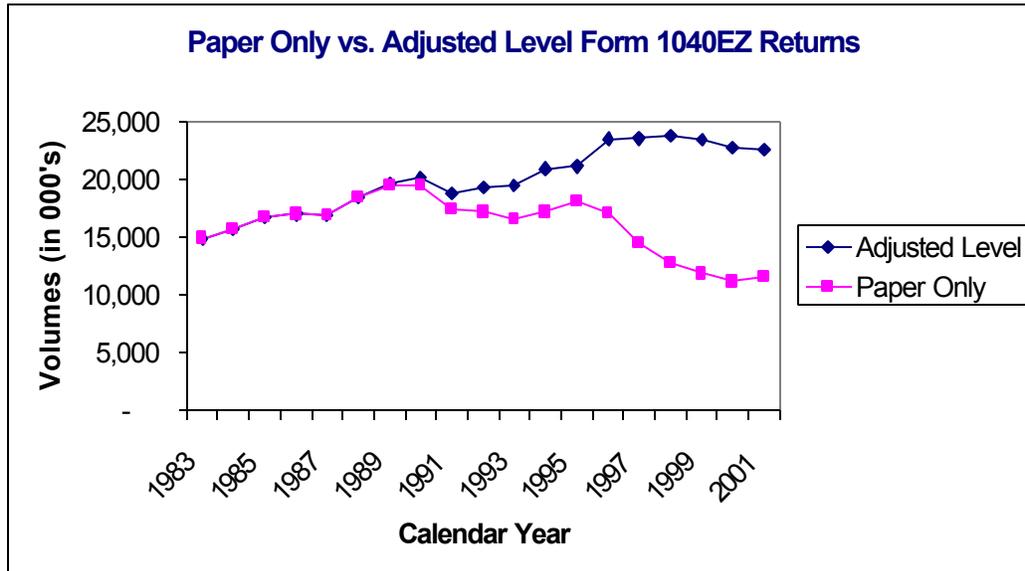
Paper Only vs. Adjusted Level Form 1040EZ Returns

DEALING WITH UNCERTAINTY IN PROJECTIONS OF ELECTRONIC FILING OF INDIVIDUAL INCOME TAX RETURN
Javier Framinan, Internal Revenue Service, U.S. Department of Treasury

Introduction

The Internal Revenue Service's National Headquarters Office of Research produces projections of various internal workload items used for resource allocation and planning purposes. A major workload item that commands a large share of IRS resources is the processing of tax returns; and an important component of that is the individual income tax return.

In 1986, the IRS introduced electronic filing as an alternative method to filing paper individual income tax returns. The intent of this new filing system was to provide convenience to the taxpayer and reduce IRS processing costs, among other benefits. By 1998, the IRS received 20 percent of its individual income tax returns electronically. In that same year, as part of the IRS Restructuring and Reform Act, the IRS and Congress set a goal of 80 percent electronic filing by 2007. (A 2000 Booz-Allen & Hamilton, *IRS Cost of Processing Electronic Returns*, estimated that the Service could reduce processing costs by between $27 million and $243 million by 2007, depending on the volumes filed electronically.) Since then, IRS projections of individual e-file have served congressional and IRS policymakers to track progress in meeting the 80 percent e-file goal. Also since 1998, the Electronic Tax Administration (ETA) within IRS has submitted to Congress an annual "Strategy for Growth" report to communicate the progress to date and future plans in promoting electronic transactions between the Service and taxpayers. These reports include historical e-file volumes and the most current IRS Office of Research forecasts.

The Need for and Use of e-file Forecasts

As e-file grows, reliable projections of its volumes become increasingly important to IRS resource allocation planners. Projections are used for e-file processing and customer service capacity planning. But more importantly (from a resource perspective), IRS wants to know the reciprocal impact on paper return volumes. Paper return and e-file processing demand different systems and resources. For each electronic return filed, there is one less labor-intensive paper Form 1040, Form 1040A or Form 1040EZ to process. There is a geographic dimension, as well. Paper return processing currently takes place in ten centers around the country, with the distribution determined by pre-established state-to-center alignments. Therefore, forecasting e-file volumes by the filer's state of origin

and 'paper form type' is critical to the paper return volume forecasting.

The e-file projections provided in the annual "Strategy for Growth" report to Congress presently do not present a scenario in which e-file reaches the 80 percent participation rate goal set for 2007. Rather, they have conveyed the shortfall expected given the current e-file environment and conditions. However, the IRS recognizes the uncertainty associated with forecasting electronic filing volumes. The general novelty of electronic commerce, an ever-changing tax preparation software industry, varying rates of adoption among professional tax practitioners, and changes and developments within the IRS relative to e-file promotion and its ability to process electronic filing have made e-file forecasting a challenging proposition. Following is a discussion of the general approach taken by the IRS to produce objective forecasts that meet the needs of its various customers, and how the IRS handles the uncertainties in those forecasts.

E-file Projection Methodology – Three Markets Considered

Individual return electronic filing has three distinct markets that require separate attention when forecasting: practitioner e-file, on-line filing, and TeleFile (Framinan, 1999). Each program began at a different time, displays a unique growth history (see Table 1 and Charts 2 through 4), and appeals to a distinct taxpayer 'market segment.'

On-Line Filing

It is easiest to begin with the most recently introduced e-filing method, on-line filing, as its growth has followed a smooth and, to date, somewhat predictable pattern. Unlike practitioner e-file and TeleFile, no major shocks or definition issues have affected its historical filing volumes outside the initial pilot years.

On-line filing refers to electronic filing of self-prepared (i.e., not professional tax-practitioner-prepared) returns. To file on-line, the taxpayer must have a computer, modem, and tax preparation software from an IRS certified private vendor. On-line filing also requires use of an IRS-accepted on-line service company or transmitter to translate the return information into an IRS readable format. On-line filing basically has the same incentives and restrictions as tax practitioner electronic filing – i.e., faster refunds, higher accuracy,

and IRS confirmation of receipt – but avoids the tax-practitioner's preparation fees.

On-line filing has grown with the proliferation of personal computers and the popularity of tax preparation software. In fact, it has experienced truly explosive growth, as evidenced by the information in Table 1 and Chart 2. In 1996, in its second year of existence, 158,000 taxpayers participated in the on-line filing program. By 1998, 942,000 were filing using this method; and by 2001, 6.8 million were. The growth pattern to date is following that of a typical product innovation diffusion, or "S," curve as depicted in Chart 1.

The "S" curve growth pattern typifies the historical usage/purchase pattern of many innovative consumer products, such as the automobile, the refrigerator, and more recently the personal computer. Adoption is slow at first, explodes, and finally slows as the market's saturation point is reached. The marketing industry uses a variety of labels to describe the different segments of the curve, and the distinct groups of consumers and overall consumer behavior it represents. For example, the beginning of the curve shows initial adoption by a few "innovators." Moving right along the curve follows the progression to "early adapters," then to an "early majority," and finally to a "late majority" and "laggards." Chart 2 shows that the adoption of on-line filing is following this "S" curve pattern.

A practical way to model the on-line filing market growth is to express the volumes in terms of participation rates. Considering participants as a percentage of the total number of potential, or eligible, filers, we defined the following on-line filing participation rate ratio:

> PR = [number of on-line filed returns] / [population of self prepared returns belonging to taxpayers that own a personal computer and have internet access]

Since there is no ready source of information for the unique market reflected in the denominator of the above ratio, we have to estimate that component. To do so, we used U.S. Census, Forrester Research, Inc., and other sources of historical and projected data on the number of U.S. households with internet access. We combined this external data with internal tax return data on the number of filers that self prepare, and made certain other assumptions, to arrive at the denominator.

To model and forecast this participation rate at the U.S. level, we used the following two-parameter-bounded logistic growth function.

$$PR(t) = u / [1 + e^{(a - b*t)}]$$

where
t = time (in calendar years)
PR(t) = participation rate at time t
u = participation rate ceiling (predetermined)
e = 2.7182 (power series expansion)
a = scale parameter
b = shape parameter

Our first step was to assume a "ceiling" participation rate u at some point in the future that by definition can not exceed 100 percent (1.0). Using survey data information, we set this upper bound, u, at 0.66. Information from a 1999 Council for Electronic Revenue Communication Advancement (CERCA) survey suggests that over 60 percent of the population of eligible e-filers stated they would file electronic returns provided the removal of all perceived barriers (including costs). We found corroborating empirical data, albeit not related to tax return filing, relative to the use of automatic teller machines (in existence since the 1970s) by individuals with bank accounts. According to a 1996 American Bankers Association and Gallup Consumer survey, 66 percent of bank customers in the U.S. had an ATM card. Similar surveys in 1993 and 1994 showed a 60 to 66 percent rate, suggesting a plateau (i.e., a "natural" ceiling) had been reached after participation growth through the 1970s and 1980s.

Given the assumed e-file participation rate ceiling u = 0.66 and the historical values of PR(t) from 1996 through 2001, we then selected values for a (the scale parameter that moves the curve up, down, left, and right) and b (the shape parameter that determines the steepness of the curve) such that the root mean squared error of the fitted values was minimized. Using a SAS grid search, we found the best fit at a = -4.0 and b = 0.6. These parameters produced fitted historical and projected on-line participation rate PR(t) values. We then multiplied the forecasted participation rates by the forecasted eligible pool of filers (supplied by the U.S. Census and Forrester, Inc. sources) to produce the nominal value return volume forecast.

Practitioner e-file
The first form of electronic filing offered to taxpayers was "practitioner e-file." From its introduction in 1986, through 1991, electronic tax return filing was possible only when done through a professional tax return preparer. In 1986 the IRS coordinated with the tax practitioner community, and set technical and

procedural standards for electronic information exchange. Although taxpayers incurred additional fees filing electronically, faster refunds provided incentive for many to e-file. By 1992, after only two years of nationwide implementation, almost 11 million taxpayers were filing electronically through a preparer.

However, as Chart 3 shows, the historical growth of practitioner e-file is not as smooth and clean as that of on-line filing. Two general considerations (i.e., interventions) help explain this messier historical pattern. First, practitioner e-file represents an aggregation of two (and ultimately more) distinct taxpayer market segments that were tapped separately and to a large degree consecutively. This has created a curve actually composed of two S curves, one on top of the other. The second consideration is the drop in volume experienced in 1995. This drop resulted from a shock to the system, precipitated by the IRS Revenue Protection Strategy's elimination of the direct deposit indicator (DDI) and the resulting reduction in refund anticipation loans (RALs) that attracted many to e-file.

What are the DDI and RAL? First, it is necessary to understand the characteristics of the taxpayers, as well as the practitioners participating in practitioner e-file, in order to explain and model practitioner e-file. From e-file inception in 1986 through 1994, taxpayers e-filing through practitioners could be characterized as young, lower income, simple return filers motivated by getting a fast refund and willing to pay relatively high preparation and RAL fees. The refund anticipation loan was an integral component of the e-file product. For an extra fee (in addition to the return preparation and transmission charges), the preparer, in coordination with a lending institution, would advance the anticipated tax refund amount to the electronic filer. The filer, in turn, agreed to have their refund deposited directly to the lending institution account. In order for this arrangement to work, the lending institution depended on the IRS providing it information, in the form of a "direct deposit indicator," to show whether the taxpayer was in fact due a refund. The arrangement gave the taxpayer an instant refund upon tax return transmittal and the lending institution direct deposit of the refund from the IRS. For the taxpayer, there were no up-front charges, as the return preparation, electronic transmission, and RAL fees all could be deducted from the RAL.

Practitioners, lending institutions, and taxpayers each had incentive to participate, resulting in fast growth in the e-file program. Chart 3 shows e-file penetration in this market segment slowing and approaching saturation by 1994, where the growth curve becomes flatter. However, refund fraud also grew in this electronic

filing-DDI-RAL arrangement. Reacting to the growth in fraud, the IRS implemented its Revenue Protection Strategy in 1995, that included more information verification and security checks and eliminated the DDI. This, in turn, caused a drop in RALs and e-file volumes.

Despite the 1995 setback, both the practitioner community and IRS continued to promote and improve the e-file program, while cooperating on efforts to combat refund fraud. Both parties recognized there were large, untapped taxpayer markets to be attracted, including market segments not necessarily attracted by the refund anticipation loans. The IRS in particular wanted to promote growth among the higher income, more complex return taxpayers, as this was where the largest savings could be realized from a paper return processing perspective. The practitioner community also contributed to this marketing, trying to attract the higher income, more complex return filers, and even the balance due filers. Practitioners reduced electronic filing fees, or packaged them differently, to attract this new market. As a result of the IRS and industry's promotional work and program improvements, renewed, sustained, and strong growth has occurred from 1996 to the present.

Methodologically, the shock associated with the Revenue Protection Strategy is easy to address relative to the construction of a forecasting model. However, on a practical level, it is difficult for the IRS to segregate and produce historical and forecasted return volume counts for the two practitioner e-file market segments identified in the above discussion. A good way to approach the practitioner e-file projections would be to consider the market segments separately. That is, we independently could model and forecast "group 1" practitioner e-file (made up of lower income, simple returns, many taking RALs), and "group 2" practitioner e-file (made up of the subsequent, higher income, more complex returns, some of which are balance due). However, splitting the practitioner e-file return counts is not readily feasible using internal IRS data, nor is there a supportive basis backed by research on how to draw the lines along these dimensions. Consequently, IRS has modeled and forecasted the aggregate practitioner e-file.

However, a geographic component to practitioner e-file participation does help alleviate the dilemma. That is, the two practitioner e-file markets are somewhat segregated along geographic boundaries. In general, the southeast region of the country was the e-file hotbed in the early years, where e-filers showed all the characteristics of the "group 1" market mentioned above. The western region of the U.S. was very slow to adopt e-file, but showed stronger growth in later years

and displayed characteristics more associated with "group 2."

IRS collects data on and makes projections for 64 IRS local geographic components (formerly known as "districts") that largely match state configurations. Forecasting practitioner e-file starts with this state-level information and builds to a national level forecast by a bottom-up approach. We developed forecasting models for each state, and summed the output to generate the national projections. To some (not well known) degree, each state is dominated by one of the market segments. Thus, fitting a two-parameter-bounded logistic growth function similar to the one used for modeling on-line filing has produced good results on a by-state level in terms of fit and reasonableness of forecasts. Summation of the state level forecasts to the national level also has produced good results, in terms of performance.

Just as for the on-line filing "S" curve model, modeling *participation rates* for practitioner e-file performed best, using the two-parameter-bounded logistic growth function model. In this case, the participation rates were defined as the ratio of return volumes filed electronically by practitioners to the total number of individual returns. The denominator includes all individual returns – i.e., practitioner prepared combined with self-prepared. We could argue the self prepared population, or definitely some portion of it, would never cross to practitioners. But we use this denominator for practical reasons, as we already produce projections of the total volume of individual returns at the state level for other customers. To account for the inflated denominator, we proportionally adjust the expected participation rate ceiling parameter, u, downward using each state's paid practitioner usage rate. For instance, approximately 55 percent of the individual income tax returns in the state of Alabama are filed through a paid tax practitioner. We multiplied this rate by the national ceiling rate (i.e., 0.55 x 0.66) to set u for Alabama at 36 percent.

Just as for on-line filing, we multiplied the projections of the participation rates by the projected denominator (i.e., total individual filings, in this case) to derive the practitioner electronic return volumes, but for each state. By this bottom-up methodology we have nominal return forecasts by state. We then summed the results to get the national volumes.

TeleFile
TeleFile employs touch-tone telephone technology to transmit returns by letting qualified taxpayers use IRS-issued customer service numbers for authentication. The IRS has limited TeleFile's availability to filers of simple returns (i.e., Form 1040EZ). This filing option caters to taxpayers that want non-paper filing, but are

unwilling to use a preparer and/or unwilling to pay transmission fees. Every year since its nationwide introduction in 1996, the IRS has mailed approximately 25 million TeleFile tax packages to taxpayers identified as eligible. In 1998, almost 6.0 million taxpayers filed this way, making up 24 percent of the individual return e-file market. However, nominal return volumes and participation rates have fallen since 1998. Some of the reasons for TeleFile's recent decline include:

- less IRS advertising and promotion,
- fewer TeleFile tax packages being mailed (due to streamlining efforts),
- problems in contracting TeleFile package printing and mail-outs, and
- loss of "cutting edge" appeal to the younger TeleFile market.

Research (see references in Framinan, 1999) suggests TeleFile achieved a 25 percent participation rate plateau under its optimal conditions. This number represents the ratio of actual users to all those eligible. Within the current, less than optimal environment that has existed since 1998, a new "natural" plateau may be establishing itself at around 20 percent. With this assumption, we employed a simple downward time trend model for each U.S. state, based on the nominal TeleFile volumes. In some cases we reduced the historical base period of the equations to reflect the current trend more accurately. We then summed these state level projections to the U.S and other subnational levels as needed.

Advantages of the Diffusion Pattern Approach
For the most part, the e-file forecasting techniques described above represent a departure from the traditional ones used to forecast IRS return volumes. Until recently, the IRS Office of Research often used more conventional approaches to e-file forecasting that included regression analysis and ARIMA modeling. In using those techniques, we produced linear models whose output then required subsequent "off model" adjustments to account for known future interventions to the e-file time trends. These interventions included such things as the introduction of the electronic signature, free on-line filing offers by a few software developers and financial service companies to certain taxpayers, and various IRS/legislative initiatives that resulted in increased refunds (and therefore increased taxpayer propensity to e-file). Estimation of the effects of these initiatives created a lot of work; and quantifying the interaction between the interventions was particularly difficult.

Moving to the "S" curve modeling methodology, which relies on the implicit assumption that improvements and promotion represent the natural course of consumer

product (or service) diffusion, simplified the task. Short of a dramatic new development, we assume the "S" curve implicitly captures the effects of the smaller refinements and expanded advertising efforts that further promote the e-file program. It has had the added benefit of a "common sense appeal" to the projection customers. The S curve evolution from early adapter, through majority, to laggards, with tapered growth at the end, has both an intuitive and historical backing.

Scenarios to Handle Uncertainty in the e-file Forecasts

Despite the new approach, our e-file projection accuracy suffers relative to that of more traditional return types, as a consequence of its novelty. Though the "S" curve modeling provides a reasonable pattern for future growth, the speed of market maturation and the ultimate saturation point (i.e., the 'u' in the function) are not certain. The total individual tax return e-file (i.e., the sum of practitioner, on-line, and TeleFile) forecasts' mean absolute percent error over the last three years for the one-year-out projection is 3.8 percent, compared to 0.3 percent for our total individual tax return (i.e., paper and e-file combined) forecasts. Some of the uncertainty comes from less obvious factors (i.e., interventions) outside the IRS's control. To what extent will the practitioner community bundle their services (including e-file) under one package and one price? How many free on-line filing packages will be made available through the internet; by whom; and accessible to what segments of taxpayers? Looking into the future raises other uncertainties. For example, what other IRS administrative changes and new tax law provisions might be enacted to promote e-filing?

To handle this uncertainty, the Office of Research moved to provide senior IRS management and external stakeholders forecasts presented in ranges. Though most of the resource planning (particularly short-term staffing) decisions are based on the forecasts developed using the methodology described in this paper, IRS Research also produces national level forecasts to reflect "optimistic" and "cautious" scenarios in addition to its official, or "likely," scenario (see Table 2). These provide IRS management a range of possible outcomes in recognition of the uncertainty, and thereby to enable them to tailor contingency plans.

The basic approach for the scenario building for the practitioner and on-line filing projections is to consider different participation rate ceilings (i.e., the u parameters) in their respective two-parameter-bounded logistic growth functions.

On-Line Filing

To the extent possible, we looked to empirical evidence to establish ranges for our upper and lower forecasts for on-line filing. We used the TeleFile experience to figure an on-line filing participation rate in an environment of limited IRS and industry promotion to produce the "cautious" scenario forecasts. The TeleFile participation rate, as calculated by the number of participants divided by the number of eligible TeleFilers *that received a TeleFile package*, was approximately 50 percent in 1999. We therefore set the ceiling participation rate u for the cautious scenario model at 0.5. Given this difference from the "likely" scenario of 66 percent, we set the upper bound participation rate ceiling roughly the same net 16 percentage points above 66 percent, and set it at 80 percent (after rounding).

We then proceeded through the same process as described earlier, conducting SAS grid searches for the optimal a and b parameters and minimizing the RMSE of each equation to find the best fits. The resulting optimal functional forms then were used to calculate the forecasts. Finally, we applied the forecasted participation rates to the forecasted denominator of the ratio – i.e., the population of self prepared returns belonging to taxpayers that own a personal computer and have internet access. The resulting forecasted nominal return counts from the fall 2001 projection cycle are presented in Table 2.

Practitioner e-file

We used the same empirically based data for the practitioner e-file calculations. However, again we had to make adjustments to account for the inflated denominator of the practitioner participation rate ratio that includes all individual income tax returns. Since we provided projection scenarios only for the national level, we deflated the participation rate ratio using the national average for paid tax practitioner usage of 53 percent. We therefore set the participation rate ceiling for the lower bound model at 27 percent (i.e., 0.50 x 0.53) and for the upper bound model at 42 percent (i.e., 0.80 x 0.53). We followed the procedures for functional optimization as for the other cases, and produced the nominal return volume forecasts presented in Table 2.

TeleFile

There is less certainty and less information on the TeleFile trend. Therefore we developed a range around the TeleFile projections using a more simple heuristic approach. Over the last three projection cycles, we over-projected the one-year-out TeleFile volume twice, and under-projected it once. We applied the average percent over-projection to our baseline forecasts to calculate the cautious scenario projections, and similarly applied the percent under-projection to the baseline

forecasts to calculate the optimistic scenario. We present the results in Table 2.

Conclusion

Forecasting electronic filing volumes for individual taxpayers is very difficult because of the novelty of the program and the impact of both big and small interventions over the years. However, for two of the three major components of e-file – i.e., on-line filing and practitioner e-file – the classic structure of the innovation diffusion "S" curve can be applied with reasonable success. Still, to handle the level of uncertainty embedded in those forecasts, IRS researchers provide senior management with a set of "cautious," "likely," and "optimistic" scenarios to provide them a fuller range of the potential future in store.

References

Armstrong, J. Scott, *Principles of Forecasting: A Handbook for Researchers and Practitioners*, Kluwer Academic Publishers, Norwell, MA, 2001.

Booz-Allen & Hamilton, "IRS Cost of Processing Electronic Returns," prepared for the IRS, March 2000.

Council for Electronic Revenue Communication Advancement, "Practitioner Survey Report," December 1999.

Electronic Tax Administration, *A strategy for Growth, IRS Publication 3187*, December 2000.

Framinan, Javier, "Review of the IRS's Individual Return Electronic Filing and Related Research," *The IRS Research Bulletin, IRS Publication 1500*, 1999.

Note

The views expressed in this article represent the opinions and conclusions of the author. They do not necessarily represent the opinion of the Internal Revenue Service.

On Line e-file Volume

Chart 2

Practitioner e-file Volume

Chart 3
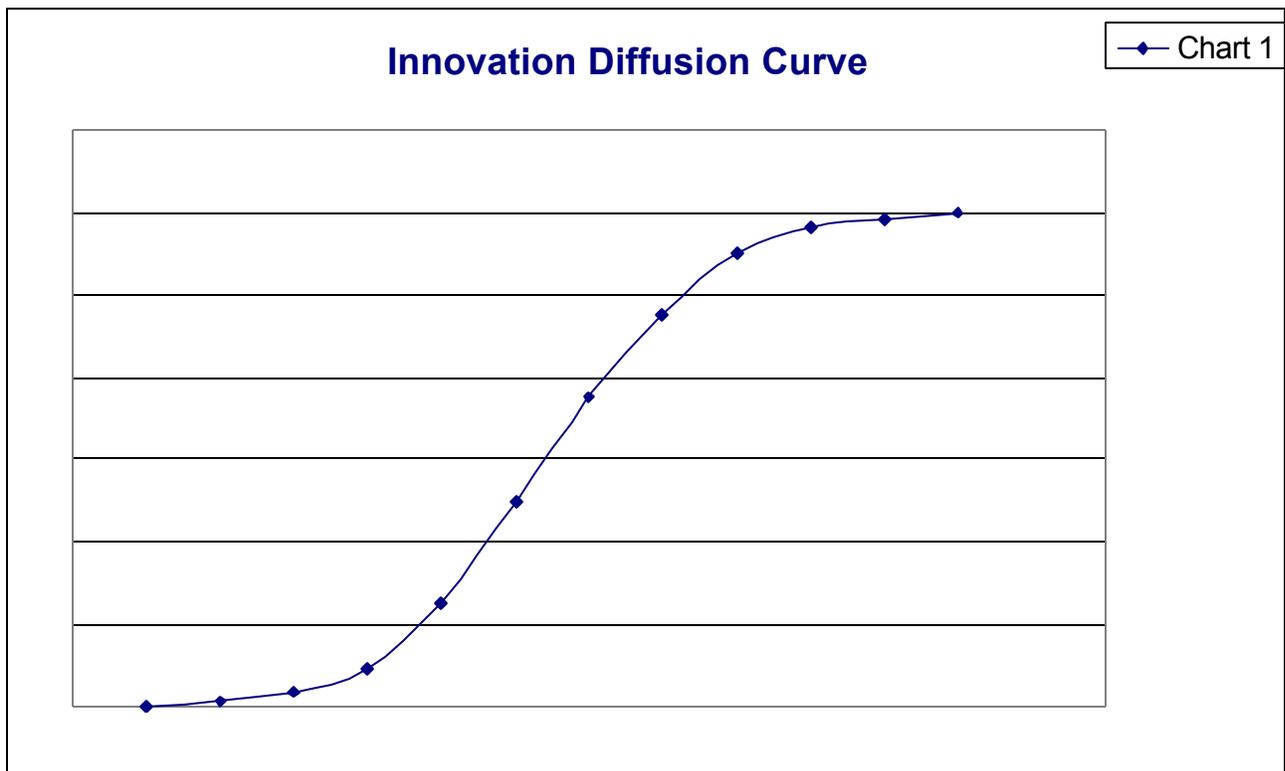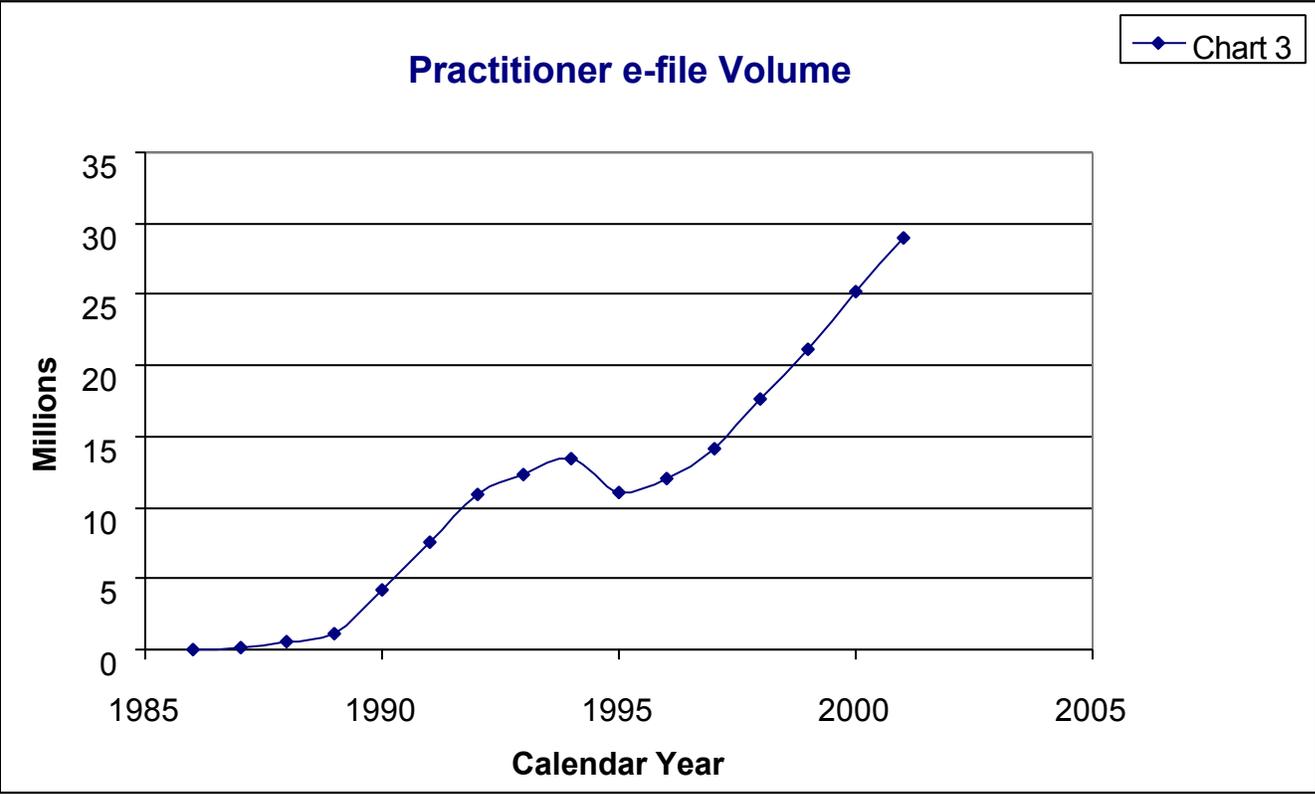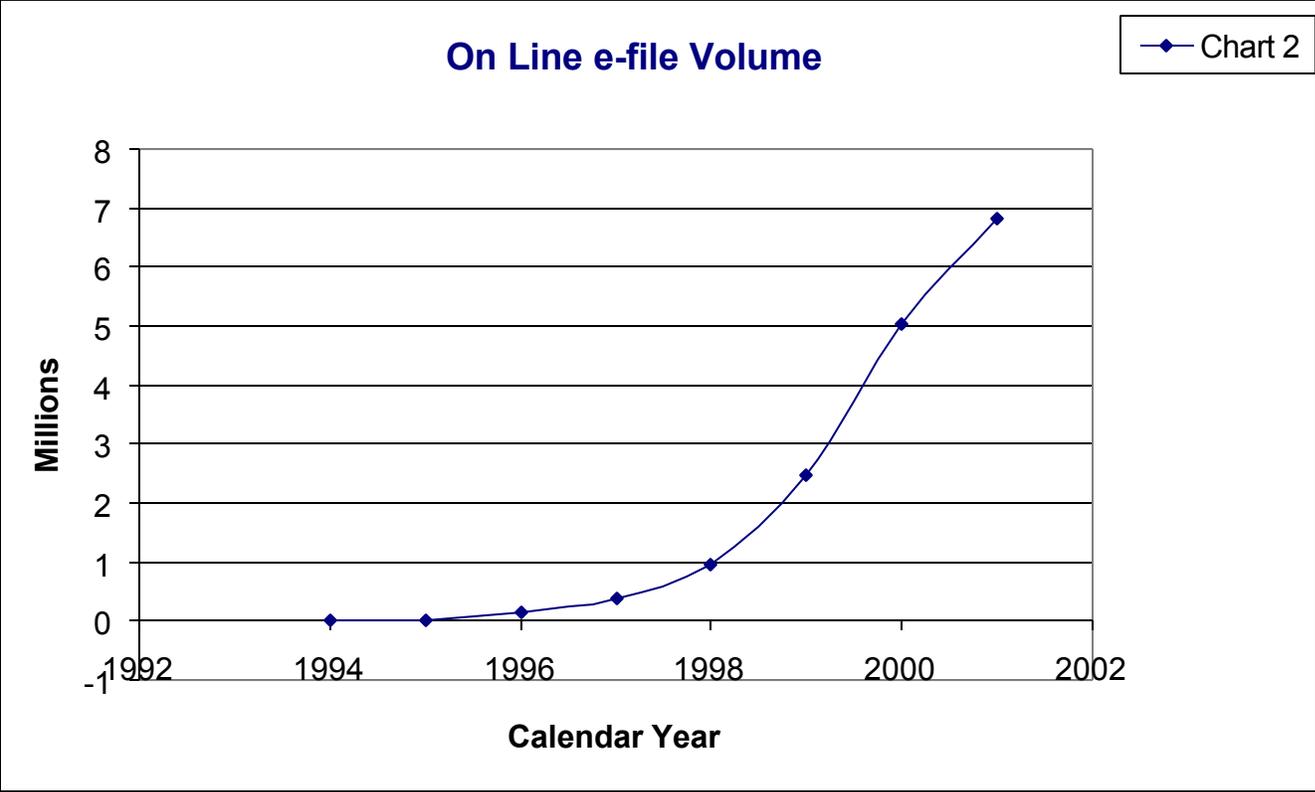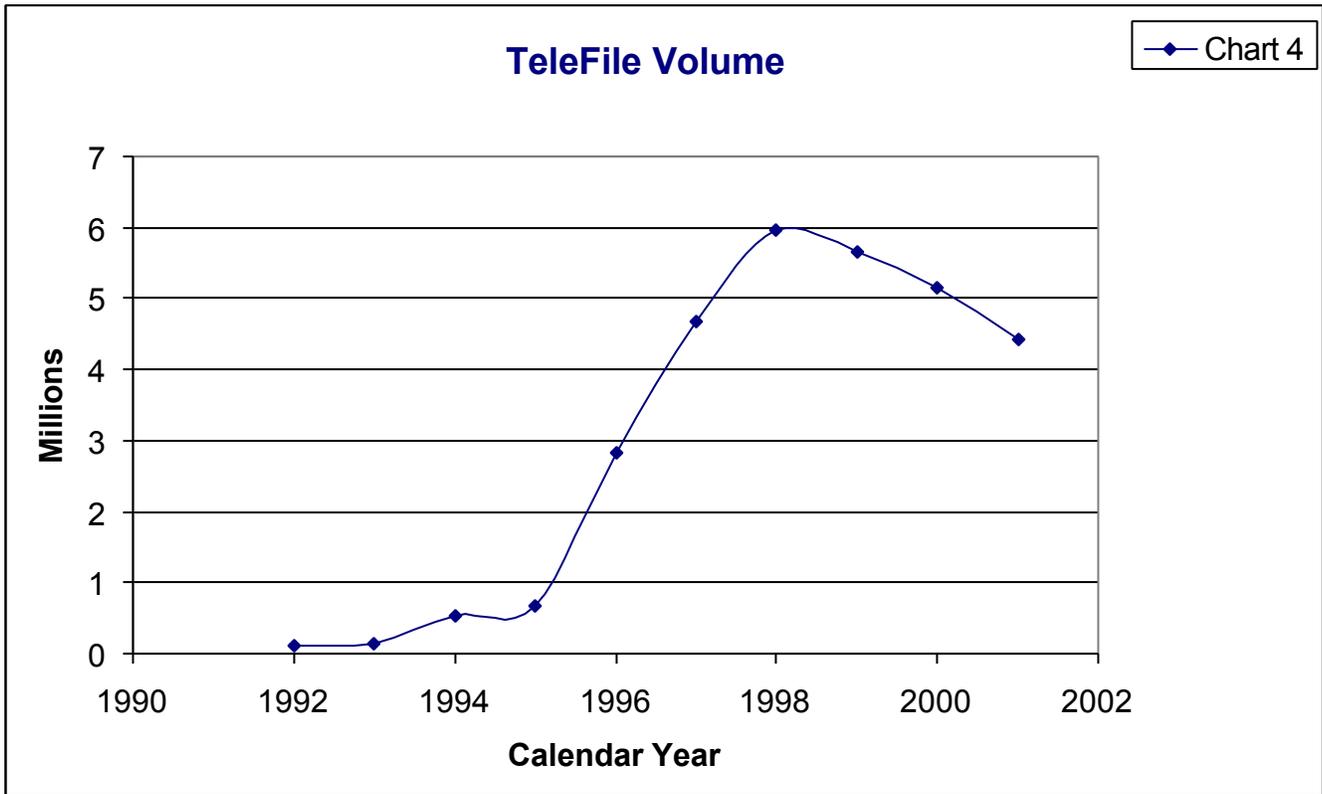
TeleFile Volume

Historical Individual Income Tax Return Volumes
Calendar Year Filings: 1986 - 2001
(in thousands)

Year	Practitioner e-file	On-Line Filing	TeleFile	Total e-file	Paper Returns	Total Individual Series
1986	25			25	101,726	101,751
1987	78			78	103,173	103,251
1988	583			583	106,445	107,029
1989	1,161			1,161	108,708	109,868
1990	4,204 **			4,204	108,101	112,305
1991	7,567			7,567	106,262	113,829
1992	10,919		126	11,045	103,674	114,719
1993	12,334		149	12,482	101,272	113,754
1994	13,502	0	519	14,021	100,663	114,683
1995	11,126	1 **	680	11,807	104,253	116,060
1996	11,971	158	2,839 **	14,968	103,394	118,363
1997	14,083	367	4,686	19,136	101,207	120,342
1998	17,684	942	5,955	24,580	97,967	122,547
1999	21,207	2,458	5,664	29,330	95,534	124,863
2000	25,216	5,026	5,160	35,402	91,695	127,097
2001	28,957	6,831	4,419	40,207	89,238	129,445

** First year of nationwide implementation

Individual Income Tax Returns: e-file Projection Scenarios
Calendar Years 2001 - 2008
(in thousands)

"Likely" Scenario

Year		Total Individual Returns	Practitioner e-file	On-Line Filing	TeleFile	Total e-file	**e-file Rate**
2001	Actual	129,445	28,957	6,831	4,419	40,207	31.1%
2002	Proj'd.	131,991	32,417	8,531	3,983	44,931	34.0%
2003		133,935	35,483	10,857	3,462	49,802	37.2%
2004		136,135	38,216	12,923	2,944	54,083	39.7%
2005		138,352	40,571	14,766	2,685	58,022	41.9%
2006		140,407	42,609	16,403	2,555	61,567	43.8%
2007		142,613	44,270	17,886	2,491	64,647	45.3%
2008		144,431	45,588	19,247	2,458	67,293	46.6%

"Cautious" Scenario

Year		Practitioner e-file	On-Line Filing	TeleFile	Total e-file	**e-file Rate**
2001	Actual	28,957	6,831	4,419	40,207	31.1%
2002	Proj'd.	31,368	8,134	3,431	42,933	32.5%
2003		33,258	9,783	2,666	45,706	34.1%
2004		34,789	11,064	1,864	47,717	35.1%
2005		36,003	12,161	1,700	49,864	36.0%
2006		37,007	13,158	1,618	51,782	36.9%
2007		37,774	14,111	1,577	53,462	37.5%
2008		38,347	15,031	1,556	54,934	38.0%

"Optimistic" Scenario

Year		Practitioner e-file	On-Line Filing	TeleFile	Total e-file	**e-file Rate**
2001	Actual	28,957	6,831	4,419	40,207	31.1%
2002	Proj'd.	33,155	8,700	4,454	46,309	35.1%
2003		37,110	11,479	4,488	53,077	39.6%
2004		40,826	14,236	4,520	59,582	43.8%
2005		44,192	16,877	4,549	65,618	47.4%
2006		47,222	19,290	4,576	71,088	50.6%
2007		49,803	21,463	4,610	75,876	53.2%
2008		51,938	23,407	4,633	79,978	55.4%

This page intentionally left blank.

ACCOUNTS RECEIVABLE RESOLUTION
AND THE IMPACT OF LIEN FILING POLICY ON SOLE PROPRIETOR BUSINESSES

Alex Turk, Ph.D., Internal Revenue Service
Terry Ashley, Ph.D., Internal Revenue Service

Introduction

The federal tax lien (FTL) is an important collection tool in resolving delinquent tax accounts. The lien helps to secure the government's right to the value embodied in the taxpayer's assets and provide creditors with important information on the credit worthiness of taxpayers. However, U.S. legislation, Internal Revenue Service (IRS) policy, and IRS budgetary shocks can alter the degree to which this collection tool is used. In recent years, the number of liens filed by the IRS has declined considerably. The obvious question is what impact has this had on the collection of delinquent accounts. It is important to know how changes in collection enforcement policy will affect the resolution of delinquent accounts.

This paper develops a model of accounts receivable resolution and uses that model to forecast the impacts of changes in the number and the timing of federal tax liens filed against sole proprietor businesses having delinquent tax accounts. We examine sole proprietor businesses having at least one module of delinquent taxes in the accounts receivable inventory. These taxpayers were in the Accounts Receivable File (ARF) for the calendar years 1996 through 1999.

Background

The IRS relies heavily on the FTL as a means of perfecting its security interest in the assets of taxpayers who owe delinquent taxes. Without the protection afforded by the lien, IRS cannot adequately establish its priority over financial institutions and other secured creditors for the equity that taxpayers have in their assets and that may be liquidated in order to satisfy their debts.

Filing fees charged by state and county recording offices for FTL's have escalated over the years to become a significant cost of doing business for IRS. These fees constitute a major expense in the enforcement budget. The upward trend of this expense has coincided with a climate of increasing budgetary pressures for the IRS since fiscal year 1992. As a result, Collection has experienced a series of changes in policy guidelines in the past decade that have set different criteria for determining whether a lien should be filed on a given case. Some criteria are based on the dollar value of the outstanding balance. However, the impact of these changes on taxpayer's propensity to satisfying their tax delinquencies is not well understood.

A significant amount of research has been done or is being done looking at accounts receivable. However, none of this research has focused on evaluating the utility of liens in resolving accounts. IRS research by Butler (1999) analyzed Individual Master File data on balance due returns for Maryland and the District of Columbia. This research calculated the conditional probability of resolving the account using tax return, tax account, and demographic information, including the level of consumer debt. This research ranked groups of accounts based on the conditional likelihood that the accounts could be resolved without intervention by collection. The research concluded that the ability to identify and distinguish among individuals who will pay or not pay balance due amounts would be an important ingredient in reducing the inventory of aggregate entity balances.

Theoretical Framework

The simple theoretical framework for this analysis has its basis in a tax payment behavior model. This taxpayer payment behavior describes the taxpayer as choosing an amount of his tax liability to pay, T_{pd}, which maximizes the following utility expression:

$$(1-\rho)U[(T_c+T_p)-T_{pd}-\theta_1\{max(T_c-T_{pd},0)\}$$

$$-\theta_2(T_p-max\{T_{pd}-T_c,0\})] + \rho U[(T_c+T_p)-T_{pd}$$

$$-\theta_1\{max(T_{tc}-T_{pd},0)+\gamma\}-\theta_2(T_p-max\{T_{pd}-T_{tc},0\})+\gamma] \quad (1)$$

where $U[\bullet]$ is the taxpayer utility function, T_c is estimated current tax liability, T_p is balance due amount, T_{pd} is the tax payment, θ_1 is the interest penalty for nonpayment of the current tax liability, θ_2 is the interest penalty on any nonpayment of any past balance due amount, T_{tc} is the actual current tax liability, γ is the failure-to-pay penalty, and ρ is the probability of audit. The tax payment, T_{pd}, includes any prepayment of taxes through withholding and any payment with the filing of the federal tax return. The term, $max\{T_c-T_{pd},0\}$, reflects

that if the tax payment is less than the tax liability for the current period, the excess tax liability faces the interest rate for the current balance due amount. The term, $max\{T_{pd}-T_c,0\}$, describes that if the tax payment exceeds the tax liability for the current period, the excess is applied to the past due balances.

We assume the taxpayer will choose a T_{pd} that will maximize Eq. (1). The T_{pd} would be the implicit solution to the following first first-order condition of Eq. (1):

$(1-\rho)U'[(T_c+T_p)-T_{pd}-\theta_1\{max(T_c-T_{pd},0)\}$

$-\theta_2(T_p-max\{T_{pd}-T_c,0\})] = -\rho U'[[(T_c+T_p)-T_{pd}$

$-\theta_1\{max(T_{tc}-T_{pd},0)+\gamma\}-\theta_2(T_p-max\{T_{pd}-T_{tc},0\})+\gamma]$ (2)

The left-hand side of Eq. (2) represents the utility gain from not incurring an audit but incurring interest on the balance due. This is weighted by the probability of not being audited. For the right-hand side of Eq. (2), this represents the utility loss from having an audit on the balance due plus sustaining a failure-to-pay penalty and potential tax adjustment as a result of the audit. This loss is also weighted by the probability of being audited. Eq. (2) describes that the marginal expected benefit of the tax payment is equal to the marginal expected cost of the tax payment.

The solution to the first-order conditions in Eq. 2 can be expressed as the simple reduced-form specification,

$T_{pd} = f[X, Prob_{res}(X)],$ (3)

where X is a vector of taxpayer filing characteristics and IRS policy characteristics, and $Prob_{res}$ is the probability of the resolving the balance due account. This model permits an analysis of the IRS compliance strategies and policies to impact the tax payment choice of taxpayers, which include Federal tax liens.

Empirical Model

This research develops an empirical model of accounts receivable resolution for the purpose of evaluating different policies for lien filing. The resolution model is defined as a function of, among other things, characteristics of the liens that are in force on the returns with outstanding balances. The resolution will be measured at the entity (case) level, as opposed to the balance for individual years. Modeling the behavior of the entity more accurately reflects the experience of

Automated Collection System and Collection Field Function personnel as they receive and process their casework. More importantly, a single FTL can cover a number of outstanding balances for different tax years.

Different FTL policy strategies are evaluated by imposing the strategies on independent variables capturing important dimensions of the policy changes. The difference in the model's estimated resolution for the actual data and proposed policy scenarios reveal the impact of the policy change.

Accounts receivable is an issue in both the dollar values and the number of taxpayers involved. Thus, resolution is defined in two ways for the purpose of this research. First, we define resolution as an ordinal variable representing the change in the entity balance for the given time period. This dependent variable takes on three discrete values that represent: (a) an increase in the entity balance; (b); a decrease in the entity balance that is not sufficient to fully resolve all modules; and (c) a decrease in the entity balance that fully resolves all modules. Secondly, we define resolution as the change in a taxpayer's outstanding balance for a given time period. We model the change in dollar value of the entity balance due on the IRS account receivable.

Specifications

We specify a resolution model, r_{it}, and a change in entity balance due model, Δb_{it}, as a system of equations. These two measures reflect one decision by the taxpayer.

Let b_{it} be the entity balance for taxpayer i at time t and let b_{it-1} be the entity balance for taxpayer i at time t-1. We define the ordinal variable r_{it} as

$$r_{it} = \begin{cases} 0 & if\ b_{it} \geq b_{it-1} \\ 1 & if\ b_{it} < b_{it-1}, b_{it} \neq 0 \\ 2 & if\ b_{it} = 0 \end{cases}$$ (4)

We assume that the probability of r_{it} is determined by assignment values for r_{it},

$P(r_{it} = 2) = \Phi(x_{it-1}\alpha),$ (5)

$P(r_{it} = 1) = \Phi(x_{it-1}\alpha + c) - \Phi(x_{It-1}\alpha),$ (6)

$P(r_{it} = 0) = 1 - \Phi(x_{it-1}\alpha + c).$ (7)

where x_{it-1} again is a vector of characteristic for taxpayer i, α is a vector of associated parameters, c is a threshold

value, and Φ is the normal cumulative distribution function. α and c are unknown parameters but can be estimated using the Probit model procedure (Green 1990, pp. 672-676). The measure for lien filing is generated by an instrument variable approach.

We can define the change in entity balance due as

$$\Delta b_{it} = b_{it} - b_{it-1}. \qquad (8)$$

We assume that the change in the entity balance can be modeled by a censored switching regression model,

$$\Delta b_{it} = x_{it-1}\beta_n + \varepsilon_{it\,n} \text{ , if } r_{it} = 0, \qquad (9)$$

$$\Delta b_{it} = x_{it-1}\beta_p + \varepsilon_{it\,p} \text{ , if } r_{it} = 1, \qquad (10)$$

$$\Delta b_{it} = -b_{it-1}, \text{ if } r_{it} = 2, \qquad (11)$$

where x_{it-1} is a vector of characteristics of taxpayer i, β_p and β_n are a vectors of associated parameters, and ε_{itp} and ε_{itn} are a random error terms. The ε_{it}'s in Eq. (9) and Eq. (10) are correlated with the errors in the resolution model. Estimation of these equations will follow a two-stage procedure described by Maddala (1983, pp. 223-234). Under this procedure, the resolution equation serves as the criterion function to generate the selectivity bias variables, ϕ_p/Φ_p, and ϕ_n/Φ_n as suggested by Lee (1982). The selectivity bias variables are included in the switching regressions to achieve consistent estimates of the parameters. Because the errors are also heteroscedastic, the data is weighted by the square of the log of the entity balance.

Data

The data for this research was constructed from the IRS Accounts Receivable Delinquent Inventory (ARDI) data. Annual observations were constructed for a random sample of individual taxpayers with delinquent accounts that had primarily sole proprietor income. This information was combined with information from the Automated Lien System (ALS) data and individual tax return information. Each taxpayer's outstanding balance was observed at the beginning of each year for years 1996 to 1999. The balance in each subsequent year was merged back to the data to determine the annual amount of change in the total balance. If a taxpayer fully resolves the balance, the taxpayer will drop out of the ARDI data. Resolution is defined as decreasing the outstanding balance. The outstanding balance may go up because of any of the following reasons: 1) The taxpayer is making no payments or the

payments don't cover the additional interest and penalties for the year, 2) the tax due on previously filed returns may have been increased as a result of an audit, but not have been fully paid at the conclusion of the audit, or 3) the taxpayer may be filing current returns without paying all the tax reported.

Descriptive statistics and variable definitions for a subset of the variables used to model account resolution are listed in the appendix. The *P(lien)* is the predicted probability of having at least one lien filed on at least one of the taxpayer's returns with a delinquent balance. This probability comes from an instrumental variable regression of a lien dummy on various case characteristics. Geographic location and time dummies are used as instruments.

Figure 1 displays the proportion of sole proprietor delinquent accounts that have a lien covering all or part of the outstanding balance. There appears to be a precipitous drop in the lien coverage over the years. This likely due in part to the impact of the Revenue Reform Act of 1998 (RRA 1998.)

Our data are a series of repeated cross sections. An alternative strategy might be to construct cohort groups and follow resolution to the end of the collection process. Since lien policy changes occur over time, one still needs several different cohorts to be able to identify the policy impact. More importantly, the account will not be written off until the collection statute expires. This can take 10 years or more. Thus, it is difficult to construct the cohorts and have an adequate length of time to observe resolution and observe changes in lien use. The repeated cross section allows us to have a better representation of the cases that are more difficult to collect.
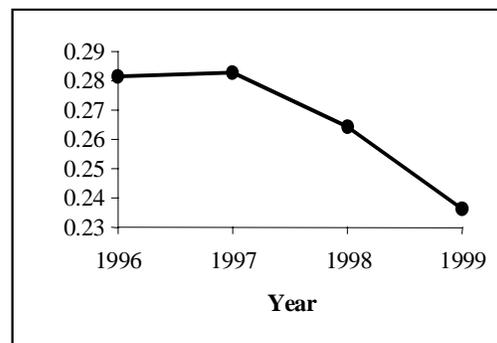

Figure 1 – Proportion of Delinquent Sole Proprietor Businesses with a Federal Tax Lien

Model Estimates

A subset of the estimated parameter for equations (5) – (7) are reported in Table 1. The estimates for equation (10) are reported in Table 2. We are imposing the assumption that there is no impact of the tax lien on conditional expected increase in the entity balance (equation (9)). Thus, we do not report estimates. The model also includes various controls for case status, source of liability, filing characteristics, as well as other characteristics of the outstanding balance.

The positive estimate for *P(lien)* in Table 1 indicates that increasing lien filing will increase the probability that a portion of the outstanding balance will be resolved within the next year. The estimated elasticity for *P(lien)* is slightly larger than 1. This suggests that a one-percent increase in lien filings overall should lead to

Table 1 - Selected Parameter Estimates for Ordered Probit Model of Case Resolution*

Variable	α	$\partial p/\partial x$	Elasticity
ABILITY	1.26E-05	1.257E-05	0.0028
	(1.73E-06)		
EMODBAL	-7.07E-07	-7.061E-07	-0.0292
	(3.85E-08)		
P(Lien)	0.6934	0.6926	1.0858
	(0.0331)		
TCLAGE	-0.11952	-0.1194	-0.0824
	(0.00504)		
TRCATAGE	-0.0936	-0.0879	-0.4983
	(0.0064)		
TRCAGE2	0.0021		
	(0.0006)		
Log L	-167730.0		
n	191,811		

*$\partial p/\partial x$'s and the elasticities are calculated using the average of each independent variable. Standard Errors are listed in parentheses.

slightly more than a one-percent increase in cases that are being resolved. The marginal effect of a lien on the dollars resolved (Table 2) is negative but not significant. Thus, these results imply that increasing lien filing increase the likelihood that the taxpayer will be resolving the outstanding balance but it not clear that it has an influence on how quickly they pay down that balance.

The variable *TCLAGE* measures when the lien, if one was filed, was imposed. The negative value in Table 1 and the Positive value in Table 2 support the notion that

Table 2 - Selected Parameter Estimates for Change in Entity Balance Due Equation – Partial Resolution*

Variable	β	Elasticity
P(Lien)	-0.0187	-0.0136
	(0.0437)	
TCLAGE	0.0289	0.0093
	(0.0083)	
MTH90	0.0041	0.0004
	(0.0220)	
MTH180	-0.0627	-0.0107
	(0.0173)	
DMRE_TAX	-0.1205	-0.0285
	(0.0255)	
ABILITY	0.87E-04	0.0011
	(0.8E-05)	
AVEMODBL	0.0067	0.0067
	(0.0041)	
Selectivity β	0.1014	0.1239
	(0.0373)	
Adj R^2	0.049	
n	47,527	

*The parentheses contain the standard errors.

the earlier liens are filed the better the chances that the outstanding balance will be resolved. This comparative static will also be exploited in our policy forecasts in the next section.

The resolution model also predicts that those sole proprietors with smaller balances in absolute terms and relative to income, have a higher chance of resolving the outstanding balance. The estimates associated with *TRCATAGE* and its square *(TRCAGE2)* suggest a probability of resolution decreases over time but the rate of decay diminishes as the case ages. It is important to note that the marginal effect will eventually become positive, but only well beyond the typical 10-year collection period.

Impact of Policy Changes

We address the impact of policy changes using the model developed above to predict what changes would occur when lien filing policy shifts. We explore two different policy scenarios: increasing the proportion of cases that have a lien in place, and filing the lien earlier for those cases where the IRS currently has liens in place. Both of these policies, based on the model's

comparative static results, will increase the number and the dollar amount of accounts receivable resolved.

The model predicts the change in the entity balance in one year's time. As such, it represents the short-run impact of a policy change. The model does not speak directly to what the long run impact is on the make-up of the ARDI inventory. Increasing lien filings would likely have impact on the size of the accounts receivable. Over time, one would expect that the number and dollar value of cases would be smaller under an increased filing paradigm. Thus, the number of resolutions should decline in proportion. In addition, the analysis here does not incorporate a process for the creation of new receivables. Collection policy and economic factors may have an influence on these forecasts. However, our analysis does not account for this variation.

Increasing Lien Filings

The proportion of sole proprietor businesses with liens in place fell to less than 24% in 1999. The question we attempt to answer is how much of an increase in resolution can we expect from increasing lien filing from this level.

We look at scenarios of increasing the percent of cases with liens by 10, 25, 50, and 100-pecentage points. This is accomplished by increasing the instrumental variable P(lien), the probability of filing a lien, by the corresponding percentage point increase. In the simulation for these scenarios however, *P(lien)* is never greater than 100%. We then compare predicted resolutions using the actual 1999 information with predicted resolutions when we impose the change in lien filing.

Table 3 reports the forecasted number and percentage of cases resolved for each increased level of lien filing. The proportion of cases that are fully or partially resolved in one year time increases from 48% to 64% in the most extreme case of filing liens on all cases. The actual data show a higher percentage of cases being partially resolved than fully resolved. However, as lien filing increases, the impact on full resolution is greater and eventually the number of partial resolutions starts to decline and actually falls below the 1999 level. This happens necessarily because we have model resolution as an ordered categorical variable.

Table 4 reports the increases in the number and percentage of case resolutions (fully or partially resolved cases) for each of the increases in lien filing. The largest impact in resolutions occurs for those sole

Table 3 - Forecasted Resolution Rates when Lien Filing is Increased

Increase in Filing Likelihood	Partial Resolution	Full Resolution	Partial and Full
1999 Level	402,833	380,164	782,997
	24.5%	23.1%	47.6%
10% Increase	409,734	408,571	818,305
	24.9%	24.8%	49.8%
25% Increase	417,148	451,655	868,804
	25.4%	27.5%	52.8%
50% Increase	421,518	523,025	944,543
	25.6%	31.8%	57.4%
Increase to 100%	403,468	644,118	1,047,586
	24.5%	39.2%	63.7%

proprietors with a total outstanding balance from $1,000 to $5,000. This is especially interesting because previous lien policy directives have focused on these lower levels of outstanding balance. In fact, around 80% of the increase in resolutions come from cases with less than $10,000 in outstanding balance. This is due partly to the larger cases are more likely to have a lien imposed, thus increasing lien filing does not affect that group as much. In addition, for the lower dollar cases, resolution is "riskier" in terms of the standard definition of risk as seen typically in insurance and financial markets. That is, risk is the inability to predict. The lower dollar cases have resolution rates around 50%, while cases with larger dollar levels have a 30-40% resolution rate. In the dichotomous case, the slope of

Table 4 - Forecasted Change in Resolutions Resulting from Increased Lien Filing - by the Amount of Outstanding Balance

		Increase in the Lien Filing Percentage			
Outstanding Balance		10%	25%	50%	100%
$0 to $1k	Cases	9,005	22,265	43,382	78,042
	%	4.0%	9.9%	19.2%	34.6%
$1k to $5k	Cases	11,788	29,183	56,327	94,692
	%	4.4%	10.9%	21.1%	35.4%
$5,k to $10k	Cases	5,081	12,486	23,612	37,454
	%	4.7%	11.5%	21.7%	34.5%
$10k to $24k	Cases	4,919	11,853	21,641	32,316
	%	5.1%	12.3%	22.5%	33.6%
> $25k	Cases	4,515	10,019	16,584	22,085
	%	5.3%	11.7%	19.4%	25.9%
Overall	Cases	35,307	85,806	161,546	264,588
	%	4.5%	11.0%	20.6%	33.8%

Table 5 - Forecasted Percentage Change in Entity Balance Due Resulting from Increased Lien Filing - by the Amount of Outstanding Balance

Outstanding Balance		Increase in the Lien Filing Percentage			
		10%	25%	50%	100%
$0 to $1k	%	2.10	5.26	10.46	19.64
$1k to $5k	%	2.02	5.06	9.98	17.64
$5k to $10k	%	1.91	4.75	9.20	15.43
$10k to $24k	%	1.72	4.22	7.92	12.51
> $25k	%	1.29	2.93	5.01	7.04
Overall	%	1.87	4.61	8.91	15.42

the distribution function is much steeper when the resolution rate is closer to 50%.

Table 5 reports the forecasted percentage change in the entity balance due by outstanding balance. The forecasted percentage change is the expected change in the balance due for 1999 subtracted from the expected change in balance due for each of the scenarios. These forecasts show comparable results to those for the case resolutions. The largest percentage gain is made in the $0 to $1,000 range of outstanding balance. This outcome reflects more fully resolving cases occurring in this range relative to the $1,000 to $5,000 range. Likewise, the greater percentage gains for all ranges of outstanding balance are made at resolution rates around 50%.

The forecasted increases in resolved cases due to the increased filing, by the status of the case in the collection process, are reported in Table 6 and Table 7. The largest percentage increases in resolution is for those cases in tolerance status and cases being worked by the call sites. Tolerance cases are not actively being pursued and thus the likelihood that we had a lien in place is very slim. Correspondingly, these cases have very low rates of resolution. The largest impact in terms of cases resolved is for those cases in the installment agreement status. However, there are close to one-half million cases in installment agreement status and the majority are resolving their outstanding balance. Thus, the expected percentage change in resolutions is relatively low.

Decreasing the Time to File Liens

The next policy scenario we explore is simply filing the

Table 6 - Forecasted Increase in Cases being Resolved Due to Increasing Lien filing - by Case Status

Case Status		Increase in the Lien Filing Percentage			
		10%	25%	50%	100%
Bankruptcy	Cases	1,273	3,056	5,655	8,425
	%	5.5%	13.1%	24.3%	36.1%
Freeze	Cases	1,011	2,434	4,514	7,422
	%	3.0%	7.2%	13.4%	22.0%
Offer in Compromise	Cases	438	927	1,438	1,722
	%	3.1%	6.6%	10.2%	12.2%
Not Collectable	Cases	3,471	7,986	13,088	16,303
	%	6.6%	15.2%	25.0%	31.1%
Tolerance	Cases	5,809	14,647	29,512	56,398
	%	5.9%	14.9%	30.0%	57.4%
Notice	Cases	3,357	8,071	14,998	25,149
	%	2.7%	6.5%	12.1%	20.3%
Installment Agreement	Cases	11,811	28,631	53,742	86,862
	%	3.7%	8.9%	16.6%	26.9%
Field	Cases	681	1,441	2,255	2,731
	%	6.1%	13.0%	20.3%	24.6%
Call Site	Cases	6,758	16,976	33,537	55,971
	%	7.2%	18.1%	35.7%	59.5%
Queue	Cases	663	1,550	2,641	3,327
	%	7.6%	17.7%	30.2%	38.0%
Overall	Cases	35,307	85,806	161,546	264,588
	%	4.5%	11.0%	20.6%	33.8%

Table 7 - Forecasted Percentage Change in Entity Balance Due Resulting from Increased Lien filing - by Case Status

Case Status		Increase in the Lien Filing Percentage			
		10%	25%	50%	100%
Criminal Investigation	%	1.02	2.68	5.67	11.22
Bankruptcy	%	1.93	4.72	8.96	14.11
Freeze	%	2.21	5.41	10.31	17.89
Offer in Compromise	%	1.98	4.31	6.89	8.49
Not Collectable	%	1.08	2.45	4.01	5.11
Tolerance	%	1.91	4.86	10.00	19.98
Notice	%	2.17	5.30	10.12	17.91
Installment Agreement	%	2.35	5.79	11.22	19.26
Field	%	1.26	2.74	4.42	5.49
Call Site	%	1.63	4.14	8.37	14.68
Queue	%	0.98	2.29	3.95	5.10
Overall	%	1.87	4.61	8.91	15.42

lien earlier in the life of the account receivable. In other words, what if we keep the number and cases that we file liens on constant but we filed the liens earlier. The model estimates indicate that filing earlier increases the likelihood of resolving the case. However, those comparative static results were fairly small in magnitude. Thus, one should expect more modest results for these policy changes.

We explore filing all liens within two years, within one year, and within 6 months of the creation of the account receivable. We also explore filing the lien in one-half of the observed time. That is, if the lien was actually filed in 18 months, what if we had filed it in 9 months. In this scenario all taxpayers that have a lien would be affected by the policy change as opposed to putting an upper bound on the filing time which affects only those cases where liens were filed relatively later. The aggregate number and percentage of case resolutions of cases with tax liens are reported in Table 8. The predicted resolution rates go up by a smaller amount than when lien filing was expanded. The predicted resolution rates for filing within 6 months are very similar to those for cutting the length of time in half. As can be seen, sole proprietors with liens have a much lower resolution rate (32% fully or partially resolved) than those sole proprietors that have not been subject to a tax lien (52% fully or partially resolved). This reinforces the notion that cases that are more difficult to resolve are more likely to be subject to liens and supports the notion that an instrumental variable approach is appropriate to measure the use of the tax lien.

Table 8 - Predicted Cases Resolved when liens are filed quicker*

Maximum Time to File	Partial Resolution	Full Resolution	Partial and Full
1999 values	72,205 18.6%	53,253 13.7%	125,458 32.3%
2 years	72,906 18.7%	54,364 14.0%	127,270 32.7%
1 Year	74,239 19.1%	56,469 14.5%	130,708 33.6%
6 Months	75,888 19.5%	59,082 15.2%	134,970 34.7%
1/2 the Time	75,288 19.4%	57,918 14.9%	133,206 34.3%

*Includes only cases where liens are filed

Table 9 - Forecasted Increase in Resolved Cases by Balance Level –Filing Liens Quicker*

Outstanding Balance		Maximum Time to File			
		2 years	1 year	6 months	1/2 time
$0 to $1k	Cases	29	81	138	113
	%	1.0%	2.9%	5.0%	4.1%
$1k to $5k	Cases	218	512	823	629
	%	1.9%	4.5%	7.2%	5.5%
$5,k to $10k	Cases	284	766	1,273	978
	%	1.7%	4.7%	7.7%	5.9%
$10k to $24k	Cases	562	1,677	2,963	2,332
	%	1.6%	4.7%	8.3%	6.6%
> $25k	Cases	720	2,214	4,314	3,696
	%	1.2%	3.7%	7.3%	6.2%
Overall	Cases	1,812	5,250	9,512	7,748
	%	1.4%	4.2%	7.6%	6.2%

* Includes only cases where liens are filed

Similar to the case where the lien filing is expanded, the impact is separated out by the level of outstanding balance and by the status of the case. The forecasted increase in the cases resolved for each level of decrease in filing time for the different level of accounts is reported in Table 9. The percentage of additional decrease in the outstanding balance is reported in Table

Table 10 – Forecasted Percentage Change in Entity Balance Due by Outstanding Balance - Filing Liens Quicker*

Outstanding Balance	Maximum Time to Filing			
	2 years	1 year	½ year	½ time
$0 to $1k	0.54%	1.46%	2.50%	3.29%
$1k to $5k	0.74%	1.77%	2.85%	3.62%
$5k to $10k	0.54%	1.49%	2.51%	3.24%
$10k to $24k	0.40%	1.22%	2.22%	2.94%
> $25k	0.29%	0.87%	1.72%	2.37%
Overall	0.39%	1.12%	2.05%	2.74%

* Includes only cases where liens are filed

10. Contrary to the policy of increasing lien filing, decreasing the time has a larger effect, in terms of the number of cases being resolved, in the higher levels of outstanding balance. This is due in part to the fact that the liens are more likely to be filed on cases where the balance is larger. Also, the smaller dollar cases tend to be more recent cases in the process. Thus, if a lien is in place, it is more likely to have already been filed relatively early.

Table 11 - Forecasted Increase in Resolved Cases by Collection Status – Filing Liens Quicker*

Case Status		Maximum Time to File			
		2 years	1 year	6 months	1/2 time
Bankruptcy	Cases	113	313	567	472
	%	1.4%	4.0%	7.3%	6.1%
Freeze	Cases	21	74	128	106
	%	1.0%	3.4%	6.0%	4.9%
Offer in Compromise	Cases	74	216	394	328
	%	0.9%	2.7%	5.0%	4.1%
Not Collectable	Cases	536	1,536	2,754	2,233
	%	1.7%	4.7%	8.5%	6.9%
Tolerance	Cases	52	172	329	276
	%	1.0%	3.4%	6.4%	5.4%
Notice	Cases	28	82	156	132
	%	1.0%	2.9%	5.5%	4.6%
Installment Agreement	Cases	227	768	1,546	1,408
	%	0.6%	2.2%	4.4%	4.0%
Field	Cases	143	402	744	608
	%	1.9%	5.4%	10.0%	8.2%
Call Site	Cases	436	1,220	2,134	1,641
	%	2.3%	6.4%	11.2%	8.6%
Queue	Cases	179	454	739	529
	%	3.5%	8.8%	14.3%	10.3%
Overall	Cases	1,812	5,250	9,512	7,748
	%	1.4%	4.2%	7.6%	6.2%

* Includes only cases where liens are filed

Table 12- Forecasted Percentage Change in Entity Balance Due by Case Status – Filing Liens Quicker*

Case Status	Maximum Time to Filing			
	2 years	1 year	½ year	½ time
Bankruptcy	0.47%	1.29%	2.34%	3.16%
Freeze	0.43%	1.63%	2.90%	3.88%
Offer in Compromise	0.57%	1.70%	3.16%	4.21%
Not Collectable	0.32%	0.93%	1.65%	2.19%
Tolerance	0.30%	1.02%	1.95%	2.60%
Notice	0.48%	1.44%	2.74%	3.69%
Installment Agreement	0.34%	1.16%	2.38%	3.38%
Field	0.36%	1.04%	1.96%	2.65%
Call Site	0.42%	1.18%	2.09%	2.73%
Queue	0.48%	1.24%	2.04%	2.56%
Overall	0.39%	1.12%	2.05%	2.74%

* Includes only cases where liens are filed

CONCLUSIONS AND DIRECTIONS FOR FURTHER RESEARCH

The model developed here supports the notion that lien-filing policy does have an impact on the size of the sole proprietor's outstanding income tax balance. The evidence here suggests that filing more liens can increase the number and dollar amount of outstanding income tax accounts that are resolved. To a lesser extent, filing liens earlier will also help to resolve the accounts more quickly. This is especially true on those cases that make it to the more advanced levels of the collection process. However, resolving the cases is not the only consideration. Some notion of burden and equity have to be considered as well as the expected cost of increasing lien filings. However, these questions are beyond the scope of this paper.

This research could be expanded by estimating resolution in the context of a proportional hazard model as opposed to one period resolution. This would help to evaluate the long run impact on accounts receivable.

The forecasted impact by case status is reported in Table 11 and Table 12. The largest impact in terms of numbers, and percentage increases, comes from cases in more advance levels of collection. For example, the largest impact occurs in the cases deemed not collectible. Cases are not classified as "not collectible" until after other collection avenues are exhausted (Call site, field, etc.). The forecasted increases for case in the call sites and in installment agreement are also relatively high. In terms of percentage increase, there is also a large impact on cases in field collection. The number of cases is not as large because there is a limited number of cases that can be worked in the field at any one point in time. The conclusions that one could draw from this is that for those cases that one expects to be difficult to collect, and thus more likely to make it to the advanced levels of the collection process, have the lien filed as early as possible. This increases the number of sole proprietors that are resolving their outstanding balance.

REFERENCES

Butler, Jeff, "Payment Dynamics of Individual Accounts Receivable and a New Look at Risk," United States Department of the Treasury, Internal Revenue Service, *The IRS Research Bulletin*, Publication 1500 (Rev. 11-99), pages 46 – 59.

Greene, William H., 1993, *Econometric Analysis*. Second Edition. Prentice-Hall.

Heckman, J. "Sample Selection Bias as a Specification Error." *Econometrica* 47, pp. 153-161.

Lee, Lung-Fei, 1983, "Generalized Econometric Models with Selectivity, *Econometrica*, Vol. 51, No. 2, pp. 507-512.

Maddala, G. A., 1983, *Limited-Dependent and Qualitative Variables in Econometrics*, Cambridge University Press.

Note

The views expressed in this article represent the opinions and conclusions of the authors. They do not necessarily represent the opinion of the Internal Revenue Service.

Appendix – Descriptive Statistics and Variable Definitions

Variable	Mean	S.D.	Definition
RBALTM1	-3.84	660.82	Change in the balance in t-1 relative to the Current Balance
DMMARJT	0.37	0.48	Dummy for filing a joint return
DMMARSP	0.04	0.20	Dummy for married and filing separate
DMHDHLD	0.07	0.26	Dummy for filling head of Household
DMWIDWR	0.00	0.02	Dummy for filing as a qualified widower
DMMG_INT	0.20	0.40	Dummy if they had Mortgage Interest deduction
ABILITY	104.4	1769.43	Total positive income divided by current balance
MATCH	0.78	0.41	Dummy for if no income tax return could be found
DMRE_TAX	0.20	0.40	Dummy if there was a deduction for real-estate taxes
RATIO_TX	0.36	0.63	Ratio of tax withholding to total tax liability on the return
DAGE0	0.35	0.48	Dummy if the account is less than one year old
PEXPIRE	0.04	0.17	Proportion of the outstanding balance associated with module that are close to expiration
EMODBAL	19391	134803	Total Outstanding balance
P(lien)	0.26	0.31	Predicted Probability of filing a lien
TCLAGE	117	262.48	Age(in days) of the case when the lien was filed, equals zero if no lien is filed
TRCATAGE	2.66	2.52	Age(in years) of the case
TRCAGE2	13.41	22.84	Age of the case squared
DCI	0.00	0.03	Dummy for Criminal investigation status
DBANK	0.03	0.18	Dummy for Bankruptcy Status
DFREEZE	0.03	0.17	Dummy for Freeze status
DOIC	0.01	0.10	Dummy of offer-in-compromise status
DCNC	0.13	0.34	Dummy for currently not collectible status
DTOL	0.16	0.37	Dummy for tolerance status
DNOTICE	0.10	0.30	Dummy for notice status
DFIELD	0.03	0.17	Dummy for field status
DACS	0.17	0.37	Dummy Automated call site status
DQUEUE	0.03	0.16	Dummy for queue status
DSFR	0.03	0.16	Dummy for major source of assessment - Substitute for return
DTDI	0.20	0.40	Dummy for major source of assessment - Tax Delinquent investigation
DDELRET	0.05	0.21	Dummy for major source of assessment – Delinquent Return
DEXAM	0.10	0.29	Dummy for major source of assessment - Examination assessment
DURP	0.05	0.22	Dummy for major source of assessment - Information returns matching
DADJ	0.02	0.13	Dummy for major source of assessment - Adjustment
DMATH	0.02	0.14	Dummy for major source of assessment - Math error
DBALD	0.76	0.43	Dummy for major source of assessment - Balance due return
D100PEN	0.01	0.12	Dummy for major source of assessment - 100% penalty
DPEN	0.08	0.28	Dummy for major source of assessment - Penalty
AVEMODBL	7.61	1.62	Ratio of the entity balance due to the number of modules
MTH90	0.11	0.32	Dummy for if age of cases is three months or less
MTH180	0.19	0.39	Dummy for if age of cases is six months or less
LNDBLM0	1.75	3.26	Log of the change in entity balance due between period t and t-1
LNBALEXP	0.51	2.10	Log of the sum of the modules close to expiration; modules over 3287 days in ARDI
Selectivity β part	1.22	0.18	Selectivity variable for taxpayer partially resolving entity balance
Selectivity β full	1.66	5.75	Selectivity variable for taxpayer fully resolving entity balance
DM1040	0.65	0.48	Dummy for presence of 1040 Form
DM1040NR	0.0001	0.01	Dummy for presence of 1040NR Form
FORMS1	0.11	0.36	Number of Schedules filed with tax return
NEW_BAL	1.73	3.16	Log of the entity balance if entity balance due was zero in period t-1

Commodity Forecasting Models

Chair: Linwood A. Hoffman, Economic Research Service, U.S. Department of Agriculture

Sources of Discontinuity and Uncertainty in Chinese Agriculture Data

James Hansen and Hsin-Hui Hsu, Economic Research Service, U.S. Department of Agriculture
Frank Fuller, University of Arkansas

China leads the world in agriculture production and consumption for most major commodities. USDA Economic Research Service develops and maintains Chinese agriculture economic models for forecasting, baseline projections, and policy analysis. This research identifies and addresses data problems and policies that cause discontinuity and uncertainty in modeling China's agriculture economic system. Different types of agriculture data problems and causes is identified for China. Problems in estimating parameters and model development is presented. Agricultural policies which altered producers economic incentives and leads to discontinuity is identified. Appropriate solutions for both data and policy problems in modeling China are presented and discussed.

Price Determination for Sorghum Barley and Oats

William Chambers
Economic Research Service, U.S. Department of Agriculture

Models of U.S. farm prices for sorghum, barley, and oats are developed specifying supply and use factors as well as corn price as explanatory variables. To account for premium opportunities available to farmers in the marketplace, the models differentiate between feed and food uses for the three grains. Model performance is tested by comparing forecasted prices with the actual prices received. More robust tests were employed using within market-year estimates of the independent variables. The models have been implemented in USDA's short-term market forecasting work and long-term baseline projections.

Rational Commodity Forecasts: Improving USDA's Cotton Analysis

Stephen MacDonald
Economic Research Service, U.S. Department of Agriculture

The reputation of USDA's crop forecasts is very good. However, USDA's international forecasting effort has been shifted away from its extensive network of overseas representatives, and new tools will have to be developed to ensure that the much smaller pool of Washington, DC analysts can continue to produce forecasts that incorporate all available information. For example, since 1996, USDA's overseas representatives have halved the number of reports devoted to cotton. This paper suggests a methodology for ensuring the rationality of export forecasts, the most vulnerable variable in the supply and demand balance sheet.

This page intentionally left blank.

SOURCES OF DISCONTINUITY AND UNCERTAINTY IN CHINESE AGRICULTURE DATA

James Hansen, Frank Fuller, and Hsin-Hui Hsu, USDA Economic Research Service and University of Arkansas

National governments have been collecting economic data series for centuries and using them for developing and implementing plans, policy analysis, program evaluation and justification. Statistical information, once made public, provides tremendous benefits to the economy because economic agents use the information to make better-informed investment, production, and consumption decisions. However, the benefits achieved from collecting and publishing economic statistics is directly proportional to the accuracy of that data. Inaccurate data may lead to poor political and private decision-making process, leaving some agents potentially worse off. Despite modern data collection practices, no country is immune to errors in the collection and reporting of economic statistical series. It is also true that the economic importance of a country in the global economy is positively related to the costs and benefits of generating accurate economic data.

China is the most populous and the third largest country in terms of physical land mass. It is likely that China will become one of the world leaders in generating economic value in the near future, and with China's entry to the WTO, China's influence on international markets can only be expected to increase. Since the introduction of economic reforms in China in the late 1970's, China's economy has become one the most dynamic and rapidly changing economies on the globe, increasing the difficulties associated with the collection of statistical data and the importance of generating accurate data. Given China's position in the global economy, it is of concern that the quality of economic data published by the Chinese National Bureau of Statistical (CNBS) and other government agencies has come under increasing scrutiny from researchers within China and in the international community. Initially, attention focused on agricultural land figures and livestock production and consumption statistics, but more recently questions have been raised about the accuracy of several macroeconomic indicators. Even Premier Zhu Rongji expressed concern about the veracity of China's offical statistics in March 2000 at the National People's Congress (Parpart). The Chinese government recognizes these problems and has taken significant steps to improve the quality of official statistics. While the quality of data is improving, researchers and government analysts must understand and work with official data, which may still have quality problems.

The purpose of this paper is to provide a brief summary of recognized problems in China's agricultural statistics and to discuss the options open to researchers for addressing these data discrepancies. Though data inaccuracies have a multitude of repercussions for analysis of China's economy and economic policies, the scope of this paper is limited to the implications for modeling and forecasting China's agricultural sector. By alerting researchers to potential difficulties in using China's published data, the authors hope to foster prudent use of Chinese statistics for agricultural research and analysis and to motivate others to devote intellectual energy to developing better methods for addressing China's data problems in the short to medium term. The remainder of the paper is divided into three sections. We begin by providing some relevant stylized facts about China's agricultural sector and by describing the data collection process in China. Next, we point to discrepancies in published statistics and suspected sources of the inaccuracies. Finally, we examine how these statistical discrepancies affect efforts to model and forecast China's agricultural production, consumption, and trade.

China's Agricultural Statistical System

One of the most frequently cited facts concerning China's agricultural situation is that China has more than 20 percent of the world's population but only 7 percent of the world's arable land (FAO). Nevertheless, many countries have far less available arable land, measured by the number of people per hectare of arable land available for agriculture production. Table 1 displays the population per hectare of arable land for several countries. Australia, Argentina, United States, Brazil, and Thailand have a large amount of arable land to support their populations; consequently, these countries are significant exporters of various crops. China has 10.3 people per hectare of arable land, which is more than double the amount of arable land available in Egypt and almost three times more than in South Korea or Japan.

The development of China's agricultural production and marketing system over the last half century has greatly influenced the statistical collection process. In particular, the collectivization of agriculture facilitated the use of a statistical reporting system based on production units. The fact that the government controlled the marketing and distribution of most agricultural products also facilitated centralized collection of data regarding food sales in urban areas. However, as the basic unit of production changed from the work team to the household and food marketing moved increasingly into the hands of private traders, the

accuracy of data collected from the reporting system suffered, forcing a greater reliance on survey data.

Table 1. People per hectare of arable land

Country	People/Ha.	Country	People/Ha.
Australia	0.4	China	10.3
Argentina	1.5	Indonesia	11.6
United States	1.6	Vietnam	13.4
Brazil	3.2	Bangladesh	16.6
Thailand	4.2	Egypt	23.5
European Union	5.1	South Korea	27.3
India	6.1	Japan	28.2

Source: United Nations, FAO, Year 1999

China's Data Collection Procedures

The Chinese National Bureau of Statistics (formerly called the State Statistical Bureau or SSB) is China's official statistical agency and is responsible for the release of all national statistics. The CNBS works closely with other ministries, such as the Ministry of Agriculture (MOA), by comparing and discussing statistics obtained from various institutions. The MOA collects more detailed statistics on agriculture and the rural economy than the CNBS. The major annual CNBS publications used by agricultural economists for conducting research are the *China Statistical Yearbook*, *Rural Statistical Yearbook*, and the *Rural Household Survey Yearbook*. The major annual MOA publication is the *Agriculture Yearbook*. In addition to this publication, there are a number of annual publications on different sectors of the agriculture economy, such as the livestock industry.

Several authors have outlined China's statistical process and how that process has evolved since the creation of Peoples Republic of China (PRC) as economic and political conditions have changed (Tuan and Crook; Barker, Sinha, and Rose; Vogel). The first official government data was published in the early 1950's. In the initial years of the PRC, collection and dissemination of statistical data was not well developed, especially at the national level. A national system for collection of statistics was established August 7, 1952 by the formation of the State Statistical Bureau (SSB). The SSB was intended to be the primary statistical agency in the PRC. The SSB collected data, analyzed statistics, and published official statistics.

Beginning with the first 5-year-plan for the years 1953 – 1957, China's economy was governed by state-planned system that continued until the late 1980's. Under the

central planning system government officials recognized the importance of obtaining agriculture data for planning and implemented statistical programs for agriculture early in the 1950s (Vogel). Agriculture was the largest sector of China's economy in early 1950's. Between 86 and 88 percent of the China's population lived in rural areas in the 1950s, and 31 percent of the total population was employed directly in agriculture (Crook, 1988; Colby, et al., 1992). The primary purpose of the government's statistical programs was to provide information to monitor how well the central government plans were being implemented at the national, provincial, prefecture, county, and commune levels. The plans implemented included 5 year, annual plans, and others based on special needs or sectors (Tuan and Crook). Village production teams were responsible to collect, compile and aggregate statistics before passing them up through the administrative levels from village to township, county, prefecture, province and finally to the national level (Vogel). This system was called the Complete Reporting System. A major change in this system occurred when economic reforms were introduced in the late 1970's. The new system, dubbed the Household Responsibility System (HRS), leased plots of land to individual households and allowed farmers to determine which crops to produce, as long as they could provide a minimum quantity of particular commodities to the government. De-collectivization and implementation of the HRS reduced the accuracy of the Complete Reporting System because the 5.6 million production teams were effectively replaced with over 200 million rural households as statistical reporting units (Vogel).

In addition to the hierarchical data collection through of the Complete Reporting System, the SSB conducted urban and rural household surveys to provide additional input for developing the government's five-year plans. The first household surveys were conducted in 1955 and 1956. The urban and rural household survey system, as well as the Complete Reporting System, was interrupted by the "Great Leap Forward" in the late 1950s and the Cultural Revolution. The last household surveys before the Cultural Revolution were conducted in 1965 (Fang, et al.). Urban and rural household surveys did not fully resume until 1980, although some initial survey work did begin in 1977-1978 (Bramall).

The next major change for the urban and rural household survey occurred when the Statistics law of the People's Republic of China was enacted in December 1983. The law was established to improve the statistical surveys and led to the formation of the Organization of Rural and Social Economic Survey within SSB. A similar organization was created for the urban surveys (Vogel). By end of 1985, both teams were operational.

The emphasis on data collection solely to facilitate government planning began to change in the mid 1980s.

As government policies were enacted that moved the system away from a command economy toward a socialist market economy, the importance of particular statistics also shifted. For example, instead of focusing heavily on agricultural production statistics, greater importance was placed on statistics that provided a better understanding of the changing economy and rural and socio-economic development. Up to 1993, the major contents of the rural household survey were comprehensive indicators such as: household size and location, use of electricity, population, labor force, land, housing, major fixed productive assets, agricultural production and sales, grain balances of rural households, income and expenditures, per capita food consumption, and the number of durable consumer goods owned (Fang, et al.). Since 1993 both urban and rural household surveys have been substantially amended to include numerous variables in the survey questionnaire. From 1993 through 1998 an additional 400 variables were added the rural household survey to capture information about the changing economy that is useful to the government in understanding these changes.

Operating independent of the CNBS, the Ministry of Agriculture (MOA) also collects rural and agriculture data. The MOA also utilizes the Complete Reporting System and household survey to develop agricultural data, but the MOA does not have authority to release rural statistical data. In 1993 the government stated that the MOA is the decision making body for rural economic policies and the SSB is responsible for collection, supervision, and release of rural statistical information (Cao). In addition to the CNBS and MOA, numerous other agencies also collect statistical information on the rural sector, which leads to significant overlap. The major difference between the MOA and CNBS statistical systems is that the MOA does not have its own personnel located at the different government levels, thus wielding less control over these government units. It should be noted that even though CNBS has official responsibility for release of official government statistics, both CNBS and MOA compare there collected statistics, and they communicate with each other prior to the release of official government statistics. Nevertheless, the CNBS has the final decision over the numbers.

Examples of China's Data Problems

The quality of Chinese statistics and the reasons for inconsistency have varied over time. Some agriculture statistics have a long history of inaccuracies, while others have developed more recently. Some statistical inaccuracy is caused by the statistical procedures used, but more important reasons depend on economic incentives and rent seeking. The source of many of China's statistical quality problems lies in the structure of institutional arrangements and the administrative system. Under the Complete Reporting System, there are opportunities for manipulation or exaggeration of the data at each governmental level to achieve personal gain or appease superiors at the next level. This is particularly true when high-level officials use statistical data as an evaluation tool for determining the promotion of lower level officials. The statistical inflation most likely occurs because of administrative pressure and is more likely to occur in the poorest and less developed regions as local level officials seek to meet official goals (Cai).

Macroeconomic Data Issues

The primary macro economic variables used in agriculture economic models are gross domestic product, consumer and producer price indexes, and demographics variables, including total population and the break down of population into rural and urban components. Depending upon the research objective, employment, labor migration, exchange rates, and other variables may be included. Because of the prominent role played by income in determining consumption, GDP growth is often the most important macroeconomic variable in agricultural models. Recently, China's GDP figures have been called into question. In China the sum of the parts can be greater than the whole—at least when it comes to the growth of the economy. For several years economists have observed that almost all provincial GDP growth figures are higher than the national growth (Parpart). For example, every principal administrative region, except one, reported economic growth rates of 8% or greater in 1998, while the national economic growth rate was 7.8% (Cao). In February of 2002, China released a GDP growth rate of 7.3 percent, which was lower than the economic growth rates declared by all provinces except Yunnan (The Economist). CNBS sample surveys for GDP growth in 1995 indicated that the national rate was 3 percent less than the rate derived from provincial GDP growth (The Economist). The difference lies in the fact that the series are collected by two different agencies using different methods. Recent research indicates that both sets of GDP level and growth rates do not appear to be accurate. According to Rawski (2001), national cumulative GDP growth from 1997 to 2001 was no more than one-third the level published by CNBS.

Inaccurate GDP levels and growth rates at both the provincial and national level have large implications for determining the level of unemployment and labor

migration within China, especially because potential improvements in income is a major cause for migration. The veracity of these economic variables have also been questioned in the past decade and a substantial amount of research literature now exists which indicate improvements are needed.

The number of people employed in agriculture, number of migrant agriculture workers, population living in rural and urban areas are quite important and especially the rate of change for these variables over the past decade. Official Chinese data for the number of farm workers may be greatly overstated with a margin of error exceeding 100 million workers (Rawski and Mead). The research also indicates that population of farm laborers has decreased at a faster rate and at a larger scale transferred in to non-agriculture occupations. By overstating the population of farm labor, research on agriculture productivity will be inaccurate and may be understated because labor intensity is actually less. Other research areas, which may be affected by inaccurate data on number of farm workers, are studies on income distribution, labor migration, changing consumption patterns in rural areas, and poverty alleviation in rural areas.

Another area of major concern by researchers is the accuracy of reported trade data. This problem exists not only with developing countries, but also developed and even United States and Canada. In 1996, the US reported a trade deficit of $39.5 billion with China while China data reported the deficit to be $10.5 billion, a difference of $29 billion. The major reason for the large difference is how commodity origins were identified when they transshipped through Hong Kong from China to US and from the US to China through Hong Kong (Feenstra, et al.). Also results using trade data in analyzing competitiveness and changes in trading patterns can be quite difficult, and may lead to misleading results.

Its important to note that the government of China recognizes the statistical methodology can be improved and are working toward this. It is also important for researchers using Chinese official statistical data series to understand China is a country under going a fast rate of change with respect to many aspects of their economy, political institution, sociological conditions, and adoption of new technology. Research by Ravallion and Chen shows that the government methods for obtaining data through surveys has not kept pace with changes and structural transformation occurring in the rural economies. However, the authors note that the quality of the raw data from China's Rural Household Survey is quite good.

Crop Area Data Issues

Crook (1991) and Smil note that reported production levels appear near actual amounts, but appear to be based on underreported area Researchers in China and abroad have generally believed cultivated area in China was underreported. From an economic point of view, there are numerous reasons why Chinese landowners, farmers, and officials might underreport the land area available to them for cultivation (Smil). Farmers have historically underreported cultivated area since ancient times to reduce taxes paid to the local government. Similarly, local officials have misreported tillable land to create a more equitable distribution of taxes paid on land area. Recognizing the variation in land quality, the cultivated area reported by local officials may be based on actual area, but may reflect the productive equivalent area in terms of a standard land quality. For example, if a farmer had 1.3 mu (15 mu = 1 hectare) of poor quality land, which was as productive as 1 mu of good quality land in that region, then the farmer's 1.3 mu would be officially recorder as 1 mu. Therefore, the area was underreported by 23 percent. This practice also increases all the yields to the levels of good or high quality land, overstating actual yields. The common multiples used were between 1.25 to 1.5 mu of poor quality land to 1 mu of good quality land. Under the communal farming system, communes might want to reduce reported area cultivated in order to reduce their state production quota. With a lower quota, less of the commune's grain is sold to the government purchase stations and taxes are lower. At the same time, commune leaders might want to match and surpass state-planning production levels in order to receive recognition; therefore, yields might be inflated.

Based on surveys conducted at the time, inconsistencies in China's cultivated area were first documented in the 1930s. The surveys also indicated that yields were over-reported (Crook 1993). A number of different surveys and studies using satellite imagery were conducted from 1980 to the present. Several of these studies were reviewed by Smil. The estimated farmland from studies conducted by satellite images and survey vary from the lowest estimate at 131.1 million hectares (mha) to the largest at 143.6 mha, which is a difference of 12.5 mha or about 9 percent. All of the land area estimates are much larger than the official statistics for cultivated area. Government statistics place China's cultivated area at 99.3 mha in 1980 and 94.7 mha in 1995.

Additional evidence of China's underreporting of cultivated area was documented by Wang Tong. He estimated that cultivated land was underreported by

about 31 percent. Interestingly, poor mountainous regions and areas where the main rural economic activity is crop cultivation were observed to exhibit the largest underreporting of cultivated area.

Most China researchers believe that official China statistics on land used for cultivated area was under reported by about 30 percent prior to 1997. This belief was supported by the 1997 National Agriculture Census. Prior the release of the 1997 agriculture census data, land used for cultivated area was reported at approximately 95 mha. This number was changed to 130 mha for 1996 following the census. With the revision of the area data, the CNBS stopped updating the cultivated area figures published annually in the *China Statistical Yearbook*. Recent editions of the *China Statistical Yearbook* report cultivated area for the year 1996 by national total and by province.

Sown area for a number of major crops is used by most agricultural economists in modeling China's agriculture. With the release of the 1997 National Agricultural Census data, the Chinese government did not revise the area sown to agriculture commodities. Sown area has gradually increased but no major revisions have occurred in this data series. In 1996 total sown area was 152 million hectares and by the year 2000 sown area had increased 2.5 percent to 156 million hectares. Some researchers are suspicious of the quality of the official statistics for sown area because they were not revised at the national aggregate level or for individual crops when new information from the agriculture census was made available. The sown area is larger than cultivate area because climatic conditions in many regions of China support cropping practices that yield multiple crops in one year. An increase in cultivated area but with identical sown area implicitly decreases the multi-cropping index.

Livestock Data Issues

Compared to the research on China's underreported cultivated, discussion of discrepancies in China's livestock inventories, production, and consumption data is quite recent. As late as 1993, researchers believed that stakeholders in the livestock industry had fewer incentives to misreport livestock data because free markets played a larger in the sector (Crook, 1993). In the mid 1990s the USDA Foreign Agriculture Service (FAS) Attaché Office in Beijing began receiving numerous questions about the reliability of published statistics for Chinese meat production. In 1997, Zhong noted that there was growing evidence of significant disparities between meat and egg consumption data derived from Comprehensive Reporting System and statistics generated from household survey data. Zhong

suggested that the data collected by the CNBS in its annual household surveys underreported meat consumption because it did not take into account food consumed away from home, nor did adequately account for the increased consumption of livestock products by migrant workers in urban areas. Fuller et al. also pointed out that the meat production statistics reported by the CNBS were inconsistent with price movements, livestock trade, and feed use estimates.

Aubert and Fuller et al. both made early attempts to reconcile the difference between meat production estimates generated by the statistical reporting system and production implied from household survey data. Both studies developed their estimates based on the premise that CNBS household survey data was the most reliable estimate for livestock product consumption in China. Therefore, consistent production estimates could be derived from the survey data by correcting for underreporting. Both studies found that there was potentially significant overreporting (20-60 percent or more) of livestock production in the Comprehensive Reporting System.

A 1998 USDA-FASonline (1998) article written by a Chinese scholar echoes the sentiment that there was significant overreporting of meat production in China. The author suggested that inflation in the data was driven by the desire of local and regional officials to improve their standing with their political superiors. The report suggested that data inflation was possible for livestock products after 1985 because meat procurement ceased and statistical checks on the output claims by local officials were minimal. Moreover, the government's increased emphasis on the growth in livestock output after 1985 meant that local officials were evaluated, in part, on their ability to meet proscribed production targets.

Colby, et al. (1999) study of China's meat statistics pointed out that meat production in the early 1980s was slightly underreported. At that time the government relied on livestock data collected from various levels of state government and from the state meat distribution system, so the increasing amount of meat slaughtered by private slaughter houses as a consequence of marketing reforms was not captured by the state reporting system. Colby et al. (1999) used data from the Ministry of Commerce to estimate food consumption away from home. These estimates were employed in constructing a third set of revised statistics that suggested overreporting was generally between 25-35 percent (except for beef and mutton). The authors also use their revised data to in a policy analysis model to examine the impacts of the data revisions on meat production and feed use. This exercise highlighted the interaction

between the level of the production data, feed coefficients, carcass conversion factors, and price response.

It was hoped that the 1997 National Agricultural Census would put to rest questions about China's livestock numbers, but the census generated nearly as many questions as answers. The census showed clearly that China's reported livestock inventories and meat production were exaggerated, but there was some disagreement about the accuracy of the census results. The CNBS did revise down animal inventory data for 1996 by 21-22 percent and red meat production data for 1996 and 1997 by 22-28 percent (FAS, 1999). However, the CNBS did not revise data for earlier years, so the new numbers create a break in the data series. The CNBS also did not make any revisions to poultry and egg statistics, which exhibited significant overreporting in data studies. Equally problematic is the fact that meat production and livestock inventories rebounded back to pre-revision levels by 1998, raising suspicions that data inflation is still a serious problem.

The most recent, and perhaps most thorough, revision of China's livestock statistics was developed by Ma et al. This work employed provincial-level data collected during China's National Agricultural Census and information about away-from-home consumption patterns gathered in surveys in 1998. Like previous authors, Ma generates a revised data series for meat and egg production and consumption from 1980 onward. Their estimates are slightly lower than figures computed by Colby et al. (1999), implying a larger degree of data inflation.

The recent research into China's livestock statistic discrepancies provides the following guidelines to modelers and forecasters who use Chinese livestock data. First, the production data generated by the statistical reporting system includes significant inflation, particularly from 1985 onward, so supply elasticities, growth trends, and productivity growth rates generated by this data are biased. Second, household consumption data collected through CNBS household surveys are generally more reliable, but they ignore food consumed away from home. Consequently, demand estimation based on this data is really an estimation of demand for food consumed at home. Increased urbanization and rising incomes have prompted significant growth in away-from-home consumption, and that trend is likely to continue. This fact should be considered in forecasts. Third, revised data series can have some value for modelers who are seeking to ascertain general trends in supply and productivity growth. Unfortunately, the deterministic nature of the process used to construct the revised data series renders them unsuitable for

estimation. In addition, they are one-time revisions that are not compatible with updates generated by the CNBS or any other statistical agency. Consequently, they may not present a viable data alternative for modelers in the medium to long term.

Data Problems and China Models:

In the post-reform period, China has become the world's largest producer and consumer of many agriculture commodities including: rice, wheat, cotton, tobacco, pork, honey, land-based aquaculture, and some specific types of vegetables and fruit. China is also the world's second largest producer and consumer of corn, poultry, and soybeans. According to statistics published by the Food and Agriculture Organization (FAO) for 1999, China share of global production and consumption is staggering for some commodities. For example, China's production of pork, rice, wheat, and corn account for 45.7, 32.8, 19.3, and 21.2 percent of the world total, respectively. China's enormous domestic production and decades of self-sufficiency oriented policies have limited China's historical trade in agriculture commodities to a fraction of total domestic consumption. Nevertheless, China's large population and agriculture production implies that very small changes in supply or demand can have large impacts on world agricultural trade and international prices.

Given the importance of Chinese agricultural markets and their potential influence on international trade, a number of agriculture economic models of China have been built and consistently maintained by both public and private institutions in different countries. The various models differ considerably as a consequence of their intended use, the individual modelers' knowledge of Chinese markets, and access to data and labor resources. A few of the models are computable general equilibrium (CGE) models based on the framework and database developed through the Global Trade Analysis Project (GTAP) (Hertel and Tsigas). CGE models have the advantage of incorporating the complete macro economy, allowing them to capture inter-industry resource flows. However, the cost of this additional information is a higher degree of commodity aggregation. Most models of China's agricultural sector are partial-equilibrium models that provide a great deal of commodity-specific information but treat other sectors of the economy as exogenous (For a comparison of China agricultural sector models see Hjort or Fan and Agcaoili-Sombilla).

Government officials and commodity organizations often use projections from these models to analyze the impacts of various domestic agricultural and trade policies and conduct research on food security for

specific countries and regions. Policy analysis using partial equilibrium models often shapes discussions in the policy formation process. The data problems discussed above raise a number of issues that modelers must consider when generating forecasts with partial equilibrium models or when evaluating the projections generated by other researchers. Failure to exercise appropriate care with the data can have enormous implications. Lester Brown's doomsday projection of China's agricultural situation is a poignant example of how simplistic projections based on questionable data can skew policy debate and research efforts around the globe.

The quality of macroeconomic data is very important for agricultural sector models because income growth is the primary driving force in consumption over the medium and long term. Inflated GDP figures will tend to dampen income elasticities derived from that data. On the other hand, using inflated GDP data in conjunction with accurate income elasticity estimates will cause forecasts of demand growth to exceed likely realizations. Modelers should evaluate whether their macroeconomic growth projections reflect the GDP inflation of the late 1990s and consider appropriate adjustments to income elasticities or consumption projections.

Problems with China's area data have a bearing on a number of important features in China agricultural models. Perhaps the most important issue is the effect that underreported area has on estimates of yield growth potential. As mentioned above, underreporting area would tend to inflate yields. If analysts factor reported yields into their assessment of yield growth potential, underreporting area would tend to bias yield growth potential downward. On the other hand, if estimates of sown area are much more accurate than cultivated area statistics, as Smil suggests, then yields may be reasonably accurate. Modelers should use other data sources to corroborate their yield estimates.

If sown area is indeed reasonably accurate, then the underreporting of cultivated area will inflate the multi-cropping index. The combination of trends in cultivated area, multi-cropping, and yields summarize the growth of output. Mixing reliable production data with less reliable area or yield data will cause one of the three components to exhibit questionable and perhaps untenable properties. When a model explicitly accounts for multi-cropping and cultivated area, researchers must examine the congruity and feasibility of their assumptions regarding the projected path for cultivated area, multi-cropping, and yields to determine whether projected output levels (or growth rates) are viable. Often some compromise between yield growth, multi-cropping trends, and loss of cultivated area needs to be

achieved to generate forecasts that reflect both the underlying trends in agricultural labor movements and technological change while yielding plausible outcomes.

Finally, problems with the China's livestock statistics create the most difficult obstacles to overcome because the data issues affect both the livestock and the grain sector. Modelers can use revised data series to assess the trend growth in meat consumption and productivity over the last decade, but the revised data is not suitable for data estimation. Consequently, demand elasticities should be derived from household survey data. One way to improve consumption forecasts is to explicitly account for away-from-home consumption. Studies are currently under way that should yield demand parameters that will be useful for this task. Good estimates of livestock product supply elasticities are not currently available, so modelers will have to continue to rely on revised data trends and good judgment in the short run. Production studies that utilize survey data are needed.

Feed demand estimates are also greatly influenced by the inflation in China's livestock statistics. Modelers have often used unrealistically low feed conversion rates for Chinese livestock because little was known about actual conversion rates. Moreover, feed demand estimates using more conventional conversion rates were not consistent observed feed use. Recent research has shed much light on feed conversion rates in China, providing more guidance for modelers (Fuller et al.; Wailes et al.). However, using more accurate feed conversion rates with inflated production or inventory data will generate implausible forecasts for feed use. As with area and yields, modelers must assess their livestock production data, feed conversion parameters, and productivity growth rates as a whole to determine the mix that is most consistent with underlying structural change in China's livestock sector, known feed conversion, and realistic future outcomes.

Summary

China's transition from a command economy to a market-oriented economy has altered the importance of economic statistics collected and published by China's National Bureau of Statistics. Data collection methods that were appropriate for a centrally-planned economy have not evolved rapidly enough to accommodate changing economic and political incentives. Consequently, several inconsistencies have appeared in China's macroeconomic and agricultural data. In this paper we briefly summarized the body of research that has developed over the last two decades to understand the nature of the data discrepancies, particularly discrepancies in GDP, cultivated area, livestock product

output. Users of Chinese statistical data need to be aware of the quality issues and the impacts the discrepancies will have on their analysis and forecasts. The discussion above points to some of the pitfalls that analysts may encounter in using China's published statistics and to ways to address or adapt to particular issues. In the short run, additional research is needed to develop better data estimates and techniques for handling existing data discrepancies. The only viable long-run solution, however, is for the CNBS to eliminate the discrepancies at the source. Fortunately, great strides have already been made to this end. Cooperative efforts between the Chinese government and several organizations in the international community hold great promise for continued progress toward developing methodologies that generate accurate data for China's economy.

References

Aubert, C. 1998. "Food Consumption and Food Production in China: Statistical Uncertainties, Educated Guesses, Reasoned Optimism." Paper presented at the EU-China Academic Network Conference, London, February 2-3, 1998.

Barker, Randolph., Radha Sinha, and Beth Rose. 1982. *The Chinese Agricultural Economy.* Boulder, Colorado:Westview Press.

Bramall, Chris. 2001. "The Quality of China's Household Income Surveys" *China Quarterly* V0, No167, September, 689-705.

Brown, Lester. 1995. *Who Will Feed China?: Wake-up Call for a Small Planet.* New York : W.W. Norton & Co.

Cai, Yongshun. 2000. "Between State and Peasant: Local Cadres and Statistical Reporting in Rural China" *The China Quarterly* v0, n163, (September) 783-805.

Cao, Qingbo., 1999. "China's Agricultural Data and Statistics System and the 1997 Agricultural Census" Agriculture in China and OECD Countries: Past Policies and Future Challenges, Organization for Economic Co-operation and Development, OECD, France, Paris.

Colby, Hunter W., Crook, Frederick W., and Shwu-Eng H. Webb 1992. "Agricultural Statistics of the People's Republic of China, 1949-90" Statistical Bulletin No. 844, Economic Research Service, United States Department of Agriculture, December.

Colby, H., Funing Zhong, and Mark Giordano 1999. "A Review of China's Meat Production Statistics. New Estimates and Implications for Trade" Agriculture in China and OECD Countries: Past Policies and Future Challenges, Organization for Economic Co-operation and Development ,OECD, France Paris.

Crook, Frederick W., 1988. "Agricultural Statistics of the People's Republic of China, 1949-86" Statistical Bulletin No. 764, Economic Research Service, United States Department of Agriculture, April.

Crook, Frederick W., 1993. "Underreporting of China's Cultivated Land Area: Implications for World Agriculture Trade" China Situation and Outlook Series, RS-93-4, Economic Research Service, United States Department of Agriculture, July.

Crook, Frederick W., 1991. "Underreporting China's Grain Area: A Short Story". China: Agriculture and Trade Report. RS-91-3, Economic Research Service, United States Department of Agriculture, July.

Fan, Shenggan., and Mercedita Agcaoili-Sombilla 1997. "Why Do Projections On China's Future Food Supply and Demand Differ? EPTD Discussion Paper No.22, International Food Policy Research Institute, Washington D.C.

Fang, Cheng, Eric Wailes and Gail Cramer., 1998. "China's Rural and Urban Household Survey Data: Collection, Availability, and Problems," Working Paper, 98-WP 202, Center For Agricultural and Rural Development, Iowa State University, Ames, Iowa.

Feenstra, Robert. Wen Hai, Wing Woo, and Shunli Yao. 1999. "The Reliability of Aggregate Statistics - Discrepancies in International Data: An application to China–Hong Kong Entrepot Trade" The American economic review. V89, no.2, 338-343.

Food and Agriculture Organization (FAO), *FAOSTAT Statistics Database.* Available online at http://apps.fao.org.

Fuller, F., D. Hayes, and D. Smith. 2000. "Reconciling Chinese Meat Production and Consumption Data."*Economic Development and Cultural Change* 49(1): 23-43.

Fuller, F., D. Hu, J. Huang, and D. Hayes. 2001. "Livestock Production and Feed Use by Rural Households in China: A Survey Report" Center for Agricultrural and Rural Development, Iowa State University, Staff Report No. 01-SR 96, November.

Hertel, T. W. and M. E. Tsigas. 1997. "Structure of GTAP," chapter 2 in Hertel (ed). *Global Trade Analysis Modeling and Applications.* New York: Cambridge University Press.

Hjort, Kim. 1999. "ChinaAG and other Agriculture Sector Models of china: Review and Assessment" Agriculture in China and OECD Countries: Past Policies and Future Challenges, Organization for Economic Co-operation and Development, OECD, France, Paris.

Ma, Hengyun., Jikun Huang, and Scott Rozelle 2002. " Reassessing China's Livestock Statistics: Analyzing the discrepancies and Creating New Data Series" (fourth coming)

National Bureau of Statistics of China. 2001. *China Statistical Yearbook 2001.* China Statistics Press, September.

Parpart, Uwe. 2002. "China: Hot economy or cooked books?" Asia Times Online, May 4th, www.atimes.china/China/DE04Ad01.html

Ravallion, Martin and Shaohua Chen. 1999. "When Economic Reform Is Faster Than Statistical Reform: Measuring and Explaining Income Inequality in Rural China". Oxford Bulletin of Economics and Statistics v61, n1 (February): 33-56

Rawski, T. G. 2001. "What is happening to China's GDP Statistics?" *China Economic Review.* 12: 347-354.

Rawski, T. G., and Robert W. Mead. 1998. "On the Trail of the China's Phantom Framers." *World Development.* 26: 767-781.

Smil, V. 1999. "China's Agricultural Land." *China Quarterly* v0, n158 (June): 414-29.

The Economist, 2002. "China: How cooked are the books?" The Economist, March 16th-22nd, page 35-36.

Tuan, Francis C., , Frederick W. Crook. 1983. "Planning and statistical systems in China's agriculture".Washington, D.C., USDA., Economic Research Service, Foreign agricultural economic report, no. 181, Supt. of Docs., U.S. G.P.O.

U.S. Department of Agriculture, Economic Research Service. China: Agriculture and Trade Report. RS-91-3, July 1991.

U.S. Department of Agriculture, 1998. "FASonline: China Livestock, Meat Production and Demand for Feed Grains" USDA Foreign Agriculture Service, CH8016A available at   http://www.fas.usda.gov /scriptsw/attacherep

U.S. Department of Agriculture, 1999. "FASonline: China Livestock Semi-Annual Report 1999" USDA Foreign Agriculture Service, CH9006 . Available at http://www.fas.usda.gov/scriptsw/attacherep

Vogel, Frederic A. 1999. "Review of Chinese Crop Production Forecasting and Estimation Methodology" Miscellaneous Publication No. 1556, National Agricultural Statistics Service, United States Department of Agriculture, June.

Wailes, Eric, Cheng Feng, Gail Cramer, and Xiaohui Zhang. 1998. "China's Livestock Feed Use Relationships: Preliminary Results from a Survey in Seven Provinces" In *Food Markets in China: New Looks and Deeper Understandings.* Proceedings of the WCC-101 Annual Meetings in Honolulu, Hawaii, January 12–13, 1998

Wang Tong, Team Leader et. al. 1992. Zhongguo kengdi dijian wenti de shuliang jingji fenxi (An Economic Analysis of the Progressive Decrease in the Quantity of China's Cultivated Land). Beijing, Jingji Kexue Chubanshe, May.

Zhong, Funing. 1997. "Will China Increase Feed Grain Imports?—An Assessment of China's Meat Production and Consumption Statistics." Paper presented to the International Food Policy Research Institute, January 29, 1997.

This page intentionally left blank.

Price Determination for Sorghum Barley and Oats
William Chambers, U.S. Department of Agriculture

Feed grains other than corn are a small but important component of production agriculture. Individually, these crops account for a small fraction of total cash grain receipts, but together they are more significant. The combined crop value of sorghum, barley, and oats averaged $1.7 billion from 1998 to 2001. In addition, sorghum, barley, and oats are a prominent part of commodity policy with significant budget exposure to the federal government.

Each month, the U.S. Department of Agriculture (USDA) analyzes major commodity markets (including sorghum, barley, and oats) and publishes annual supply, demand, and price projections for the current year. Additionally, once a year, USDA publishes 10-year baseline projections for the agricultural sector that include commodity supply, demand, and prices. This report examines some of the factors that influence U.S. farm-level prices for sorghum, barley, and oats. The models in this paper forecast season average prices received by farmers. This paper mirrors the work done by Westcott and Hoffman, which analyzed price determination for corn and wheat.[1] The models developed here provide an analytical framework for forecasting prices and allow consistency checks to be made for supply, demand, and price forecasts. The following sections provide background information for the minor feed grains sector, discuss the important supply and demand factors, develop statistical forecasting models, and provide an evaluation of each model's effectiveness.

Background

Corn is the dominant feed grain, while sorghum, barley, and oats make up a relatively small proportion of the total feed grain complex. The main use of feed grains is as an energy source for livestock, although many feed grains have significant food and industrial uses. Some uses are specific to a feed grain, such as malting barley, but in general grains are highly substitutable for one another when used in animal feed. Because of this, prices for the different feed grains are closely related to one another and are heavily affected by corn supply and use.

Sorghum is the most prominent of the three minor feed grains:

- Farmers' cash receipts for sorghum averaged nearly $921 million per year in 1998-2001.

- Sorghum tolerates hot and dry conditions better than corn, and although it is grown throughout much of the United states, production is concentrated in the Southern Plains.

- Kansas and Texas are the largest producing states and together account for more than 70 percent of total U.S. sorghum production.

Barley is the next largest minor feed grain:

- The average annual value of barley production in 1998-2001 was $617 million.

- Barley production is concentrated in the northern and western part of the country.

- In 2000/01, North Dakota was the largest barley producing state accounting for 30 percent of total production.

Oats is the least prominent minor feed grain:

- The average annual value of oats production in 1998-2001 was $180 million.

- Oats used to be one of the most important grains produced in the U.S. but has steadily declined in significance since the 1950's.

- Oats do well in cool climates and are grown in the upper third of the United States.

- The U.S. imports about 30 percent of total oats supply, primarily from Canada.

Price Determination Factors for Sorghum, Barley, and Oats

Prices for agricultural commodities are determined by the interaction of supply and demand. The relationship between supply and demand is reflected in carryover stocks, which are inversely related to price. If total use rises relative to supply, carryover stocks will decline and

---

[1] The article Price Determination for Corn and Wheat: The Role of Market Factors and Government Programs can be accessed on the web at http://www.ers.usda.gov/publications/tb1878/.

farm prices will tend to rise. By contrast, rising carryover stocks will tend to decrease farm prices.

Since feed grains are highly substitutable with one another, prices are influenced by the interaction of supply and demand for the entire feed grain complex. Corn, with a value of production of more than $18 billion, is the most prominent feed grain in the U.S. and plays a crucial role for all of the minor feed grains. Prices for sorghum, oats, and barley are strongly correlated with the corn price. This is especially true for sorghum, which is primarily used as feed.

Another important aspect of price determination for minor feed grains is that price premiums are associated with the higher quality grades used for food or industrial purposes. This applies in varying degrees to the different feed grains and is most notable for barley where a significant portion of the grain is used for malting. The premiums that these grains obtain for food and industrial uses differentiates them from the other feed grains.

Supply factors

The components of supply are beginning stocks, imports, and production. To a varying degree, the production of sorghum, barley, and oats has declined for the past decade. This drop has been largely caused by competition from other crops that provide farmers with higher returns, especially corn and soybeans.

Beginning stocks. Ending stocks from the previous year become the current year's beginning stocks and augment current production in determining total supply.

Imports. Imports are important for barley and oats, but not a factor for sorghum. The U.S. imported 28 million bushels of barley in 1999/00, which accounted for 6 percent of total supply. Most of these barley imports came from Canada and were for malting. U.S. barley imports were relatively small until fusarium head blight (FHB) became a problem in the Northern Plains in the early 1990's. This disease can result in lower yields and quality problems, including vomitoxin, a toxic byproduct of FHB. Vomitoxin forced brewers to identify alternative sources of malting barley and many brewers turned to Canada and the Western United States.

The U.S. imports nearly 100 million bushels of oats annually. Canada is the primary foreign oats supplier to the U.S., followed by Finland and Sweden. The United States became a net oats importer in the early 1980's after a long decline in production that began in the 1950's. Canadian production has expanded throughout the 1990's to meet U.S. demand. U.S. imports from Finland and Sweden tend to be larger when the North American oats crop is small.

Production. Production is the primary component of supply and is determined by harvested acreage and yield per acre. Planted acreage is a reflection of producers' expected returns for a given commodity compared with expected returns for competing crops (this incorporates not only price but also expected yields and government benefits). Agronomic considerations, such as crop rotations, can also influence farm plantings. Acreage planted to sorghum, barley, and oats has declined throughout the 1990's while corn plantings have increased. This has expanded the relative importance of corn in the feed grain complex. A major factor explaining the reduction in minor feed grain acreage is improved varieties of competing crops—especially corn and soybeans—which are now economically grown outside of the traditional corn belt.

Changes in farm policy, especially the move towards planting flexibility, are another important factor explaining the shift in acreage away from minor feed grains. Government programs have been increasing planting flexibility over the past fifteen years, and the 1996 Farm Act enabled farmers to grow almost any crop on their contract acreage without losing program benefits.[2] This flexibility, along with improved corn and soybean varieties, enabled some farmers to shift production away from sorghum, barley, and oats.

Yields are affected by many factors including climate, weather, farm management practices, crop variety, and soil type. Trend yields are a good composite indicator of productivity gains associated with from production practices, management skills, technology, and input use. However, in any given year weather events are crucial and can push yields above or below trend. Major deviations from trend yields can have a significant impact on prices.

National average yields for sorghum, barley, and oats have increased over time, but the rate of increase is less than it is for corn. Research into new hybrids is a major factor leading to higher yields, and there is generally more research conducted on corn than the other feed grains. The emergence of agricultural biotechnology could lead to further increases in crop yields. Biotech crop varieties are designed to be insect resistant and/or

---

[2] For more information about farm policies, see the ERS farm and commodity policy briefing room at http://www.ers.usda.gov/briefing/FarmPolicy, or *Feed Grains: Background and Issues for Farm Legislation* at http://www.ers.usda.gov/publications/fds-0701-01.

herbicide tolerant, or could have enhanced specific end-use characteristics desired by consumers. However, biotech varieties of sorghum, barley, and oats are not yet commercially available.

U.S. production of all feed grains averaged nearly 270 million metric tons in 1998-2000. In this time period, corn accounted for nearly 92 percent of total feed grain production followed by sorghum at 5 percent, barley at nearly 3 percent, and oats at less than 1 percent. In 1998-2000, corn production averaged more than 9.7 billion bushels, sorghum was 528 million bushels, barley was 317 million bushels, and oats was 154 million bushels.

Demand Factors

Major components of demand for sorghum, barley, and oats include food, seed, and industrial (FSI) uses, feed and residual use, exports, and ending stocks. Domestic use for minor feed grains has been declining in tandem with declines in production.

Food Seed and Industrial (FSI). FSI use varies significantly between sorghum, barley, and oats. Sorghum can be milled into the same types of products as corn, including food grade sorghum, sweeteners, and ethanol, but FSI represents a small portion of total sorghum supply and use. By contrast, FSI is a major component of domestic barley utilization, mainly for malt production. Malting barley is of a higher quality than feed barley and fetches a significant price premium in the marketplace. Oats are also commonly produced for food, but the bulk of oats are used as animal feed. Oats FSI increased significantly in the second half of the 1980's due to growing consumer awareness of health benefits provided by oat bran.

Seed use is a relatively small component of total demand, and reflects the amount of land planted to the crop and per-acre seeding rates. Seeding rates vary across states due to different soil types and production practices, but change slowly over time. Therefore, national average seeding rates tend to be fairly stable, and seed use tends to rise or fall with acres planted.

Feed and Residual. Feed and residual is the primary demand component for sorghum and is also important for oats and barley. Data for this category are obtained by subtracting FSI uses, exports, and ending stocks from total supply. As a result, some variation reflects unaccounted statistical measurement errors in other categories of supply and demand. Because the data for the other categories are collected by the commerce department or obtained from surveys, feed and residual

is truly a "residual" component of the supply and demand balance sheet.

Feed use is closely related to the number of animals on feed as well as the price of competing grains, including feed wheat, non-grain feeds, and numerous byproducts. Feed and residual use of sorghum averaged nearly 255 million bushels annually in 1998-2000 and accounted for 48 percent of total sorghum use. Feed and residual barley use averaged 140 million bushels in 1998-2000, accounting for 40 percent of total barley use. Feed and residual use of oats averaged 188 million bushels in 1998-2000 and accounted for more than 70 percent of total oats use.

Exports. Exports are very important for the sorghum industry, modestly important for barley, and virtually non-existent for oats. More than 40 percent of the domestic sorghum supply is exported, the bulk of which is used as animal feed. Mexico is by far the most important destination for U.S. sorghum accounting for more than half of the total. Japan is the second largest importer of U.S.-grown sorghum. Barley exports vary significantly from year-to-year but averaged nearly 40 million bushels in 1998-2000. The Middle East, especially Saudi Arabia, is the most important destination for U.S. barley exports. However, the EU is the primary feed barley supplier to the Middle East. Exports of malting barley are relatively small, in part because of different varietal preferences in offshore markets and strong competition from the EU, Canada, and Australia.

Ending Stocks. The difference between total supply and total use results in ending stocks, which become a supply component for the next crop year. Ending stocks are inversely related to price and as such are a key market situation indicator. If total use rises relative to supply, ending stocks decline and commodity prices will tend to rise. On the other hand, if supply rises relative to total use, prices tend to decline as ending stocks build. A summary measure of the 'tightness' of supplies is the stocks-to-use ratio, or ratio of ending stocks to total use.

Pricing models for minor feed grains

As is evident from the above discussion, the fundamental markets for the three minor feed grains are different, and therefore the statistical price determination models will be different. However, there are key similarities between these commodities, and the specifications will have some common variables. In general, the most important independent variables used in this analysis are the corn price, the stocks-to-use ratio, and a measure of the proportion of the grain used for food and industrial uses.

Since feed inputs are highly substitutable and corn is the dominant feed grain, the corn price helps to determine prices for all feed grains. This relationship is especially strong for sorghum, which has a relatively minor food and industrial component and thus competes more directly with corn in the feed market. It is less important for oats and barley because their food and industrial demand is independent of the corn supply and generates a premium, which increases the average farm price. Note that no policy variables are included in the regressions. However, since the price of corn is influenced by government policy it is a good proxy for policy changes that occurred within the estimation period.

The stocks-to-use ratio is a key factor of supply and demand and is commonly used in price forecasting models. The stocks-to-use ratio is defined as the stocks of a commodity at the end of a marketing year divided by total use during the period. The stocks-to-use ratio encapsulates both supply and demand pressures. All other things the same, when the stocks-to-use ratio increases (decreases) prices will tend to fall (rise).

In this analysis, the inverse of the stocks-to-use ratio is used. This provides similar statistical information but the expected sign on the coefficient is positive so when the inverse of the stocks-to-use ratio increases prices are expected to rise. In addition, the cumulative stocks situation for all feed grains affects the price of individual feed grains. Because of its dominant position in the market, ending stocks of corn have a major impact on the price of corn and the prices for all of the minor feed grains. Including corn price in the analysis captures the adverse effects of increasing feed grain stocks on the price of individual feed grains.

Food and industrial use is important because these products are generally of higher value than feed products (there are also high quality feed products that earn premiums). We included a variable that reflects food and industrial use for each of the different grains. Food and industrial use is most important for barley and oats. In particular, food grade oats and malting barley command premiums in the marketplace that are critical determinants of the season average farm price.

Model Implementation

Ordinary least squares was used to estimate the statistical models. The data for all of these pricing models can be obtained from the Feed Situation and Outlook Yearbooks or from the ERS corn briefing room (http://www.ers.usda.gov/Briefing/Corn/). In the following sections, we describe the different models for sorghum, barley, and oats.

Table 1--Summary of variable definitions

Variable name	Definition
SP	Sorghum price
CP	Corn price
ABP	All barley price (includes malting and feed barley)
OP	Oats price
SUR	Stocks-to-use ratio
FAIU	Food alcohol and industrial use
FEU	Feed use
EX	Exports
USE	Total use of the commodity
S	Total supply of the commodity

Sorghum Price Equation

The estimated regression equation for sorghum prices is shown below. Variable definitions are provided in table 1.

(1) $SP = -0.0754 + 0.9261(CP)^* + 0.5945(1/SUR)^{**}$

$R^2 = 0.956$
F-Value = 238.226
Standard error of regression = 0.0892
Durbin-Watson statistic = 1.8719
Estimation period: 1975-1999
* Significant at the 95 percent level or better.
** Significant at 90 percent level or better.

Nearly 96 percent of the variation in sorghum price is explained by the independent variables in the equation.

Figure 1
**Sorghum Prices: Actual and Model Estimates**
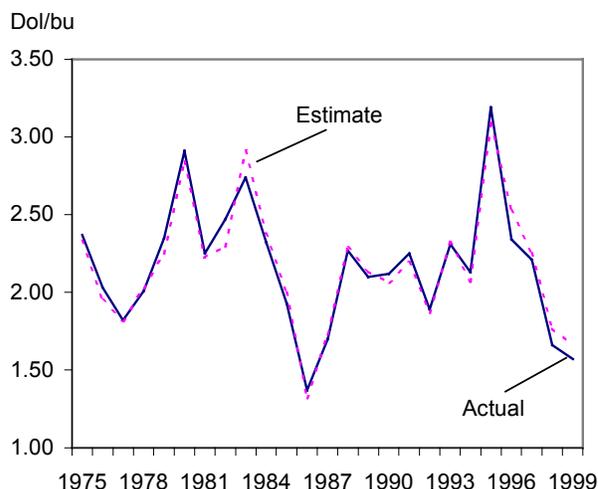
Table 2--Standardized coefficients and elasticities for the different commodities

Independent Variable	Standardized Coefficient	Elasticity
---------------Sorghum---------------		
CP	0.94	1.02
1/SUR	0.09	0.02
---------------Oats---------------		
CP	0.53	0.71
FAIU/FAIU(-1)	0.65	1.50
1/SUR	0.33	0.43
---------------Barley---------------		
CP	0.77	0.67
(FEU+EX)/USE	-0.18	-0.34
Ln(S/S(-1))	-0.29	0.00

Each coefficient has the expected sign. The coefficient for corn price greatly exceeds the 1-percent significance level. The coefficient for 1/SUR exceeds the 10-percent significance level. The F-value for the equation is very large exceeding the 1-percent significance level. The Durbin-Watson statistic indicates that first order auto correlation is not a problem.

Table 2 shows the standardized coefficients and elasticities for the different price equations, which describe the relative importance of the independent variables in each regression model.[3] Because the uses for corn and sorghum are so similar, corn price is the most important independent variable in the sorghum price equation. The inverse of the stocks-to-use ratio is also important but does not explain nearly as much variation as corn price. A likely explanation is that the sorghum stocks-to-use ratio does not account for the stock situation of other feed grains, especially corn.

---

[3] Standardized coefficients describe the relative importance of the independent variables on a multiple regression model. The standardized coefficient adjusts the estimated slope parameter by the ratio of the standard deviation of the independent variable to the standard deviation of the dependent variable. A standardized coefficient of 0.7 means that a change of 1 standard deviation in the independent variable will lead to a change of 0.7 standard deviation in the dependent variable. A standardized coefficient is very similar to an elasticity, which adjusts the slope parameter by the ratio of the means of the independent and dependent variables. An elasticity measures the effect on the dependent variable of a 1 percent change in an independent variable.

Because the two grains are so closely related and corn is such a dominant factor in the feed grain complex, the corn stocks-to-use ratio is highly relevant for sorghum price. Although we did not include corn stocks in the regression, much of this information is captured with the corn price variable. Finally, note that we did not include any variables for food and industrial uses of sorghum. This is because food and industrial use is a minor category for sorghum and the model presented explains much of the variation.

Oats Price Equation

The estimated regression equation for oats prices is shown below. As with the sorghum equation, the variables are defined in table 1.

(2) $OP = -2.416 + 0.441(CP)^* + 2.179(FAIU/FAIU_{-1})^*$
$+ 20.626(1/SUR)^*$

$R^2 = 0.763$
F-Value = 21.504
Standard error of regression = 0.1825
Durbin-Watson statistic = 2.038
Estimation period: 1975-1999
* Significant at the 95 percent level or better.

More than 76 percent of the variation in oats price is explained by the model. Each of the coefficients has the expected sign, and each coefficient is significant at better than the 1-percent level. The F-value also exceeds the 1-percent significance level. The Durbin-Watson statistic indicates that first order auto-correlation is not a problem with the equation.

Figure 2
Oats Prices: Actual and Model Estimates

Dol/bu
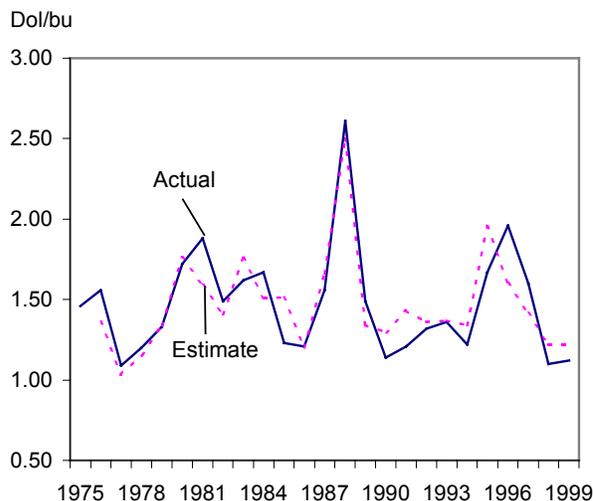
As with the sorghum equation, the corn price is a very important variable in determining the price of oats. However, the relationship between oats and corn prices (with a correlation of 0.54) is not nearly as pronounced as that between sorghum and corn prices (correlation of 0.973). This is because oats and corn are not as closely substitutable as sorghum and corn. The most obvious difference is that oats have a food component that differs from corn. Feed use is also different for oats. Horse enthusiasts commonly feed oats, a large proportion is fed on farm, and very little is used for poultry. Because of these differences, the corn price does not affect the oats price as much as was the case with sorghum.

The oats price model introduces an additional variable that shows the ratio of this period's and last period's food alcohol and industrial use (FAIU). This is a measure of growth in non-feed uses. As FAIU increases relative to the prior year, prices tend to increase because they are higher-valued end uses. If FAIU decreases from a year earlier the impact on price will be smaller because there are fewer food alcohol and industrial products being consumed (holding other variables constant).

Barley Price Equation

The estimated regression equation for the all-barley price is shown below. Definitions of the variables are in table 1.

(3) $ABP = 1.4931 + 0.6321(CP)^*$
$-0.0125(FEU+EX)/USE^{**}$
$- 0.9455Ln(S/S_{-1})^*$

$R^2 = 0.819$
F-Value = 30.178
Standard error of regression = 0.157
Durbin-Watson statistic = 1.919
Estimation period: 1975-1999
* Significant at the 95 percent level or better.
** Significant at the 90 percent level or better.

Nearly 82 percent of the variation in barley price is explained by the independent variables of equation (3). The statistical significance on the coefficient for corn price greatly exceeds the 1-percent level. The significance of the coefficient for the log of the ratio for last year's and this year's supply ($Ln(S/S_{-1})$) is at nearly the 1-percent level. The coefficient that shows the ratio of total (domestic and export) feed use to total use ($(FEU+EX)/USE$) is significant at the 10 percent level. As with the other two models, the F-value is very large and greatly exceeds the 1-percent significance level.

Figure 3
Barley Prices: Actual and Model Estimates
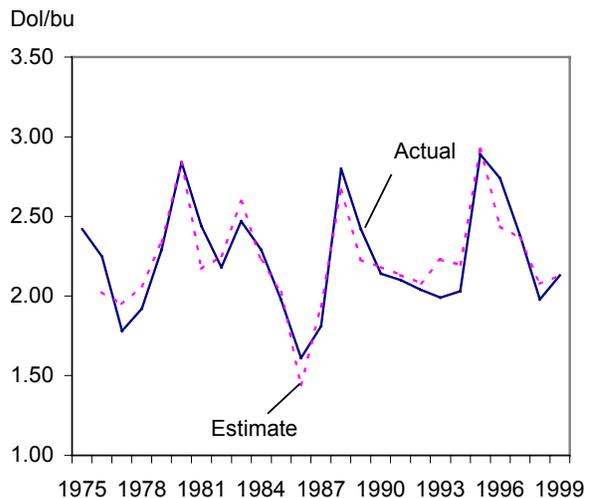
Finally, the Durbin-Watson statistic indicates that first order autocorrelation is not a problem.

As was the case with sorghum and oats, the corn price is important in determining the price of barley. Corn and barley prices have a correlation of 0.817. There is more variation between barley and corn prices than was the case with sorghum and corn because the barley price used here is a composite of both malting and feed barley.

The stocks-to-use ratio was not statistically significant in the barley model and was therefore not included in the equation. The barley stocks estimate maintained by USDA combines both malting and feed barley, and this may explain why the stocks-to-use ratio was not statistically significant. Malting barley stocks are clearly important in determining malting price, and feed barley stocks are important in determining feed barley prices. However, the correlation with price may be diminished when feed and malting barley stocks are combined.

Two new variables are introduced into the barley pricing model. The first ($(FEU+EX)/USE$) approximates the ratio of total feed use to total use because the vast majority of barley exports is feed barley. This is an important variable because it differentiates between feed and malting use. Since feed barley is cheaper, than malting barley, when this variable increases (decreases) the all-barley price will tend to fall (rise). The other variable that we introduced ($Ln(S/S_{-1})$) measures the percent change in supply. The results show that when

Table 3--Turning point errors for the commodity pricing models

	Lagged actual prices are used in the turning point error definition	Lagged predicted prices are used in the turning point error definition
Sorghum	2 (1990, 1999)	1 (1990)
Barley	1 (1993)	3 (1982, 1993, 1994)
Oats	4 (1976, 1981, 1984, 1996)	6 (1981, 1984, 1985, 1992, 1996, 1999)

this variable increases (decreases) barley prices will tend to fall (rise).

Evaluation of the Price Models

Figures 1, 2, and 3 show graphs of historical prices for sorghum, barley, and oats over the model estimation period of 1975-1999, along with the predicted values derived from estimated equations (1), (2), and (3). In general, the price models track actual prices quite well. The sorghum model has the tightest fit followed by barley and then oats. Most differences between the model estimates and the actual prices are less than 15 cents for sorghum, 25 cents for barley, and 30 cents for oats.

The models capture turning points fairly well. A turning point error can be defined statistically when the inequalities in (4) and (5) hold.

$$(4)\ (Predicted_t - Actual_{t-1})(Actual_t - Actual_{t-1}) < 0$$

$$or$$

$$(5)\ (Predicted_t - Predicted_{t-1})(Actual_t - Actual_{t-1}) < 0$$

Predicted prices are derived from the models, and actual prices are those prices received by farmers as reported by the National Agricultural Statistics Service. The subscripts "t" and "t-1" represent current and lagged time periods, respectively. Defined in this way, the statistic measures whether predicted year-to-year changes from the models are directionally the same as changes in actual prices. Turning-point errors can occur in two ways: first, when actual prices indicate a turning point but predicted prices do not and, second, when actual prices do not indicate a turning point but predicted prices show a turning point. The different definitions for the occurrence of a turning point in

equations 4 and 5 relate to whether the change in the predicted price is measured relative to the previous year's actual price (equation 4) or the previous year's predicted price (equation 5). Both measures are useful, but the appropriate measure depends on the intended use of the model. For short term forecasting applications of the models where the previous year's actual price is known, the former measure is more appropriate. For longer term forecasts where the previous year's actual price is not known, the latter definition is better. Since these price models are intended for both short term and long term forecasting, both definitions will be used. A breakdown of the number of turning point errors and the years in which they occurred is provided in table 3. In general, the sorghum model had the fewest turning point errors followed by barley and oats.

Table 4 shows the mean absolute errors, standard deviation, and mean absolute percentage errors for the three models for the full estimation period (1975-1999) and a selected sub-sample covering the decade of the 1990's. The general results show that sorghum was the

Figure 4
**Sorghum: Model performance using September and January WASDE**

Dol/bu
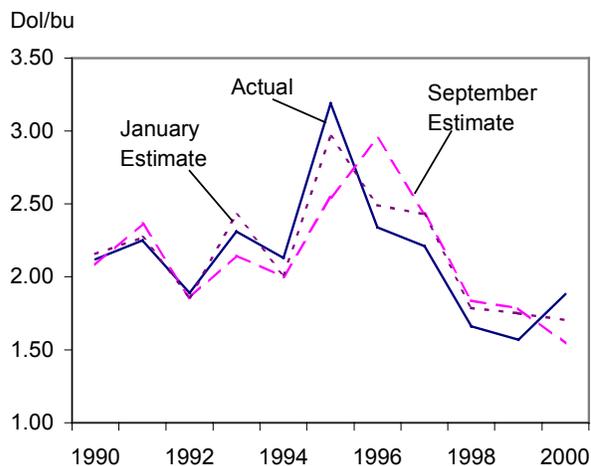
Table 4--Sorghum, barley, and oats model performance measures, selected periods

Time Period	Mean absolute error			Standard deviation			Mean absolute percentage error		
	Sorghum	Barley	Oats	Sorghum	Barley	Oats	Sorghum	Barley	Oats
	Cents per bushel			*Cents per bushel*			*Percent*		
1975-1999	6.7	11.5	13.6	8.5	14.6	17.0	3.1	5.3	9.2
1990-1999	7.4	9.7	15.9	9.2	14.1	19.2	3.5	4.4	11.2

strongest model, followed by barley and oats. This result mirrors the results from the statistical analysis (especially the R^2 statistic) as well as the turning point analysis described above. Interestingly, model performance in the 1990's has not been as good as for the full sample period for sorghum and oats but has been better for barley.

We also analyzed model performance when the independent variables were themselves unknown. This mimics the situation of the World Agricultural Outlook Board, which develops price forecasts based on projected supply and demand quantities in a marketing year. To do this, we took past World Agricultural Supply and Demand Estimates (WASDE) reports for the months of September and January during 1990-2000 and used the forecasted variables in these reports as the independent variables in the models. We stopped with the January WASDE because we wanted to focus on reports that are early in the crop year. Figures 4, 5, and

6 compare these forecasted prices (using information provided in the September and January WASDE reports) to the actual season average farm price for that crop year.

In general, the model did a pretty good job of forecasting price. The use of forecasted independent variables obviously increased the forecast error, but the models would appear to be useful for forecasting prices. For the September forecasts, the barley model performed better than the sorghum or oats models. The high forecast error for sorghum in 1995 and 1996 is the main reason for the relatively poor performance of the sorghum model. Note that the sorghum forecast is based to a large degree on the corn market situation at the time, which changed later in the crop year. In September 1995, the 1995/96 corn crop was projected at 7.8 billion bushels and the corn price was projected at $2.55-$2.95. However, the actual 1995/96 corn crop was 7.4 billion bushels—nearly 6 percent lower than the

Figure 5

**Oats: Model performance using September and January WASDE**
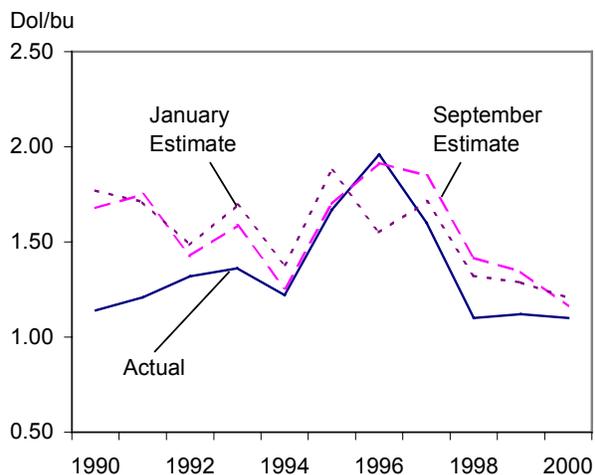
Figure 6

Barley: Model performance using September and January WASDE
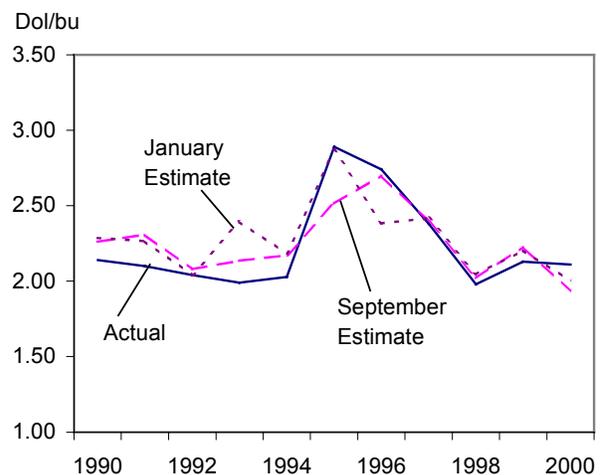
Table 5--Sorghum, barley, and oats model performance measures using WASDE forecasts

Forecasts as of:	Mean absolute error			Standard deviation			Mean absolute percentage error		
	Sorghum	Barley	Oats	Sorghum	Barley	Oats	Sorghum	Barley	Oats
	Cents per bushel			Cents per bushel			Percent		
September WASDE	24.4	12.7	21.6	31.4	16.4	18.8	10.9	5.6	17.6
January WASDE	12.7	13.9	27.4	14.3	18.0	24.9	6.1	6.3	21.1

September projection—and the actual corn price ended up at $3.24. Because of this, the model understated in September what the final sorghum price would be.

A similar issue arose with the 1996/97 sorghum crop. As of September 1996, corn production was forecast at 8.8 billion bushels and price was projected at $3.00-$3.40. However, actual 1996/97 production was 9.2 billion bushels and this larger supply lowered the corn price to $2.71, which had a bearish impact on the sorghum market.

Another issue to point out is that the oats forecast seems to be biased above the actual oats price (this is true for both the September and January estimates). For forecasting purposes, it is important to keep in mind that the oats model has a relatively large forecast error and seems to have an upward bias.

With respect to the January estimates, the sorghum model performed best followed by barley and then oats. For sorghum, the January forecast was significantly improved over the forecast using September estimates because the projected season average corn price in January for 1995/96 and 1996/97 is much closer to the actual corn price.

Interestingly, the barley and oats models had a higher forecast error than was the case previously. This was unexpected because market information is more complete in January than it is in September (this is particularly true for barley and oats because their crop year ends on May 31). Still, the independent variables used are forecasts. Also, note that the standard error for the barley and oats regressions are reasonably low but still larger than was the case for sorghum. This suggests that there is additional market information that is not reflected in the models (this is especially true for oats).

Conclusions

This paper analyzed some important factors that determine prices for the minor feed grains. The models analyzed here utilize the price of corn because corn is the primary feed grain and the corn supply and demand situation has a profound impact on all of the feed grains. The food and industrial demand component is especially important for barley and oats. Grains used for food and some industrial purposes are generally of higher quality and provide more value to final users. Therefore, the statistical model for barley and oats incorporate variables that adjust for these effects, and illustrate that food and industrial uses of these grains are very important in determining the average price of those commodities. These variables also highlight important market differences between the minor feed grains.

In general, the sorghum and barley models did a better job of forecasting prices than did the oats model. These models had higher R^2 statistics, fewer turning point errors, and lower mean absolute deviation. Forecast error increased for all models when they were tested with forecasted independent variables, but again the sorghum and barley models generally outperformed the oats model.

Even with possible forecast errors, these models provide a useful foundation for analyzing sorghum, barley, and oats prices. The relatively simple structure of the models and their small data requirements lend themselves to price-forecasting applications in conjunction with market analysis of supply and demand conditions. These models provide a statistical framework for forecasting prices as well as a tool for making consistency checks for supply, demand, and price forecasts.

This page intentionally left blank.

Rational Commodity Forecasts: Improving USDA's Cotton Analysis

Stephen MacDonald, Economic Research Service, USDA

Introduction

Agricultural markets differ from other commodity markets in having the U.S. Government play a significant role in the creation, validation, and publication of market information. Crop surveys by USDA's National Agricultural Statistics Service (NASS) are virtually the sole source of trusted information about the state of U.S. major field crops during the growing season. At the same time, USDA publishes estimates of U.S. and world supply and demand that are in some cases unique, but in virtually all cases are the benchmarks for other estimates (Vogel and Bange). In addition to forecasts, USDA maintains historical estimates of these variables, in the case of cotton supply and demand, going back to 1960.

In addition to cotton, USDA updates forecasts and historical estimates of supply and demand for 33 other commodities, with balance sheets for more than 100 countries in some cases. This process is coordinated by the World Agricultural Outlook Board (WAOB), which chairs 9 commodity committees with representatives from other USDA agencies: the Economic Research Service (ERS), Foreign Agricultural Service (FAS), Farm Service Agency (FSA), and Agricultural Marketing Service (AMS).

With the revolution in telecommunications of the last few decades, the market's reliance on USDA information, and the staff-years required by USDA to produce it, has lessened. However, periodically questions arise about the declining contributions of various agencies to this process, and the ability of the system to continue functioning in a changing environment requires frequent reexamination of methodologies.

This study assesses the accuracy of USDA's world cotton forecasts in an effort to develop templates for incorporating information about USDA's past forecasts into its current forecasts. For the first time, a database of USDA's cotton forecasts by country since 1993 has been developed, and the rationality of USDA's forecasts could eventually be improved if these past forecasts became an active part of the information set used by USDA's forecasters.

Utilizing Past Forecasts

Studies of forecast rationality typically assume that past forecasts are included in the forecaster's information set. While intuition suggests this is a reasonable assumption, collecting and reviewing information is never cost-less. Since past forecasts only become useful after a significant sample has been accumulated the discounted value of archiving forecasts may be relatively small. Developing a database from past published forecasts is also difficult: published documents do not include historical data for many variables. Historical data can on occasion be revised significantly, perhaps changing the meaning of the forecast. Thus, with limited resources, it may not have been rational to maintain or develop large databases of forecasts. The WAOB maintains a database of forecasts published in the World Agricultural Supply and Demand Estimates (WASDE), but only for a subset of countries and years.

Since 1989, USDA has electronically published its full international commodity databases each month. Since 1993, the agency publishing this data, ERS, has saved copies for virtually every month. Since the files saved include the full historical estimates current at that time as well as the forecasts of the coming year, the information set used in the forecast is more completely represented. This is true for other commodities as well as cotton, and the experience in putting the database together and the tools used to analyze the forecasts could prove useful to USDA's forecasts of other commodities.

For the record, there were some problems with stored records—no records were found for October 1993, March 1999, and October 2001. Some data were corrupted in one month's records (August 1993), and naming conventions changed over time, meaning additional work will be necessary to create database queries that can draw time series for certain countries over the entire period. Some of these problems may not occur in the records for the other commodities.

Assessing USDA's Cotton Forecasting

USDA's annual cotton data are based on an August-July marketing year, based on the

cropping cycle of the Northern Hemisphere, which accounts for 90 percent of the world's cotton output. Every July, USDA publishes its first estimates of cotton supply and demand by country for the coming year.

This study is largely confined to analysis of USDA's total world forecasts. Ultimately, the forecasts should be assessed on a country by country basis since USDA's world forecasts are largely the simple sum of its individual country forecasts. However, rather than try to assess and summarize the roughly 6000 forecasts this would entail (12 months, 100 countries, 5 variables), this study analyzes various accuracy measures and methodologies for the total world estimates that can later be applied to the country level estimates.

Figure 1 provides an overview of the database, illustrating the errors in USDA's world consumption forecasts and estimates through their first 12 revisions for marketing years 1993-99. Forecast errors likely have a smaller range over a long period than do forecast levels if a variable has a steady trend, and looking directly at the errors helps intuit accuracy. Each point in the graph indicates how that month's estimate of that year's world cotton consumption differs from USDA's current estimate for that year as of April 2002. Each line represents the first 12 months of USDA's forecasts and estimates for world cotton consumption for a given marketing year, beginning with 1993 and ending with 1999. Marketing year 2000 cannot be assessed since experience has shown revisions will be large through July 2002.

From this graph, one can make three observations about the accuracy of USDA's world cotton consumption forecasts in 1993-99:

1. Forecasts tend to be too high at the beginning of the year
2. They are more accurate later in the year, although non-zero errors are common
3. They tend to be too low at the end of the year

Historical Revisions

In many cases there is no source for "final" estimates, particularly for ending stocks, but also


Figure 1: World Consumption Forecast Errors
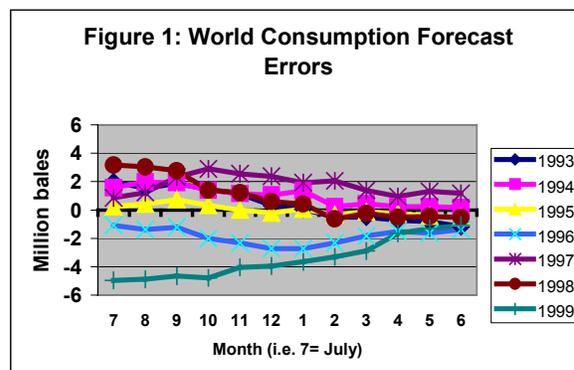
for consumption. In the United States, government agencies collect and publish data on cotton production, trade, consumption, and ending stocks. For some countries, industry associations can be relied upon for estimates of production and consumption, and in those cases USDA's estimated ending stocks are the residual of the rest of the balance sheet. However, in many cases there is no authoritative source of historical consumption data. In these cases, consumption and ending stocks are estimated together, based on production and trade data, which are more commonly available.

For these countries, new information, or the perception that estimated cotton supplies in a given country are wrong due to cumulative errors in consumption estimates, can result in significant revision of the "final" estimates years after the fact. Occasionally sources of production data may change as well.

Figure 2 illustrates these points. By different definitions of reality, USDA may or may have been extraordinarily inaccurate in its 1995 estimate of world consumption. Errors in the graph are the Forecast minus the Actual, with actual defined as the 60th estimate, in this case the June 2000 estimate of 1995 consumption. The year markers indicate July of every year.

The plain line (labeled "World") traces out the smallest errors of the 3 lines during the first few months of forecasting, but then suggests USDA still overestimated 1995 world consumption by about 1 million bales (about 1 percent) several years after 1995 ended. The line with diamond markers (labeled "W – China"), which is the world excluding China, indicates that by May 1997 there was very little error in USDA's estimate for total consumption outside of China. However, it suggests that through February 1997 there was a large, 2 percent, error.

Figure 2: World Cotton Consumption Errors, 1995
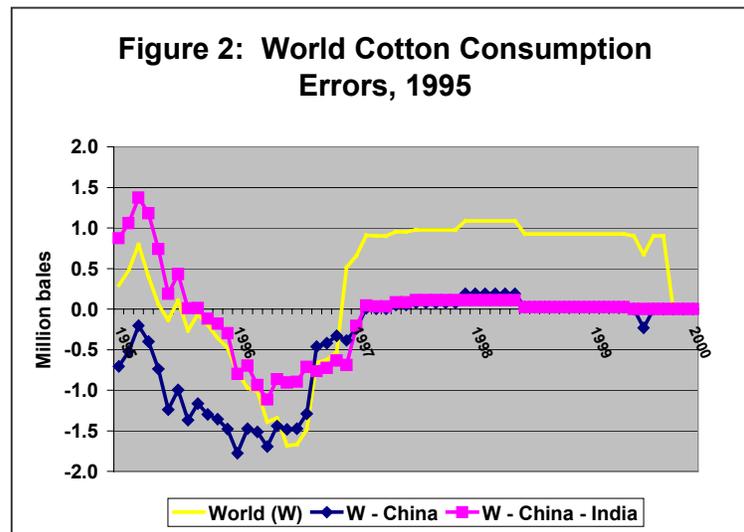
The line with square markers (labeled "W – China – India), which is the world minus China and India, differs because India's 1995 consumption estimate was revised 8 percent in March 1997, as part of a long term revision of the data for India. While this line maintains a large inaccuracy several months later than the second line (until July 1997), it is more accurate during most of the preceding period.

This illustrates the lags in finalizing estimates even in the absence of long-term revisions, and the impact of these revisions. Typically, a given marketing year's consumption is revised significantly through its 25th estimate.

The long-term revisions raise interesting questions about measuring accuracy: if USDA is forecasting within a certain information set which is invalidated years later, should forecast performance be judged solely within that set? Or should the long run error be included in the measure of accuracy as well?

Time Series Properties

1993-99 was a period of stagnant growth in global cotton consumption and production, so it is not obvious that the variables analyzed are non-stationary: here they may appear to be mean-reverting, but over longer periods this does not seem to be case. Other commodities, like grains and oilseeds, have more consistent world consumption gains, and country-level cotton variables also often have strong trends. Given the small sample size (7 to 8) it would be difficult to formally test for non-stationarity, let alone estimate cointegrating relationships, but

the time series properties of the series being forecast cannot be completely ignored.

First differencing, or examining forecasts as changes rather than as levels, is a good compromise between the need to avoid spurious regressions and the lack of sufficient data to formally measure all aspects of the relationship between the forecasts and the actual realizations (Granger and Newbold).

First differencing also may abstract from some of the problems raised by long-term adjustments to the estimates of the actual realizations. For example, the revisions of cumulative small errors will have a greater impact on the annual levels than on the trends, so USDA's ability to predict the presence and magnitude of turning points could still be analyzed through first differences.

In the cases examined here, early season forecasts were more highly correlated with realizations when examined as first differences than as levels. This would not be the case with highly non-stationary variables and the decision to examine forecasts in terms of differences or levels will have to be made in the context of how to best use the information provided to improve future forecasts.

USDA publishes its forecasts as levels, so a given month's forecasted change was calculated as the difference between a given month's forecast and that same month's estimate of the previous year. This is preferable to using the current estimate of that previous year due to the possibility of long term revisions since that initial publication.

Results

Tables 1 and 2 illustrate the accuracy of USDA's world cotton estimates over their first year during 1993-99. Table 1 summarizes for the initial July forecasts and Table 2 summarizes for the estimates published the following June for the same marketing years. Virtually every measure indicates improved accuracy from the first forecast to the twelfth: Mean Absolute Percent Error (MAPE) declines, correlation rises, and Mean Percent Error (MPE) shrinks, except for ending stocks.

Table 1--First World Forecast's Accuracy, 1993-99

Variable	MAPE	Correlation	MPE
		Percent	
Area	2.0	92	0.8
Production	3.5	77	1.5
Consumption	2.6	43	1.1
Imports	5.9	71	-1.2
Exports	4.3	69	0.7
Ending Stocks	12.2	34	-1.8

Table 2--Twelfth World Forecast's Accuracy, 1993-99

Variable	MAPE	Correlation	MPE
		Percent	
Area	0.7	99	-0.5
Production	1.3	99	-1.2
Consumption	1.0	94	-0.7
Imports	3.0	96	-0.6
Exports	1.2	99	0.4
Ending Stocks	3.3	99	-3.3

MAPE is similar to Root Mean Square Error (RMSE), but gives each error equal weight. Correlation provides further insight into the ability of forecasts to track realizations, although it can be influenced by variability. MPE indicates simple bias. MAPE is probably the most important of the three measures.

Area is the most accurate initial forecast, since cotton has largely been already planted by July. Ending stocks is the least accurate. Ending stocks are largely estimated as residuals to the other components of the balance sheet, and apparently their errors are not offsetting. Note that while their MPEs indicate that production and consumption have largely offsetting errors, the import and export errors would both serve to drive ending stocks lower.

Interestingly, imports are the second least accurate variable, according to MAPE. This study was undertaken in part with this hypothesis in mind--that inefficient forecasts of ending stocks result in inaccurate estimates of imports by USDA. While historical data exist for imports, trade is typically more variable than consumption, making forecasts difficult. The correlation for consumption is lower than the correlation for imports, so it is not unequivocally clear that import forecasts are less accurate than consumption. However, by the 12th forecast, imports begin to vie with ending stocks for the rank of least accurate.

Ending stocks are typically too low by about 3 percent even in the twelfth month, but note that none of these biases are statistically significant. The source of this ending stock error bears examination, although, as the next tables show, it does not stem from the significant revisions to India's and China's data.

Table 3--First World (minus China & India) Forecast's Accuracy, 1993-99

Variable	MAPE	Correlation	MPE
		Percent	
Area	2.2	97	1.5
Production	4.2	89	4.2
Consumption	2.4	73	1.8
Imports	3.1	88	1.2
Exports	3.2	85	1.0
Ending Stocks	8.2	72	6.9

Table 4--Twelfth World (minus China & India) Forecast's Accuracy, 1993-99

Variable	MAPE	Correlation	MPE
		Percent	
Area	1.0	98	-0.1
Production	0.8	100	-0.8
Consumption	0.9	92	-0.2
Imports	2.4	86	-0.5
Exports	3.4	92	0.6
Ending Stocks	4.1	97	-3.5

Removing India and China results in a smaller range of errors across the balance sheet in the first month for MAPE and correlation, and a few slightly higher errors in the twelfth estimate. The biggest deterioration in the 12th estimate comes in exports, perhaps reflecting China's role as a swing exporter.

The understatement of stocks in the 12th estimate remains above 3 percent, indicating the problem is not simply a function of long-term

revisions to China and India. It is possibly the result of long-term revisions to other countries, but even if this is the case it would be worthwhile to consider why long-term revisions tend to be in the same direction.

Production tends to be 4.2 percent too high in the first estimate (again, not in a statistically significant manner), accounting, perhaps, for overestimated stocks. Note that area is only overstated by 1.5 percent, consistent with USDA's normal weather assumption, or with the notion that negative yield shocks have greater unanticipated consequences than do positive ones.

Imports' first-estimate errors shrink with the removal of China and India as measured by MAPE and correlation, and the bias reverses. Evidently, USDA is underestimating China's and India's imports at the beginning of the year, but overestimating those of the rest of the world.

By the twelfth estimate, imports are slightly underestimated and the impact of this on ending stocks is reinforced by slightly overstated exports, just as they were when China and India were included. Similarly, an underestimation of yields is further depressing the stock estimate, with and without China and India.

Efficiency

The behavior of forecast errors can be examined by more sophisticated means. The most sophisticated would be to determine if other information available at the time the forecast is published could reduce errors. This is typically referred to as testing for "strong-form" efficiency. A significant hurdle to doing this is establishing a database of expectations for other variables--it can be daunting enough to establish and use a database of the forecasts one wishes to study. However, the large number of forecasts in this database can be used for both purposes. Furthermore, similar databases for other commodities may prove even more useful for this purpose. This is an area for future research and is not explored here.

Two "weak-form" efficiency or rationality tests were performed: checking for serial correlation of the errors and for the presence of forecast-improving linear transformations. Serial correlation is regarded as irrational in the literature, although the discussion earlier in this paper argues for a more bounded definition of rationality. Serial correlation was measured by Durbin-Watson statistics. Forecast-improving linear transformations were found by attempting to reject the joint hypothesis that a=0 and b=1 in the regression A = a + bF, where A = the actual realization of the variable and F = USDA's forecast (Theil). Note that b should also be significantly different from zero for the test to indicate inefficiency rather than inaccuracy.

Table 5--Inefficient Forecasts, World, 1993-99

Variable	Month	Linear	Serial
Imports	11	Yes	--
Imports	12	Yes	--
Imports	1	Yes	--
Exports	1	Yes	--
E. Stocks	2	Yes	--
E. Stocks	4	Yes	Positive
E. Stocks	5	Yes	--
E. Stocks	6	Yes	--

Table 6--Inefficient Forecasts, World minus China & India, 1993-99

Variable	Month	Linear	Serial
Area	9	Yes	--
Area	10	Yes	--
Production	11	--	Positive
Consumption	11	Yes	Positive
E. Stocks	11	--	Positive
E. Stocks	5	Yes	--
E. Stocks	6	Yes	--

Tables 5 and 6 summarize the instances of forecasts with significant (5 percent level) inefficiency. Significant forecast-improving linear transformations were found for less than 10 percent of this set of forecasts, mostly for ending stocks. There were even fewer instances of significant Durbin-Watson statistics.

The presence of unit roots would reduce the reliability of the standard statistical tests on parameter values used here. Thus, in some cases where a forecast fails an efficiency test, the forecast may in fact be efficient. As noted earlier, the small sample size makes it difficult to address this possibility.

Excluding the ending stock forecasts, inefficiency seems most prevalent during the middle of the year, November to January. This might reflect a point in the forecasting cycle when the forecasts are accurate enough that random errors have been reduced but sufficient

information is not yet available to overcome some persistent errors by USDA.

Some intuition regarding this inefficiency can be obtained by noting that when a forecast is expressed in terms of change—and the mean of the forecast equals the mean of actual change—then the parameter estimate for b in A = a + bF indicates a tendency to over- or under-estimate the magnitude of change. If b > 1, then change is under-estimated, or forecast conservatively.

Inefficient forecasts of world exports might stem from conservative forecasting. The global levels of imports and exports in a given year are strongly related, so import forecasts might also be conservative. Imports and exports have different global totals due to data collection problems common to all merchandise trade and the tendency of cotton to gain weight during shipping. The difference has been as great as 8 percent of world trade in recent years and has even changed sign (typically, imports are larger). Thus the performance of USDA's forecasts of imports and exports differ despite the strong relationship between them.

For January world exports, a regression of actual changes on forecast changes, A = a + bF, yields the following estimates (standard errors in parentheses):

$$A = 8.35 + 1.44 \, F$$
$$(180.5) \, (0.13)$$

The estimate b = 1.44 suggests that USDA's January forecasts understated annual changes by 30 percent during 1993-99. The low MPE of this forecast (0.4 percent), and the even lower estimate for a (0.03 percent of the mean of 1993-99 exports), suggests that b = 1.44 can indeed be interpreted as indicating underestimated change. Further research is necessary to explain this tendency.

Conclusions

Assessing forecast accuracy is not necessarily straight-forward. The choice of how to define the forecast is important, and for global commodity supply and demand data there are important choices about how to define the final realization of variables. If forecasters' performance is to be judged by forecast accuracy, these choices have to be examined and clearly articulated.

World estimate totals embody offsetting errors in the estimates of their component countries and an initial examination indicates that most country level forecasts achieve lower accuracy. The desire to forecast in round numbers may also introduce errors-in-variables problems into efficiency tests of forecasts for small countries, interfering with the parameter estimates. Furthermore, examination at the country level will necessitate even closer attention to the problems introduced by long-term revisions.

The slight (not significant) bias in the world ending stock estimates may indicate statistically significant bias at the country level. If so, then it should be determined if this is due to long-term adjustments or disappears shortly after the 12th monthly forecast. Alternatively, the insignificant global bias may not be attributable to specific countries, but could represent a slight tendency across a large number of country forecasts. This issue is worth addressing regardless of its source.

Increasing the sample of forecasts could change some results, which readers should bear in mind. For example, using the WAOB's database to extend the sample back to 1982 results in the July world production forecast's MPE dropping from 1.5 percent to zero, and consumption's from 1.1 percent to 0.2 percent. Interestingly, the MPE for ending stocks remains at −1.8 percent even for this larger sample.

In addition to extending this analysis to country level forecasts, extending efficiency tests to "strong-form" tests should be pursued. Also, examining the behavior of the forecasts beyond the 12-month period would be useful, as well as analyzing the role of previous year errors on current year errors.

References

Granger, C.W.J. and Newbold, Paul, *Forecasting Economic Time Series*, London, Academic Press. 1986.

Theil, Henri, *Economic Forecasts and Policy*, Amsterdam, North Holland. 1958.

Vogel, Frederic, and Bange, Gerald, *Understanding USDA Crop Forecasts*, Miscellaneous Publication No. 1554, U.S. Department of Agriculture. March 1999.

FEDERAL
FORECASTERS
CONFERENCE