PAPERS
and
PROCEEDINGS

the 14th

FEDERAL FORECASTERS CONFERENCE

2005

Sponsored by

Bureau of Health Professions
Bureau of Labor Statistics
Bureau of Transportation Statistics
Department of Veterans Affairs
Economic Research Service
Internal Revenue Service
International Trade Administration
National Center for Education Statistics
U.S. Census Bureau
U.S. Geological Survey
U. S. Postal Service

~Partnering Organization~
Center for Economic Research,
The George Washington University

ANNOUNCEMENT

The 15th
Federal Forecasters Conference
FFC/2006

Will be held

September 28, 2006

In

Washington, DC

More information will be available in the coming months.

Federal Forecasters Conference - 2005

Papers and Proceedings

Edited by

Mitra Toossi, Chair
Bureau of Labor Statistics
U.S. Department of Labor
Washington, DC

Bill Lawhorn
Bureau of Labor Statistics
U.S. Department of Labor
Washington, DC

Willliam J. Hussar
National Center for Education Statistics
U.S. Department of Education
Washington, DC

Edited and Prepared by

Marybeth Matthews
Marion Kolsch
Healthcare Analysis and Information Group
Veterans Health Administration
Office of the Assistant Deputy Under Secretary for Health
for Policy and Planning
U.S. Department of Veterans Affairs
Milwaukee, Wisconsin

Donald Stockford
Policy Analysis Service
Veterans Health Administration
Office of the Assistant Deputy Under Secretary for Health
for Policy and Planning
U.S. Department of Veterans Affairs
Washington, DC

Federal Forecasters Conference Governing Board

Federal Forecasters Consortium Governing Board

(Front Row) **Norman C. Saunders**, Bureau of Labor Statistics; **Caribert Irazi**, U.S. Census Bureau; **Brian Sloboda**, Bureau of Transportation Statistics; **Kathleen Sorensen**, National Cemetery Administration; **Peg Young**, Bureau of Transportation Statistics; and **Elliott Levy,** International Trade Administration

(Back Row) **William Hussar**, National Center for Education Statistics; **Donald Stockford**, Department of Veterans Affairs; **Bill Lawhorn**, Bureau of Labor Statistics; **Mitra Toossi,** Bureau of Labor Statistics; **Russell Geiman,** Internal Revenue Service; **Malcolm C. Harris**, U.S. Post Office; and **Jeff Busse**, U.S. Geological Service

(Not Pictured) **Frederick L. Joutz**, The George Washington University; **Stephen A. MacDonald**, Economic Research Service; **Ching-li Wang**, U.S. Census Bureau; **Louis Kuta**, National Center for Health Workforce Analysis, and **Terri Manzi**, Internal Revenue Service

Foreword

The 14th Federal Forecasters Conference (FFC/2005) was held April 21, 2005 in Washington, DC, and was a great success. In fact, FFC/2005 has continued a long string of successful conferences that began in 1988 and that have brought wide recognition to the importance of forecasting as a major statistical activity within the federal government and among some of its partner organizations. Over the years, the federal forecasters conferences have succeeded most at providing a forum for practitioners and others interested in the field to organize, meet, and share information on forecasting data and methods, the quality and performance of forecasts, and major issues impacting federal forecasts. In particular, the theme of FFC/2005 was "Capturing the Impact of International Trends in Our Forecasts," in recognition of increasing globalization and the growing impact of global market forces, international developments, and cross-national interdependencies in forecasting activities related to population, migration, economics, business, technology, education, the military, and politics.

As part of the lead-in to FFC/2005, federal department, agency, and other sponsors of the federal forecasters conferences re-organized themselves into the Federal Forecasters Consortium, evidencing the increasing levels of collaboration and networking activities of a core group of federal forecasters. The need for a consortium has become increasingly clear over time and attests to the evolution of federal forecasting activities and the strengthening identification that federal forecasters have with what they do and how they do it as well as their concern for who, what, when, where, how, and why their forecasts impact. Members of the Federal Forecasters Consortium agree that we can determine alternative futures, despite myriad uncertainties, with forecasts and perhaps ensure we select the best future possible for everyone. Our forecasts can excite the imagination, raise fundamental questions, help us understand complex systems and deal with uncertain realities, and actually provide us control over the way things turn out in the end.

The papers and presentations in this FFC/2005 proceedings volume relate to forecasting demographic and attitude shifts, health care environment trends, and socio-economic and business-related factors and events, and contribute to the ultimate goal of attaining meaningful, timely, and visionary forecasts that can ensure a better future for all. After all is said and done, this is the truly great success of FFC/2005 and the truly great promise of FFC/2006.

Acknowledgements

Many individuals contributed to the success of the 14th Federal Forecasters Conference (FFC/2005). First and foremost, without the support of the cosponsoring agencies and the dedication of the Federal Forecasters Consortium Governing Board, FFC/2005 would not have been possible.

Mitra Toossi of the Bureau of Labor Statistics (BLS) chaired the morning program. Kathleen Utgoff, Commissioner of BLS, gave the welcoming remarks. Brian Sloboda, of the Bureau of Transportation Statistics, presented certificates to the winners of the FFC/2004 and FFC/2005 forecasting contests. Russell Geiman, of the Internal Revenue Service, and Jeff Busse, of the U.S. Geological Survey, announced the FFC/2003 best conference paper awards. Kathleen Sorensen, of the U.S. Department of Veterans Affairs, organized the morning panel. All the members of the Federal Forecasters Organizing Committee worked hard to provide support for the various aspects of the conference, making it the success it was.

Special thanks go to Brian Boulier, Robert Trost, and Frederick L. Joutz of The George Washington University for reviewing the papers presented at the 13th Federal Forecasters Conference and selecting the winners of the Best Conference Paper awards for FFC/2003.

Special thanks go to Gwendolyn Coleman, Vanessa Sandige, Lilia George and Erma McCray, all of the Economic Research Service, for directing the organization of materials into conference packets and staffing the registration desk. In addition, special thanks also go to Mary Jane Slagle from the Census Bureau for managing the FFC 2005 name tags.

Special thanks go to Marybeth Matthews and Marion Kolsch of the U.S. Department of Veterans Affairs for producing the conference program, the FFC Directory, and this publication, the 2005 papers and proceedings. Also, special thanks go to Donald Stockford of the U.S. Department of Veterans Affairs for his related efforts. Additionally, special thanks also go to the staff of the Bureau of Labor Statistics Conference and Training Center, who once again helped to make the day go smoothly.

Finally, we thank all of the presenters, discussants, and attendees whose participation made FFC/2005 another successful conference.

2005
Federal Forecasters Conference
Forecasting Contest

WINNER

Terry Schau
U.S. Department of Labor

First Runner Up

Mirko Novakovic
U.S. Department of Labor

Second Runner Up

Peggy Podolak
U.S. Department of Energy

2003
Best Conference Paper

WINNER

"Data Consistency Issues in Projecting Births and Population Under Age 1 by Race"

Myoung-Ouk Kim and Ching-li Wang
Population Division, U.S. Census Bureau

Honorable Mention

"Forecasting the Counter-Cyclical Payment Rate for U.S. Corn: An Application of the Futures Price Forecasting Model"

Linwood Hoffman
Economic Research Service

"Business Cycle Analysis for the U.S. Transportation Sector"

Kajal Lahiri and Wenxiong Yae (Suny-Albany)
And Peg Young (Bureau of Transportation Statistics)
Bureau of Labor Statistics

Charter of the Federal Forecasters Consortium

The Federal Forecasters Consortium is a collaborative effort of agencies in the United States Government, as well as other interested parties in the academic and not-for-profit communities, who share an interest in the practice, planning, and use of forecasting activities by and within the Federal Government. In this context forecasting is taken to mean advance planning, decision-making, and the description of expected outcomes, all for unknown future situations. The art of forecasting encompasses many disciplines and utilizes many tools, all applied with the intent of predicting and evaluating alternative futures.

The Consortium provides an environment in which forecasters can network, present papers, take courses, attend seminars, and otherwise improve their ability to prepare meaningful and timely forecasts of occurrences in today's complex and changing world.

The primary objectives of the Consortium are as follows:

1. To provide a forum for forecasters to exchange information on data issues and data quality, on forecast methodologies, and on evaluation techniques.

2. To promote an ongoing dialogue about various forecasting topics among professionals from a variety of disciplines.

3. To build a core network of professionals whose collaboration furthers the use of forecasting as an important planning tool in the 21st century.

4. To expand the network of forecasters by seeking sponsorship from agencies in all parts of the Government and by actively seeking out and fostering working relationships among government, private, and academic communities of forecasters.

5. To provide both formal and informal opportunities to learn about general forecasting methodologies or about new techniques still in experimental stages.

6. To discuss data presentation and dissemination issues.

Membership

The role of member organizations is to provide support and advice to the Federal Forecasters Consortium Governing Board in promoting, planning, and conducting the periodic Federal Forecasters Conference, annual forecast methodology workshops, and such seminars and presentations as are deemed necessary and useful by the Board.

Any government agency may seek to become a member of the Consortium by satisfying the following criteria:

1. Provide support to the Federal Forecasters Consortium in the form of financial support, in-kind contributions, or person-hour support for the programs of the Consortium.

2. Name one or more representatives to the Consortium Governing Board who shall regularly attend and participate in the meetings of the Consortium.

Any not-for-profit or academic organization with an interest in the purposes and goals of the Consortium may become an associate member of the Consortium by satisfying the same criteria.

While there is no intent to exclude agency representatives from the Governing Board if their management is unwilling or unable to formally commit to support for the organization, we feel that it is equally important for the largest participating agencies to understand, acknowledge, and support in a more formal way the activities of the FFC. If it is not against current policies of these agencies, a Memorandum of Understanding is one appropriate way to show high-level agency support of the Consortium.

Governing Board

The Federal Forecasters Consortium Governing Board shall consist of one or more individuals from each of the member agencies and associate members. These individuals are named to the Board by their respective organization or agency. Those agencies designated as "sponsoring agencies" as of January 1, 2003, shall continue in that role so long as they continue to support the Consortium as they have prior to that date.

The chairperson, recording secretary, and other committee assignments are chosen from and by the Governing Board on an annual basis.

The role of the Governing Board is to plan the annual conference, locate resources to conduct the conference, deliberate on issues affecting its operations, promote collaboration among forecasters, organize and present annual forecasting workshops, and support an ongoing seminar series focusing on topics of interest to forecasters.

The Governing Board will meet at least four times a year and an annual report will be prepared by the Board and provided to all member organizations once each year by the last calendar working day of the month of January.

Contents

FORECASTING STRATEGIC ISSUES FACING THE VETERANS HEALTH ADMINISTRATION I

PROJECTIONS AND ESTIMATES IN SUPPORT OF FEDERAL TAX ADMINISTRATION

FORECASTING KEY ASPECTS OF IMMIGRATION

ADDITIONAL DOCUMENT(S)

Panel Discussants

Don McPartland
Internal Revenue Service
U.S. Department of the Treasury

Data-Driven Tax Administration in a Global Environment

Large multinational corporations are engaged in a vast array of economic activities around the world. As a result, U.S. corporation income tax rules are highly complex and administering them effectively is a significant challenge for IRS. Data analysis is a critical component of tax administration and enforcement. For example, tax return data recently compiled by the Internal Revenue Service (IRS) show that U.S. based corporations are rapidly shifting profits into low-tax countries. While this tax avoidance strategy is not illegal, some methods for accomplishing this income shift may be improper. The IRS is engaged in a number of efforts to detect and address abusive tax practices by large corporations, including making inappropriate use of income-shifting. This presentation by the Acting Director for Strategy, Research and Program Planning for the IRS Large and Mid-Size Business Division highlights the importance of research activities to tax administration, and demonstrates how research can affect strategic and business plans that drive IRS operations.

John H. DeYoung Jr., W. David Menzie, and Pui-Kwan Tse
U.S. Geological Survey
U.S. Department of the Interior

Look Who's Coming to Dinner—China's Voracious Appetite for Minerals

Non-linear increases in consumption of mineral commodities related to stages of economic growth have been identified in several developed and developing countries.

Since the 1980s, economic growth in China has been between 7 and 9 percent annually, doubling the economy every 8 to 10 years. China has been undergoing industrialization, moving through a series of stages that include development of infrastructure, followed by development of light manufacture, development of heavy manufacture, increased consumption of consumer goods, and, finally, by the development of a service economy. Based upon the experiences of the Federal Republic of Germany and Japan during the post-World War II period, and of the Republic of Korea during the period 1970–95, changes appear to begin roughly at 5-year intervals and each of the stages takes about 20 years to complete—with the stages overlapping. During each stage of economic development, consumption of particular mineral commodities rises dramatically.

Rapid changes in mineral (and energy) consumption are creating conditions where reliable information for economic and national security planning and developing public policies will be increasingly important.

Robert Bednarzik
Public Policy Institute
The Georgetown University

Offshoring as an Issue in Forecasting IT Job Growth

This paper discusses the restructuring in the information technology (IT) sector in the United States and what is known about the number and likelihood of IT jobs moving offshore. A comparison of U.S. economic recoveries is undertaken to sort out whether the current slow job growth is related to offshoring. It presents a synthesis of studies that have estimated and forecasted the number of IT sector jobs moving offshore.

James Gillula
Global Insight

Measuring International Economic Impacts on the U.S. within a Global Model

In response to the growing need to understand the impact of international economic developments on the U.S., Global Insight recently developed a new tool to supplement its traditional forecasting work with the Quarterly Model of the U.S. Economy. A Global Scenario Model was designed and constructed to capture the linkages between the U.S., 15 other major countries, and the rest of the world grouped into 7 regions. This presentation will briefly describe the structure of this new model and illustrate how it is being used by presenting the results of economic scenarios that have been analyzed using it.

Jeffrey S. Passel
Pew Hispanic Center / Pew Research Center

International Migration to the United States: Impacts, Measurement, Models, and Forecasts

In the past generation, the importance of international migration as a driver of population change in the United States has increased dramatically. Its impact can be felt not only in the amount of population growth but also in the country's racial/ethnic composition, age structure, and geographic distribution. Yet, U.S. population projections, with rare exceptions, have not been particularly accurate in forecasting migration nor have they incorporated models or assumptions based on either theoretical or realistic understanding of the processes involved in international movement. Forecasters have been hampered by inaccurate measurement of migration to the United States, lack of consensus on appropriate models, and contradictions between official policy and migration realities.

This paper begins with data illustrating the critical nature of international migration for understanding past and prospective demographic change in the United States. It then moves on to describe the problems in measuring international migration and how inaccurate measures for the 1980s and 1990s have hampered forecasting activities. A range of potential models for understanding migration are briefly discussed with their contradictory assertions. Finally, the presentation concludes with an assessment of the practical difficulties facing forecasters relating to international migration and some suggestions for the future.

Conference At A Glance

Morning Session

8:00 a.m. – 12:00 p.m.

Registration

Conference Center Lobby
8:00 a.m. - 9:00 a.m.

Welcome

9:00 a.m. - 9:05 a.m.

Opening Remarks

9:05 a.m. - 9:10 a.m.

Award Presentations

9:10 a.m. - 9:30 a.m.

Panel Discussion

9:30 a.m. - 12:00 p.m.

Afternoon Concurrent Sessions

1:00 p.m. – 4:30 p.m.

Concurrent Sessions I

1:00 p.m. - 2:30 p.m.

Concurrent Sessions II

3:00 p.m. - 4:30 p.m.

CONCURRENT SESSIONS I

This page intentionally left blank.

Research Results From the Bureau of Labor Statistics Projections Program

Session Chair: Norman C. Saunders, Bureau of Labor Statistics (saunders.norman@bls.gov)

Alternative Approaches to Measuring Shortages of Skilled Workers

Michael Horrigan, Assistant Commissioner, Office of Consumer Prices and Price Indexes
Bureau of Labor Statistics (horrigan.michael@bls.gov)

One of the most significant challenges facing the U.S. labor market is the need to accurately measure the gap that may exist—either currently or in the future--between the skill levels required by employers and the skill levels possessed by the labor force. Using Current Population Survey (CPS) data for 1994 and 2000, this article first examines an equilibrium approach that compares the labor market outcomes of observed interactions of labor demand and supply to identify and develop empirical measures of the degree of skill shortages in U.S. labor markets. Several general types of labor market outcomes are explored. Then, to capture the skill dimensions associated with these labor markets, data on occupation and educational attainment are used to identify a new typology of relative skill clusters across occupations--clusters that form a natural hierarchy of occupational groups reflecting increasing levels of skills. The question of whether or not objective data-driven guidance can be provided as to which of these occupations will experience *shortages of high-skilled* workers' in the future is also examined.

A Century of Occupational Change

Ian Wyatt, Bureau of Labor Statistics (wyatt.ian@bls.gov)

The past century has seen an economic transformation. This study attempts to quantify that transformation in terms of its impact on occupational distribution and coverage. Some of the changes and trends are obvious, but the scale of the numeric change is still interesting. Other changes and trends are far from obvious, including a few occupations that have remained surprisingly consistent over the entire period. Earlier data was derived from Census occupation data.

Foreign Trade in Goods and Services: Data Development for Country-Specific Analyses

Mirko Novakovic and Betty W. Su, Bureau of Labor Statistics

Over the past two decades, globalization and international competition have played an important role in U.S. economic activity. Countries have intensified their links with the global economy through trade and investment. This study analyzes the trade data on both exports of goods and services as well as imports of goods and services from 1983 to 2003, and develops detailed descriptions on industry-basis by 14 selected countries and world major regions to examine possible enhancement to that trade data. This study may provide an essential resource for further research to support understanding of outsourcing issues.

This page intentionally left blank.

Foreign Trade in Goods and Services: Data Development for Country-Specific Analyses

Mirko Novakovic and Betty W. Su
Office of Occupational Statistics and Employment Projections
Bureau of Labor Statistics, Department of Labor

Introduction

In classical trade theory and according to the law of comparative advantage, even if a country has an absolute disadvantage in producing certain goods compared to other countries, it will be beneficial for it to engage in trade while specializing in the production of the goods it has a comparative advantage in, as a result of which the total output of the traded goods will rise.

The aim of this paper is to view the exports and imports of goods and services relative to the rest of the world and to try to relate these to the classical comparative advantage theory. Results will then be examined by breaking the "rest of the world" into separate countries. Trade with these countries represents about 80 percent of total U.S. goods trade. Services trade between the U.S. and the selected trading partners will also be evaluated. Finally, effort will be made to address the issue of outsourcing while it is also acknowledged that the measurement of affiliate trade may contribute to more exactness in the determination of the effect of foreign trade on overall employment.

Foreign Trade Data

For this study, foreign trade data were accessed from the Trade Policy Information System (TPIS), a system developed by the International Trade Administration, Department of Commerce. On the basis of the TPIS 10-digit Harmonized code as well as the North American Industry Classification System (NAICS), approximately fifteen thousand export and twenty thousand import merchandise-codes were retrieved for each of the 14 selected countries for the period 1983-2003. For services, the main set of data relied upon is the 1992-2003 set of unaffiliated data provided by the Bureau of Economic Analysis (BEA) of the Department of Commerce. The services data are directly available through the Survey of Current Business. The trade data were then distributed across 200 industries based on the historical input-output table, and then grouped into 11 industrial sectors based on the type-of-produce criterion.

Goods of exports and imports

Table 1 shows the export shares of industries over the years. Attention has been given to observing those industries deemed to be "traditional", such as Industrial sector 1 (agriculture, forestry, fishing and hunting), and Industrial sector 4 (textile product and apparel), and those industries that the United States has an absolute and comparative advantage in producing and exporting, such as Industrial sector 8 (computer, communications, and audio and video equipment), and Industrial sector 9 (semiconductor, navigational, electro-medical, and appliance).

For most goods-producing industries, results meet the expectation set by classical trade theory. Thus, a slight drop in export shares of Industries 1 and 4 may be justified, i.e., it confirms expectations. However, for Industries 8 and 9, the share of exports increased in both industries through 2000, but following a marked decrease for the next 3 years. A further examination is needed. It should be noted that data shown in this article are in nominal terms. Ultimately, one also may need to look at prices and the real exchange rate. Falling prices of computers and the fall of the overvalued dollar would contribute to a decline in the export share of computers. What also comes to mind is the issue of affiliate trade as companies move overseas. While their income does contribute to the overall U.S. balance of payments, the industry drop in trade from the U.S. is a trade and employment loss for the particular industry. The measurement of foreign trade today includes only cross-border trade, which consists of trade with unaffiliated companies or trade within multinationals. Not treated as part of U.S. international transactions is trade between affiliates of multinational companies, which includes the following:

1. Sales to foreigners by foreign affiliates of U.S. companies, and
2. Sales to U.S. residents by affiliates of foreign companies.[1]

Industry import shares are given in Table 2. From 1993 to 2003, shares of imports have risen mainly in Industries 6 and 9. The slightly falling share of imports of Industries 1 and 4 may be unexpected but may also be the result of a low relative value of trade in these items. Thus, the shares might have changed slightly for some industries but in general, rather than shares, one may want to look at growth rates associated with a set of trading partner countries and how these rates may have changed regarding particular industries.

Trade information regarding a vast number of trading partners was considered. The countries chosen represented around 80 percent of U.S. import trade as well as U.S. export trade.[2] Tables 3 and 4 consider total average growth rates of exports and imports of the United States as well as U.S. growth rates with respect to a group of trading partner countries.

The export growth rates are examined for two periods: 1992-2003 and 1997-2003. From Table 3, the average rate of growth of exports prevailing in the period 1992-2003 was 1.4 percent for agriculture, 3.3 percent for textiles, 1.9 percent for computers, and 6.2 percent for semiconductor and appliances. For computers, this growth rate in U.S. trade with China was 13.2 percent annually, from 1992 to 2003. In the same period, the average annual rate of growth of computer exports to India was 17.5 percent. (Low growth rates were marked for exports to some other countries like the European Union, 0.2 percent, and Canada, 1.1 percent.)

As shown in Table 4, imports of goods from all countries for the four goods categories, from 1997 to 2003, grew positively but at a decreasing rate, as compared with that for the period 1992-2003. The relative decline in imports is seen even in textiles. Although the prevailing growth rates in the period 1992-2003 were higher implying that the period 1997 to 2003 had marked only a temporary decline. To examine the results by country, the annual rate of growth of imports of textiles from both India and China were above 9 percent for the period 1992-2003. In the same period, U.S. imports of computers and semiconductors and appliances were strong. The average annual rate of growth of computer imports from China was 28.7 percent, while from India it was 4.9 percent. Semiconductors and appliances imports from China grew at 17.9 percent annually, while from India this rate was somewhat lower at 13.1 percent. It is not entirely explainable why there is a strong growth of imports of these items. Under the assumption of the functioning of the theory of comparative advantage, these goods would be exported rather than the reverse. All this may be subject to interpretation and perhaps suggests the necessity of more rigorous work with the existing data.

Services of exports and imports

When discussing exports and imports services, the first thing to consider is the data availability itself. The main reliance is the unaffiliated data provided by BEA & NIPA (National Income and Product Accounts). The data are directly available through the Survey of Current Business. The modern data concerns could be formulated as follows:

1) It has become an issue as how U.S. jobs are affected by offshoring/outsourcing because information may be nonexistent particularly regarding affiliate purchases in the United States. With measurement of purchases of affiliates, we would have more complete information regarding total employment lost and gained in foreign trade. (i.e., loss of jobs related to outsourcing may be mitigated by purchases/investment of foreign affiliates in the U.S.)

2) Information technology has made it easier for companies to engage in foreign trade than before. While indeed we do have the comparative advantage in trading high-tech goods and services products, with rapid development and transfer of technology, this advantage is perhaps rapidly changing. In particular the other countries may be gaining comparative advantage in their trade of such items as TVs, cars, even computers, as well as in certain computer programming skills.

3) That is why advancement in measurement is necessary to be even more in sync with industrial development. Better measurement will contribute to a more precise industry definition which would likewise contribute to a better understanding of some services issues, as for example the issue as to what kinds of jobs are being outsourced.

Services export and import data are each broken down by NIPA into seven same categories. These are represented in Tables 5 and 6. In the services cited as the sector mostly related to modern technology issues is that of other private services (OPS). The growth and distribution of OPS may themselves be broken down into five different categories. These categories are: financial services, insurance, telecommunications, business professional and technical services (BPT), and other unaffiliated services. Telecommunications and BPT cover legal services, accounting and advertising services, and "other BPT services".

From Tables 5 and 6, we see how OPS have been gaining in significance over the years. Other private services in exports from 1992 to 2003 grew at an average annual rate of 9.3 percent (Table 5). At the same time, the growth rate of imports of OPS was even higher or 11.5 percent (Table 6). Given these high growth rates, the share of OPS in total exports and imports has increased tremendously. The share of OPS in total services exports in 1992 was 27.0 percent; it rose to 41.9 percent in 2003. At the same time, the share of OPS in total services imports was 21.1 percent in 1992 and up to 32.9 percent in 2003.

By connecting to the BEA's website, we obtained OPS export and import data by country, the same countries selected to observe export and import goods transactions. The OPS data were used representing about 60 percent of total OPS data (affiliated and unaffiliated) in the case for both receipts and payments. Our analyses indicate that both receipts from and payments to some of the U.S. trading partners exhibited high OPS growth rates from 1992 to 2003. The results are visible from Table 7 (details of OPS receipts) and Table 8 (details of OPS payments). We see that the growth of unaffiliated trade receipts has been steady, while payments data also show a similar trend.

We see that OPS imports from China grew at a solid rate, 7.5 percent annually from 1992 to 2003. At the same time, OPS imports from India grew at an astounding rate of 21.0 percent. The growth rates information may be a bit more disturbing when looking at the data for the period 1997-2003. U.S. OPS receipts growth to India on an annual basis was 21.0 percent, but the rate of increase of payments to India was 35.6 percent. Upon further observing the existing OPS data, it may be said that the volumes of U.S. foreign trade transactions with India may be relatively low compared to the same foreign trade transaction volumes observed in U.S. trade with Canada, or even with Europe, all of which perhaps makes a dramatic growth rate change in U.S. trade with India more likely. More affiliated data would also be helpful as it would give a better indication as to the total industrial employment created and lost.

Some final thoughts on outsourcing

Trade "conundrum": constant modernization by developed and developing countries. What do we see in today's world? We see developed countries trade industrial goods including computers and other products listed in the IT sector. But, with science and technology developing, developed and developing countries both sell more sophisticated products. For example, while the U.S. focuses on producing more sophisticated products including computers and other machines, the developing countries like India and China produce and export not only agricultural products, textiles, furniture, and household appliances, but also industrial machinery, TVs, and computers, etc.

Services jobs, outsourcing and developed infrastructure. The production and trade of industrial goods throughout the world has always necessitated the existence of a developed infrastructure. The infrastructure represented a formidable cost to investors despite the relatively lower labor costs prevailing in some countries. In the services domain, particularly with the development of today's information technology, everything is happening much quicker than before. Unlike the case of the goods industries, finished products (i.e., services) may be sent easily and with 'reckless speed' from country to country. And while there is less of a need for built infrastructure there is a need for educated and skilled workers. Thus, for example, a well-trained computer operator from China or India, with the most important capital good—computer—beside him/her, may be sitting at home and working jointly with his/her far away colleague whenever time permits. Once finished, the program is sent to the employer over the internet.

Some services jobs/occupations simply may be dying out. For the United States it has been noted that some services jobs/occupations simply may be dying out due to the new opportunities for cheapening the process of production brought about in the "computer age". For example, "… bank tellers have been replaced by automatic teller machines; receptionists and operators have been replaced by voice mail and automated call menus; back-office record-keeping and other clerical jobs have been replaced by computers; layers of middle management have been replaced by better internal communications systems."[3] Putting it even more bluntly, "...jobs are not simply being transferred overseas; they are being consigned to oblivion by automation and the resulting reorganization of work processes".[4]

Services, goods sector jobs and outsourcing. Finally, it may be that outsourcing has become an issue related to both goods and services. In the "old fashioned way", we may consider only services to be obtained by a U.S. company from some other countries. In recent literature, jobs of computer programmers, telephone operators, and bank tellers, etc., were termed to be "dirty jobs which basically have the aim of reducing costs."[5] "However, in the near future, it may appear that "outsourcing" more and more relates to lost jobs in the manufacturing sector. According to Business Week, companies like Dell, Motorola as well as Phillips, for example, are buying complete designs (which practically mean finished products) from the Asian developers, "tweaking them to their own specifications, and slapping on their own brand names."[6]

Government unemployment-financial aid programs, the (temporary) solution? Ultimately, if the jobs of blue-collar and white-collar employees are threatened due to the impact of foreign trade, the question is whether one needs to deal with this problem by perhaps developing like never before government unemployment-financial aid programs.[7] Before this, the issue may be that of assessing the availability of solid goods and services trade data to see how the current economic issues actually relate to trade data availability.

Table 1. Share of Exports of Goods by Industrial Sector, 1983, 1993, and 2000-03

Percent Distribution	1983	1993	2000	2001	2002	2003
Total exports of goods	100.0	100.0	100.0	100.0	100.0	100.0
1. Agriculture, forestry, fishing and hunting	14.4	6.6	4.1	4.5	4.8	5.4
2. Mining, utilities, and construction	3.4	1.3	0.9	1.0	0.9	1.1
3. Food, beverage, and tobacco product	5.7	5.6	4.3	4.7	4.6	4.7
4. Textile product and apparel	1.7	3.0	2.9	2.8	2.8	2.7
5. Wood product, paper, and printing	3.6	4.2	3.5	3.4	3.4	3.4
6. Petroleum, chemical, plastic, & rubber product	14.8	14.7	15.9	16.5	17.3	18.6
7. Metal and machinery	18.0	18.0	17.3	16.7	16.3	15.8
8. Computer, communications, and audio and video equipment	6.8	8.5	9.6	8.6	7.3	6.5
9. Semiconductor, navigational, electromedical, and appliance	11.6	15.6	20.2	18.8	18.4	18.5
10. Transportation equipment and furniture	15.4	18.2	17.3	18.7	19.9	19.0
11. Others	4.8	4.2	4.0	4.3	4.1	4.4

Source: TPIS

Table 2. Share of Imports of Goods by Industrial Sector, 1983, 1993, and 2000-03

Percent Distribution	1983	1993	2000	2001	2002	2003
Total imports of goods	100.0	100.0	100.0	100.0	100.0	100.0
1. Agriculture, forestry, fishing and hunting	4.2	2.9	2.1	2.1	2.2	2.2
2. Mining, utilities, and construction	18.1	8.3	7.1	7.2	6.9	8.8
3. Food, beverage, and tobacco product	4.3	2.9	2.3	2.6	2.8	2.9
4. Textile product and apparel	7.3	9.9	8.4	8.9	8.9	8.7
5. Wood product, paper, and printing	4.3	3.9	3.3	3.4	3.4	3.3
6. Petroleum, chemical, plastic, & rubber product	12.3	9.7	12.7	13.3	13.5	14.5
7. Metal and machinery	15.3	14.0	12.7	12.1	11.6	11.4
8. Computer, communications, and audio and video equipment	5.9	11.2	11.0	10.3	10.9	10.4
9. Semiconductor, navigational, electromedical, and appliance	10.5	16.7	20.0	18.3	17.8	17.1
10. Transportation equipment and furniture	16.2	18.7	18.2	19.3	19.6	18.3
11. Others	1.6	1.8	2.3	2.3	2.3	2.4

Source: TPIS

Table 3. U.S. Exports of Goods, Annual Rate of Growth

1997-2003	Annual Growth Rates of U.S. Exports to:				
	All Countries	India	EU15	Canada	China
1. Agriculture, forestry, fishing and hunting	0.4	24.6	-4.0	7.3	22.2
4. Textile product and apparel	-1.8	-3.3	-6.0	-1.5	17.0
8. Computer, communications, and audio and video equipment	-6.2	11.4	-5.5	-7.3	11.2
9. Semiconductor, navigational, electromedical, appliances	1.1	6.7	2.5	2.6	20.5
1992-2003					
1. Agriculture, forestry, fishing and hunting	1.4	16.0	-1.0	7.2	18.0
4. Textile product and apparel	3.3	7.2	-1.3	3.0	14.6
8. Computer, communications, and audio and video equipment	1.9	17.5	0.2	1.1	13.2
9. Semiconductor, navigational, electromedical, appliances	6.2	11.7	4.5	6.3	20.0

Source: TPIS

Table 4. U.S. Imports of Goods, Annual Rate of Growth

1997-2003	Annual Growth Rates of U.S. Imports to:				
	All Countries	India	EU15	Canada	China
1. Agriculture, forestry, fishing and hunting	2.2	8.4	3.4	13.2	16.8
4. Textile product and apparel	5.3	7.7	0.8	-2.2	7.9
8. Computer, communications, and audio and video equipment	6.3	-25.0	3.6	-6.8	26.6
9. Semiconductor, navigational, electromedical, appliances	4.2	12.0	6.9	3.0	13.8
1992-2003					
1. Agriculture, forestry, fishing and hunting	4.8	10.8	4.8	12.1	8.5
4. Textile product and apparel	6.8	9.2	3.8	5.8	9.6
8. Computer, communications, and audio and video equipment	8.0	4.9	5.3	3.6	28.7
9. Semiconductor, navigational, electromedical, appliances	8.7	13.1	8.4	6.5	17.9

Source: TPIS

Table 5. U.S. Exports of Services, 1992, 1997, and 2003

	Exports of Services	Services Components as a Share of Total Services Exports			Average Annual Rate of Growth	
		1992	1997	2003	1997-2003	1992-2003
	Total	100.0	100.0	100.0	3.0	5.0
1	Transfers under U.S. military agency contracts	6.4	6.0	3.7	-4.9	0.0
2	Travel	29.2	27.4	20.2	-2.1	1.5
3	Passenger fares	8.9	7.8	4.9	-4.7	-0.5
4	Other transportation	11.5	10.1	9.9	2.8	3.6
5	Royalties and license fees	11.1	12.4	15.1	6.4	7.9
6	Other private services	27.0	31.6	41.9	8.0	9.3
7	Other--part of which must be U.S. government miscellaneous services	6.0	4.7	4.3	1.4	1.8

Source: Survey of Current Business

Table 6. U.S. Imports of Services, 1992, 1997, and 2003

	Imports of Services	Services Components as a Share of Total Services Imports			Average Annual Import Growth	
		1992	1997	2003	1997-2003	1992-2003
	Total	100.0	100.0	100.0	7.3	7.1
1	Transfers under U.S. military agency contracts	11.2	6.8	9.6	13.6	5.6
2	Travel	31.2	30.4	21.6	1.4	3.5
3	Passenger fares	8.6	10.6	8.0	2.5	6.4
4	Other transportation	19.3	16.9	17.1	7.5	5.9
5	Royalties and license fees	4.2	5.4	7.6	13.8	13.0
6	Other private services	21.1	26.0	32.9	11.6	11.5
7	Other--part of which must be U.S. government miscellaneous services	4.5	4.0	3.2	3.5	3.9

Source: Survey of Current Business

Table 7. Total Receipts, 1992, 1997, and 2003

Other Private Services	Exports, i.e., Receipts (Millions of Current Dollars)				Rates of Growth	
	1992	1997	2003		1992-2003	1997-2003
Total affiliated and unaffiliated	50,600	84,500	13,4000		9.3	8.0
Total unaffiliated	33,467	57,005	85,368		8.9	7.0
Unaffiliated: India	539	646	2,030		12.8	21.0
EU15	9,061	17,738	30,038		11.5	9.2
Canada	2,596	3,916	6,887		9.3	9.9
China	784	1,383	2,439		10.9	9.9

Source: NIPA

Table 8. Total Payments, 1992, 1997, and 2003

Other Private Services	Imports, i.e., Payments (Millions of Current Dollars)				Rates of Growth	
	1992	1997	2003		1992-2003	1997-2003
Total affiliated and unaffiliated	26,100	44,600	86,300		11.5	11.6
Total unaffiliated	15,625	25,942	50,332		11.2	11.7
Unaffiliated: India	108	141	877		21.0	35.6
EU15	6,120	10,014	22,750		12.7	14.7
Canada	1,294	2,347	4,030		10.9	9.4
China	107	394	237		7.5	-8.1

Source: NIPA

Footnotes

[1] Maria Borga and Michael Mann. "U.S. International Services," Survey of Current Business, October 2004.

[2] The selected countries chosen were: Canada, Mexico, EU15 countries (including: Belgium, France, Germany, Italy, Luxembourg, the Netherlands, Denmark, Ireland, the United Kingdom, Greece, Portugal, Spain, Austria, Finland, and Sweden), Japan, China, Korea, Taiwan, Singapore, Malaysia, Philippines, Thailand, India, Brazil and Argentina.

[3] Brink Lindsey. "Job Losses and Trade: A Reality Check," Cato Institute.

[4] Ibid.

[5] Dean Foust, Michael Eidam, Spencer E. Ante, and Manjeet Kripalani in Bombay. "Outsourcing Food Chain," Business Week, March 11, 2004.

[6] Pete Engardio and Bruce Einhorn. "Outsourcing Innovation," Business Week, March 21, 2005.

[7] For further discussion, see Robert W. Bednarzik, "Restructuring in the U.S. Information Technology (IT) Sector: Should We Be Worrying About Offshoring?" Georgetown Public Policy Institute, Georgetown University.

This page intentionally left blank.

Forecasting Strategic Issues Facing the Veterans Health Administration I

Session Chair: Robert E. Klein, Office of the VA Actuary, Department of Veterans Affairs (robert.klein@va.gov)

Forecasting the COLA

Stephen Meskin, Chief Actuary, Department of Veterans Affairs (stephen.meskin@va.gov)

Benefits for disabled veterans increase each calendar year in accordance with Social Security COLA. The COLA for a given year is calculated in mid-October of the prior year when the September CPI-W is released by the Census Bureau. It is desired to develop a forecast of the COLA in the first week of October in order to meet financial reporting requirements. Time series models are discussed which are able to forecast the COLA within a margin of 0.1 over 90% of the time.

Time Series VS. Chain Ratio Forecasting

Rick Bjorklund, Carl Newman, Don Stockford, Laura Bowman, VHA Office of the Assistant Deputy Under Secretary for Health for Policy and Planning, and Rick Andrews, Louisiana State University

This paper describes methods and results of a forecasting competition held by the VHA to validate an existing forecasting model. Forecasting competitions typically involve competing statistical methods, models and experts. But how would powerful statistical forecasting methods stand up to forecasts generated based on simple ratios that relied on a deep and broad understanding of the markets and the business? To our knowledge this type of forecasting competition has never been held before. Both opponents performed at extremely high levels and when the competition was over both competitors had produced forecasted results that came within 1% of actual results. The competition and implications are discussed.

VA Health Care and U.S. and Global Aging

Donald Stockford, Veterans Health Administration, U.S. Department of Veterans Affairs

In this paper, the aging total veteran and VA health care enrollee populations are placed within the context of U.S. and global aging phenomena. For example, the "aging boomer" phenomenon includes veterans and veterans enrolled for VA health care. "Boomers", born between 1946 and 1964, begin to turn age 65 and become Medicare age-eligible in 2011. "Boomer generation" veteran and Vietnam Era veteran groups largely overlap. A projected new peak in the age 65 or over veteran population in 2013 is largely due to the aging of boomer/Vietnam era veterans. There are consequences for VA enrollment and health care.

Proposed Forecasting Methodology for Pharmacy Residency Training

Dilpreet K. Singh, Lori Golterman, Gloria J. Holland, Linda D. Johnson, Evert M. Melander, and Karen M. Sanders

The Veterans Health Administration (VHA) is the largest single provider of health care professions training in the United States. A significant percent of all pharmacy residents trained in the United States receive part of their training at VA. A quantitative performance measure has been established to serve as an indicator of the quality of the pharmacy training program at each VA facility. This paper includes a proposed criteria based forecasting methodology to determine future allocations of pharmacy trainee positions. This methodology will include factors, such as results of the Learners' Perceptions Survey, accreditation status, workload, etc.

This page intentionally left blank.

Time Series VS. Chain Ratio Forecasting

Rick Bjorklund, Carl Newman, Don Stockford, Laura Bowman, VHA, Office of the Assistant Deputy Under Secretary for Health Policy and Planning (ADUSH) and Rick Andrews, Louisiana State University

Introduction

Background: Forecasting competitions typically pit powerful statistical models and those at the leading edges of applying and developing them against one another. In practice, forecasting sales, units, budgets and other measures are often based on management's intuitive metrics that reflect an intimate understanding of markets, businesses, and their drivers.

Determining how to acquire forecasted information requires consideration of several factors associated with the cost of information. First and foremost, there is likely a cost associated with inaccurate forecasts. An internal forecasting staff increases fixed costs, overhead, and the breakeven point, all of which reduce profit in a for-profit organization. If those fixed costs can be shifted to a contractor they become a variable cost; reducing overhead and the breakeven point, thus potentially producing an economic benefit for the organization.

The Veterans Health Administration's (VHA) Office of the Assistant Deputy Under Secretary for Health (ADUSH) Policy and Planning contracts with an organization that has a deep and broad understanding of VHA's markets and business to develop future budget requirements. The contractor applies VHA's health care management metrics in the form of simple, intuitive ratios that are linked together creating "chain ratios" to forecast VHA enrollment, unique patients, and cost per patient. When multiplied together these "chain ratios" produce estimates of VHA's future budgets.

The described situation sets the stage for three interesting questions suggesting a unique forecasting competition. (1) Can forecasts made by a widely used, sophisticated statistical forecasting software package (i.e., ForecastPro) improve upon forecasts produced by management's intuitive metric (as operationalized by contractor: Milliman, Inc.)? (2) Could a "World Class" forecasting expert construct models that would improve upon management's intuitive metrics and/or those produced by ForecastPro? (3) What insights into the cost of information and organizational implications were gained from this competition?

The authors believe this is a unique forecasting competition since to our knowledge, for the first time, sophisticated statistical models are pitted against intuitive management metrics in a forecasting competition.

Results: Results of the forecasting competition appear in Exhibit 1. Using a rule of thumb that forecasts containing less than 10% error are considered "good", all participants in the forecasting competition produced exceptional forecasts (small deviation from actual results). ForecastPro produced the best enrollment forecast, with expert forecasts next but containing twice as much error. Milliman produced the best unique patient forecast followed closely by those of the expert, only about .5% behind.

Methods

Participants: There were three participants in this forecasting competition:

(1) Milliman, Inc, the VHA contractor, has been servicing the health care and insurance industries for decades and, since 1999, providing annual budget forecasts based on enrollment and unique patient projections to VHA. Milliman does not provide monthly forecasts.

(2) Professor Rick Andrews, Robert S. Greer, Sr. Alumni-Endowed Chair of Business Administration and Professor Department of Marketing, Louisiana State University, Baton Rouge, LA. Dr. Andrews has published leading articles on times series forecasting that are widely cited in the forecasting literature and has participated in a number of statistical forecasting competitions. Dr. Andrews provided custom models to forecast monthly and annual VHA enrollment and unique patients.

(3) ForecastPro Software: VHA staff experienced with statistical forecasting techniques and having experience reviewing alternative software products purchased ForcastPro from Scientific Systems, Cambridge, MA. ForecastPro software has the

capability of considering most time series (e.g., State Space components, Box-Jenkins, Exponential Smoothing, Moving Average, etc.), econometric (simple and step-wise) and combinations of time series and econometric techniques (e.g., using time series techniques to model the error term from econometric models). Forecast Pro offers the option of several modes of model selection. An "expert selection" mode selects the best (smallest forecast error) technique for fitting a set of data and forecasting future outcomes. VHA staff attended a training course on the software and underlying statistical techniques included in the software. ForecastPro in "expert selection" mode provided monthly and annual forecasts of VHA enrollment and unique patients.

Setting: The competition between statistical models and management metrics began in June '04, using monthly data from the period Jan '01 to Sep '03. The competition involved forecasting annual enrollment and unique patients for two fiscal years, FY '03 and FY '04. FY '03 was considered the *experimental phase* since actual year end data was known. FY '04 was considered the *test phase* since, at the time of the competition, actual results were not known.

Statistical models included both econometric and time series. Econometric techniques required independent variables that explained variation in the dependent variables (i.e., enrollment and unique patients). Macroeconomic variables available on a monthly basis such as real income, relative prices, (either proxies for, or actual data), etc were used as independent variables in the econometric models. Time series models relied on monthly enrollment and unique patient series over the Jan '01 – Sep '03 period.

Data: The VHA Enrollment and Unique Patient series are displayed in Exhibit 2. Enrollment shows a trend with no apparent seasonality, while Unique Patients shows a trend with strong seasonality.

The unique patient series is a count of unique patients seen by VHA physicians throughout the fiscal year. The count is cumulative, beginning with patients remaining in VHA facilities at the Fiscal Year start and increasing as more patients are seen throughout the year.

The enrollment series does not appear to have seasonal or cyclical trends and might be estimated by a simple percent increase or linear regression model with dummy variables to account for policy changes which occasionally drive small, but distinct increases and decreases. The cumulative nature of both series suggested that differencing the data may improve results; hence forecasts were made with both the natural and differenced data with the objective of selecting a data format that minimized forecast error.

Exhibit 3 shows the same data using a holdout sample and comparing "forecasts" to actual data.

Since the competition pitted statistical models against management metrics as operationalized by Milliman, error rates (the difference between known actual and forecasted results) were calculated from FY '03 forecasts produced by the management metrics. These error rates would eventually be compared to those produced by the statistical models. At the time of the forecasting competition, June '04, only the actual results for FY '03 were available. The actual results for FY '04 would not be available until approximately January '05.

For econometric techniques, data on monthly macro-economic independent variables over the period Jan 2001 through Sep 2003 were located from public data provided by the Bureau of Labor Statistics and the Bureau of Economic Analysis consistent with two criteria; they (1) were thought to be related to the number of enrollees or the number of unique patients in a macro-economic sense (i.e., including independent variables such as real income, relative cost of health care, etc) and (2) were or could be independently forecasted into the future. Broad categories of independent variables which met this criteria were:

~ National Unemployment
~ Civilian Benefits
~ Veteran Unemployment
~ Personal Income and Expenditures
~ Medical Consumer Price Index
~ VHA Healthcare Expenditures

From the set of independent variables, ForecastPro identified the six that were statistically significant for the purpose of forecasting Veteran Enrollment and Unique Patient population. Most were sub-categories of Personal Income and Expenditures: Income (Current transfer receipts from business); Personal Savings as a percentage of personal disposable income; non-farm income; and disposable personal income per capita (2000) dollars. In addition, VHA medical expenditures and Enrollment (previous observations) were considered statistically significant.

For the time series techniques, enrollment and unique patient data were obtained for monthly time periods from Jan 2001 through Sep 2003. For this set of data, it was our experience that, in general, econometric techniques were not able to compete with time series techniques for forecasting enrollment and unique patients and so our discussion of them will stop here.

In the experimental phase the actual enrollment and unique patient data were divided into two periods: model estimation (Jan 2001 – Sep 2002) and hold out samples (Oct 2002 – Sep 2003). Enrollment and Unique Patient monthly data were submitted to ForecastPro; the model best fitting the data was selected and used to forecast enrollment and unique patients for the period Oct '02 – Sep '03.

The same data were provided to the forecasting expert and he was asked to divide the data into the same two periods and develop a statistical time series model to forecast the hold out sample. The expert used a time series model called the Basic Structural Model:

$$y_t = \mu_t + \gamma_t + \varepsilon_t$$

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t$$

$$\beta_t = \beta_{t-1} + \varsigma_t$$

$$\gamma_t = -\sum_{j=1}^{s-1} \gamma_{t-j} + \omega_t$$

Where:

y_t is the series to be forecasted;

μ_t is the level or mean of the series, which varies over time;

β_t is the slope of the trend of the series, which varies over time;

γ_t is the seasonal component of the series, which varies over time;

$\varepsilon_t, \eta_t, \varsigma_t, \omega_t$ are random noise components, whose variances are determined from the data.

The series is thought to be composed of a level component (which may trend over time), a seasonal component, and a random noise component. The slope of the trend may also change over time. The seasonal component is constructed such that the seasonal factors sum to zero over any given year.

Unique patient forecasts were made for both the raw and differenced Utilization series and the forecasts for the differenced series was preferred. This made sense due to the cumulative nature of the Utilization series.

Even though Enrollment has little, if any, seasonality a Monthly Basic Structural Model which includes seasonality was used because (1) there may be some unseen seasonality and (2) it does no harm to include a seasonal component when there is really no seasonal pattern in the data.

Experimental Phase: In the experimental phase both ForecastPro and the world class forecasting expert became familiar with the idiosyncrasies of the enrollment and unique patient data series. The experimental phase consisted of a total of 33 months of data (January 2001 through September 2003). As noted above, the data series were divided into model estimation (January 2001 – September, 2002 or 21 months) and forecast periods (October 2002 – September 2003 or 12 months). That is, the parameters of alternative models were estimated using the January 2001 – September 2002 data and the model with the least error selected. This was used to forecast monthly enrollment and unique patients over the October 2002 – September 2003 period.

Policy changes that impacted the enrollment series in the first few months of the forecast period for FY '03 were calculated and applied to the annual enrollment forecast. Generally, given enough observations, time series methods are sufficiently sensitive to pick up subtle changes in trends. However, in this situation, the policy change (i.e. halting enrollment of veterans with Priority 8 status defined as veterans both with income above VA and geographic means thresholds and who do not have service-connected disabilities, catastrophic disabilities, or conditions related to chemical exposures during service) occurred during the forecast period. To account for its impact, the impact on overall enrollment growth was calculated and subtracted from the annual totals forecasted by ForecastPro and the world class forecasting expert. Milliman had accounted for the policy change in their enrollment forecast, although their method is unknown.

Results of the experimental phase are in Exhibit 4 and the fitted/projected data are in Exhibit 5.

The results from the experimental phase by all participants were very good. The purpose of the experimental phase was to gain familiarity with the idiosyncrasies of the data series and clearly the expert's forecasts set the standard for both enrollment and unique patients. For enrollment his results were about three times more accurate than Milliman and about nine times more accurate than ForecastPro. For unique patients, the expert's forecast was about three times more accurate than either ForecastPro or Milliman.

Test Phase: In the test phase, the forecast for FY '04, statistical models used the full set of available data (Jan '01 – Sep '03) to estimate model parameters and select the best model. The test phase was also distinguished from experimental phase on another dimension: Both ForecastPro and the world class forecasting expert had no information on Milliman error rates since actual results from FY '04 were unavailable. In the experimental phase, there was as much focus on developing models and producing forecasts as there was on improving the error rate experienced by Milliman. In the test phase there were no target error rates to improve upon, therefore, the only focus was on providing the best forecast.

The final results (summarized) from the test phase are presented in Exhibit 1. The full results, including monthly forecasts provided by the statistical models, are presented in Exhibits 6 and 7.

At the time the forecasts were made there was confidence and expectation that the statistical models would improve upon forecasts using management metrics. While expectations were met for enrollment, it appears as though, for unique patients, the last month of FY '04 contained some surprises that were not sensed by the statistical models.

Discussion

At the outset of this paper we indicated that forecast accuracy and the cost of that information were issues in selecting a forecasting method. To assess accuracy of Milliman's forecast of VHA unique patients and enrollment, both powerful statistical techniques and a world class forecasting expert tried to improve upon the Milliman forecasts. Milliman's forecasts are based on simple management metrics developed from a deep and broad understanding of the industry, markets, and business. We have characterized this as a chain ratio approach. For each of VHA's 157 facilities, the Milliman metrics forecast demand and associated cost of providing healthcare services to meet that demand for 56 groups of health care services. This information is aggregated to facility, regional, and national levels to produce estimates of demand and the cost of providing services to meet demand.

The most successful statistical techniques employed were time series techniques. This is a broad categorization of statistical techniques ranging from, for example, simple moving average, to exponential smoothing, to Box-Jenkins and State-Space. Depending on the data series, there may be a large number of different models that fit the data. The challenge is to select the model with the greatest likelihood of success when forecasting into the future. The statistical model exhibiting the least error for a given time series was selected.

The differences between forecasting accuracy of the statistical models and management metrics were negligible. However, what is the cost of this information? What is the cost associated with forecasting error? Answers to either or both of these questions might provide an obvious direction in terms of which method to select.

In this situation the cost of information is difficult to quantify. VHA's contract with Milliman includes not only basic forecasting services discussed in this paper, but analytical services to address questions raised by Congress, OMB, and senior VHA management. The knowledge and understanding to accomplish both tasks are interdependent and continually leverage each other to provide ever increasingly accurate forecasts and value added products.

However, an assessment of how information is used and the consequences of misinformation can be made.

- Use of Information: Forecasts of the demand for VHA health care have to be viewed in context of all demands on the Federal Budget. VHA obtains funding from Congress and the level of funding is often associated with the political and economic climate, as well as tradeoffs Congress must make between other budget demands and VHA's budget request. Political and economic situations may be the most important driver determining the budget received by VHA. However, the use of an independent actuary to forecast enrollment and utilization has gained credibility in the eyes of OMB and Congress. VHA has found this independent assessment to be a powerful resource in defending and protecting its budget requests.

Related costs that might be useful to consider include:

(1) Potential costs associated with the forecasting function and the accuracy of the forecast

- <u>Organizational cost</u> to VHA of building and maintaining an in-house forecasting capability with credibility equal to that of current contactor

(2) Potential costs associated with a forecast of demand that falls below actual demand:

- <u>Patient Incurred Cost</u>: In the VHA patients with veteran status are not turned away, and so there is no patient incurred cost if the forecasted level of demand is lower than actual demand.
- <u>Organizational cost to implement "Cost take-out" strategy</u>: The organization's ability to continue to reduce cost to close the gap between funds received from Congress and the cost of providing health care is not known.

(3) Potential cost associated with forecasts of demand that exceed actual demand:

- <u>Interest Cost</u>: In periods where the government is running a deficit, the cost to the US Treasury of issuing more debt than was actually required.
- <u>Opportunity Cost</u>: Cost associated with not providing government services where they were needed as a result of providing them to VHA where they were not needed.

(4) Potential opportunity cost associated with the value of monthly forecasted information.

- <u>Opportunity Cost</u>: One might assert that because the time series techniques provided monthly enrollment and unique patient information that they would intrinsically be more useful to management, since mid year policy changes might be implemented if the environment was changing quickly. While the Milliman model does not provide monthly forecasts, the ADUSH office develops and provides to senior management monthly forecasts based on historic monthly changes and Milliman end of year forecasts. These roughly linear interpolations are robust estimates of actual monthly enrollment and unique patient results.

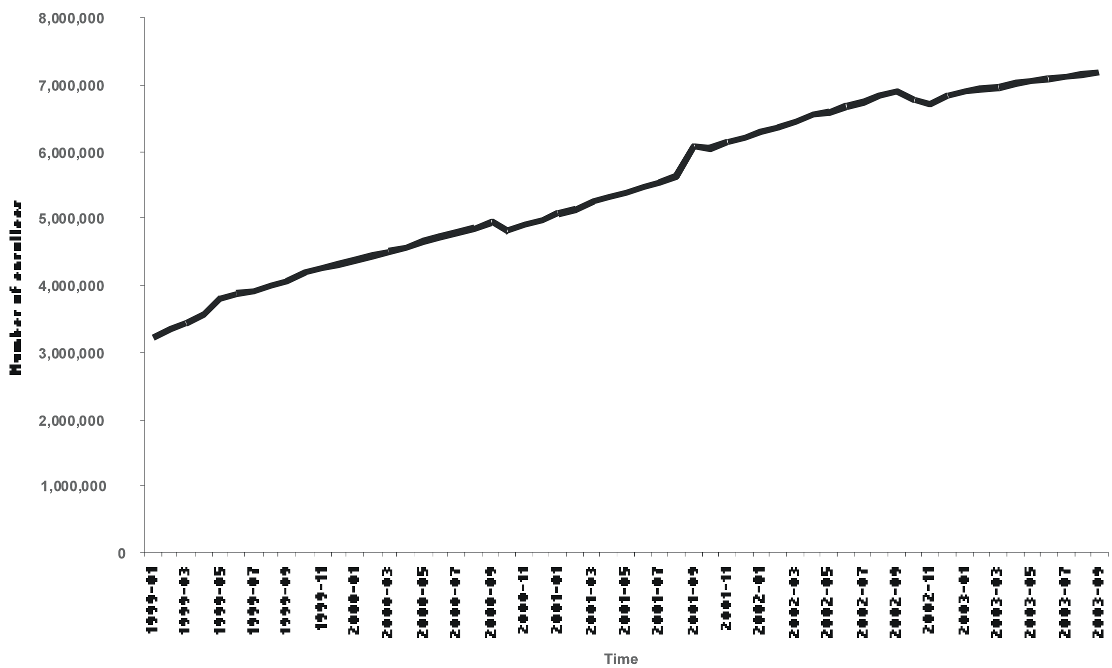
At this time, the Milliman model seems the most efficacious to meet all of VHA's needs. The error rate is either better or close to statistical models. The forecasting work is incorporated into larger analytical tasks which the statistical models cannot accomplish, therefore there is minimal contracting expense saved in bringing this activity in-house. And, at this time, VHA does not have the resources to build the in-house expertise needed to duplicate Milliman's resources. Conversely, the statistical models helped to validate Milliman's forecasts, as well as proved to be good independent forecasting tools that may be used to pull quick, internal forecasts for planning purposes.

Exhibit 1
Forecasting Results
September 2004

Participants	Enrollment			Unique Patients		
	Actual	**Forecast**	**Error**	**Actual**	**Forecast**	**Error**
Milliman (Chain Ratio)	7,419,852	7,632,416	2.9%	4,657,997	4,701,689	**.94%**
ForecastPro (Statistical)	7,419,852	7,323,246	**1.3%**	4,657,997	4,843,207	4.0%
"Expert" (Statistical)	7,419,852	7,235,220	2.6%	4,657,997	4,725,132	1.4%

Exhibit 2 (time series)
January 1999 thru September 2003

<u>Enrollment</u>



Unique Patients

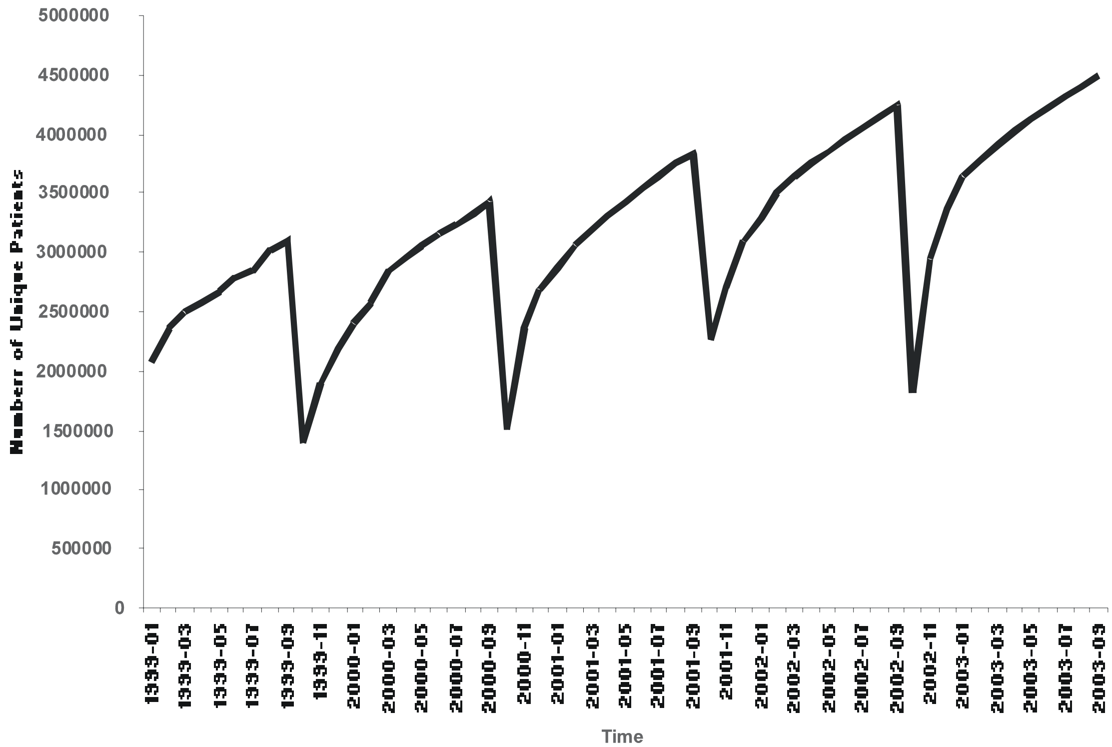
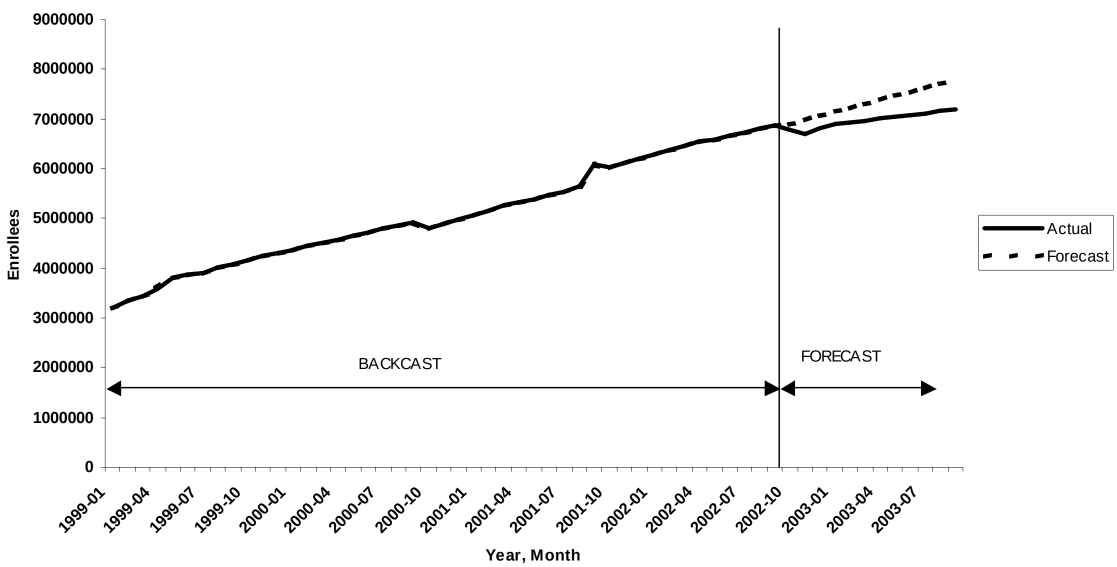


Exhibit 3
Hold Out Sample: *October 2002 thru September 2003*

Enrollment



Unique Patients

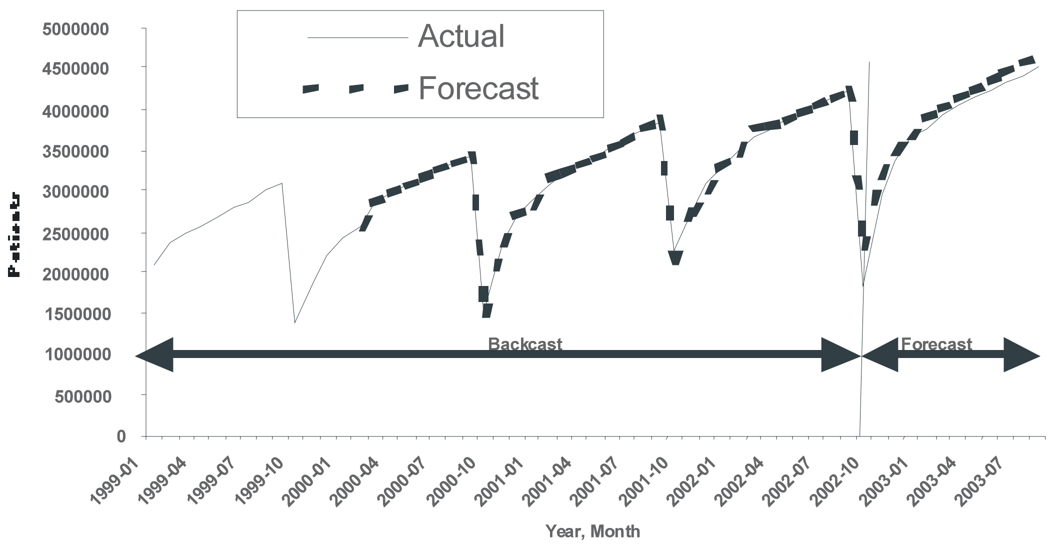
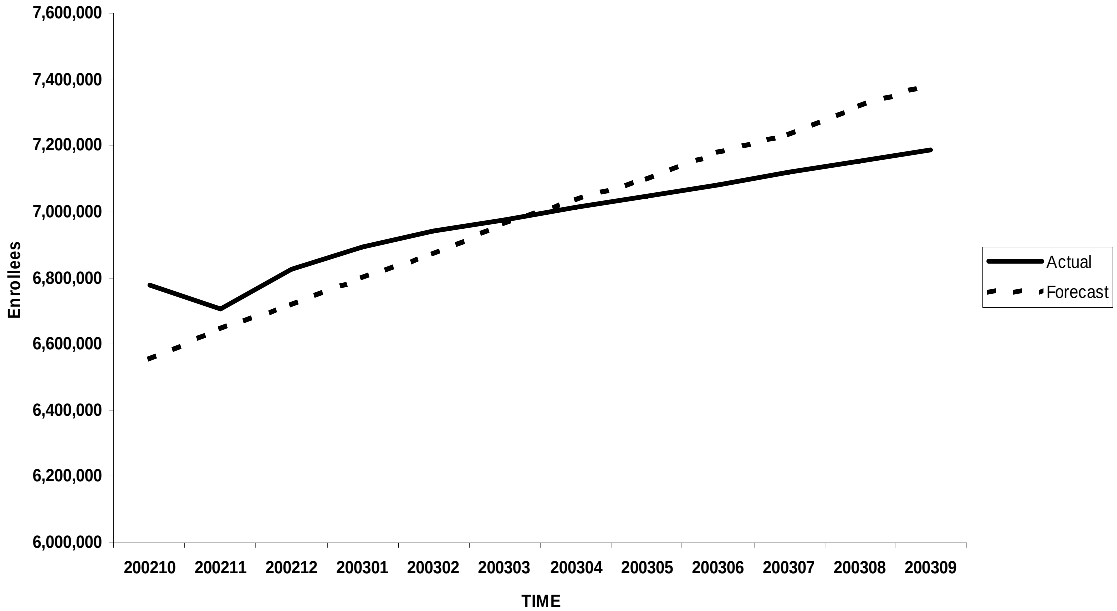


Exhibit 4
Forecasting Results
Experimental Phase
September 2003

Participants	Enrollment			Unique Patients		
	Actual	Forecast	Error	Actual	Forecast	Error
Milliman (Chain Ratio)	7,186,823	7,350,999	2.28%	4,494,425	4,673,503	3.98%
ForecastPro (Statistical)	7,186,823	7,643,147	6.33%	4,494,425	4,843,207	3.63%
"Expert" (Statistical)	7,186,823	7,243,530	**.78%**	4,494,425	4,725,132	**1.21%**

Exhibit 5
Experimental Phase: "Expert" Fitted/Projected Results

Enrollment



Unique Patients

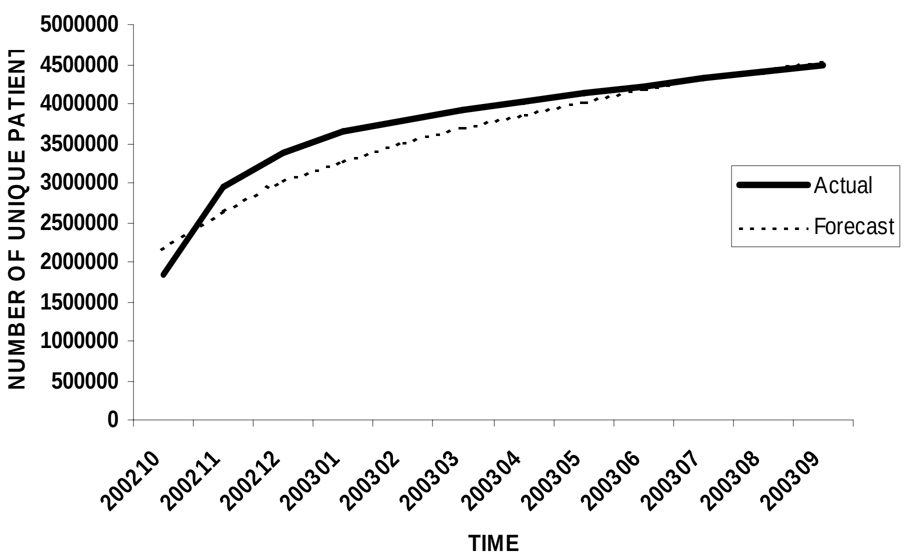


Exhibit 6
Cumulative Enrollment Forecast

Month FY'04	Actual	ForecastPro		Expert		Milliman	
		Forecast	Error	Forecast	Error	Forecast	Error
Oct	6,980,784	7,140,067	2.28%	7,006,821	.3%		
Nov	7,026,638	7,017,382	0.13%	7,028,162	.02%		
Dec	7,059,922	7,064,970	0.07%	7,060,772	.01%		
Jan	7,102,938	7,063,886	0.55%	7,099,683	.04%		
Feb	7,143,477	7,124,404	(0.26%)	7,116,626	(.04%)		
Mar	7,193,825	7,146,652	(0.66%)	7,133,568	(.08%)		
Apr	7,228,292	7,157,797	(0.99%)	7,150,510	(1.1%)		
May	7,266,837	7,186,834	(1.1%)	7,167,452	(1.4%)		
Jun	7,302,297	7,208,863	(1.3%)	7,184,394	(1.6%)		
Jul	7,341,564	7,211,270	(1.8%)	7,201,336	(1.9%)		
Aug	7,378,194	7,235,313	(1.9%)	7,218,278	(2.2%)		
Sep	7,419,852	7,323,246	(1.3%)	7,235,220	(2.6%)	7,632,416	2.9%

Exhibit 7
Test Phase
Unique Patient Forecast

Month FY'04	Actual	ForecastPro		Expert		Milliman	
		Forecast	Error	Forecast	Error	Forecast	Error
Oct	2,006,581	2,218,150	10.5%	2,113,792	5.3%		
Nov	3,162,588	3,130,125	(1.0%)	3,190,337	.9%		
Dec	3,467,040	3,552,963	2.4%	3,484,381	(.5%)		
Jan	3,784,683	3,769,280	.4%	3,765,662	(.5%)		
Feb	3,986,198	3,977,568	(0.2%)	3,872,595	(2.9%)		
Mar	4,146,894	4,152,578	0.1%	4,034,188	(2.8%)		
Apr	4,244,821	4,280,188	0.8%	4,150,784	(2.3%)		
May	4,329,908	4,394,395	1.5%	4,251,218	(1.8%)		
Jun	4,417,602	4,508,669	2.1%	4,349,428	(1.6%)		
Jul	4,499,171	4,612,245	2.5%	4,450,483	(1.1%)		
Aug	4,580,766	4,735,576	3.4%	4,568,330	(0.2%)		
Sep	4,657,997	4,843,207	4.0%	4,725,132	1.4%	4,701,689	.94%

VA Health Care and U.S. and Global Aging

Donald Stockford
Office of the Assistant Deputy Under Secretary for Health for Policy and Planning
U.S. Department of Veterans Affairs

Introduction

It's a remarkable time in the history of the world. The global population is at its highest level ever, although the pace of global population growth peaked some time ago and, everywhere all over the world, populations are aging – there are rapidly increasing numbers of elderly persons in virtually all developed and developing countries alike.

These simple facts have profound implications for the future of the world, the U.S., and the VA and VA health care system – yes, the VA and VA health care system. As such, it is timely and informative to look at how the U.S. is doing relative to the rest of the world and to seek some strategic insights into how global and U.S. population and aging phenomena are impacting VA and the future of VA health care.

An increasingly elderly U.S. population means U.S. health care must evolve to meet new challenges and not remain static. Of course, the VA is still the largest health care delivery system in the U.S. Half of all veterans enrolled for VA health care currently are age 65 or over. VA will continue to be concerned about elder veteran care and long-term care for veterans, in particular, even as it deals with increasing levels of younger Gulf War Era veterans and their very different and special needs, because VA is responsible for the health care of veterans over the lifespan.

Consider the Baby Boomer Generation, persons born between 1946 and 1964. Most Vietnam Era veterans are baby boomers, and VA must evolve to meet the challenge of boomer/Vietnam Era veterans starting to turn age 65 in the year 2011 and how their special needs will be different than those of the more elderly World War II or Korean War veterans who came before them or the much younger Gulf War Era veterans who come after.

Still, the baby boomer story is really in many ways just a blip in much longer and more profound global and U.S. population aging trends that are ongoing and that will continue well into the foreseeable future. And it is increasingly true, as Marshall McLuhan stated: "We now live in a global village"; and everything and everyone have the potential to impact everything and everyone else in ever new and fascinating ways. VA is situated in a critical place in the world story and what happens with VA can be turned into lessons for everybody everywhere.

The World's Oldest Man is a U.S. (& P.R.) Veteran

On January 17, 2005, Guinness World Records announced the world's oldest man and, in doing so, proclaimed America's (and the world's) oldest veteran. Guinness announced that Emiliano Mercado del Toro of Puerto Rico was the world's oldest man at 113 years and 149 days of age (born August 21, 1891). He entered the U.S. Army in 1918 and was in training in the final months of World War I. At this writing, he is still alive and still the world's and America's oldest man and oldest veteran.

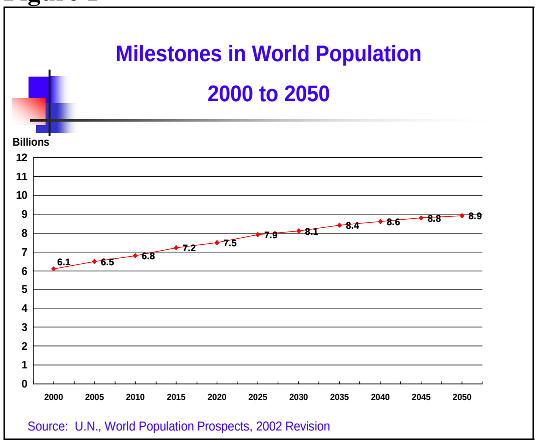
This story suggests that there might be an important link between age and veteran status. And, indeed there is. We discover, just for example, that in 2005 most elderly males in the U.S. are veterans of active duty military service - a little known but important fact. There are other related facts and issues discussed below.

Wars are often international events (World War I, World War II, etc.), and one begins to wonder whether there are no doubt similar
stories in other countries of the world. Although we do not include international veteran data in this paper, we nevertheless may have a clue now
when we think of aging in the world that the Baby Boom is just part of a much longer term aging story.

Milestones in World Population

The world population hit 6 billion in 1999, double what it was in 1960,[1] and it is currently (in 2005) at about 6.5 billion (**Figure 1**). The time it took to grow from 5 billion (in 1987) to 6 billion (in 1999) was only 12 years, the shortest period between any billion increase.[1] However, the pace of global population growth peaked in this time period (1987-1999),[1] and a reduced pace of global population growth has followed and will continue for some time to come (**Figure 1**). The declining pace of global population growth is due primarily to declines in fertility.[1]

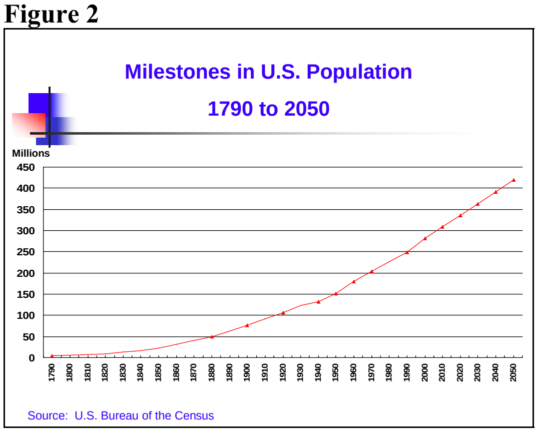
Figure 1



Milestones in U.S. Population

The U.S. population continues to grow at a rapid pace.[2, 3] From 1790 to 2000, the U.S. population grew by 71 times, from 3.9 million to 282.1 million; and from 2000 to 2050 it is projected to grow by 1.5 times again, from 282.1 million to 419.9 million (**Figure 2**). Although fertility rates in the U.S. have declined long-term and seem recently to have stabilized, increasing life expectancies and immigration are also key factors in overall U.S. population growth as is the age-sex structure of the population; there are relatively high and increasing numbers of younger people, including young adults and women of child-bearing age, but there are also increasing numbers of older people in the United States. There is an equalization of the generations taking place, whereby older persons are increasingly comprising a larger component of the total population.[4]

Figure 2



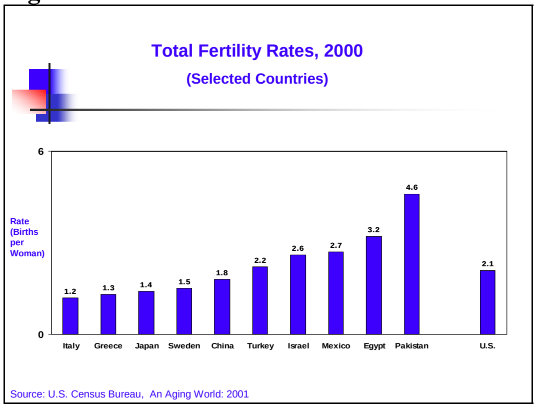
Fertility Rates

One of the major factors in long-term population trends is the "Fertility Rate". The fertility rate measures, on average, the number of children each woman would have in her lifetime, assuming that current age-specific birth rates remain constant throughout her child-bearing years. In a population with approximately equal levels of in and out migration, a sustained fertility rate of about 2.1 or slightly less can gradually lead to zero population growth (i.e., birth rate=death rate).

In the United States, the fertility rate has been declining since at least the year 1800, when women had on average 7.0 children.[4] By 1945, the U.S. fertility rate had declined to about 2.4. With the post-World War II Baby Boom (1946-1964), the U.S. fertility rate started to rise again but declined rapidly again afterwards, only to rise slightly and begin to level off from about the mid 1980's through the 1990's. This may be a consequence of the economic expansion during that time period. Nevertheless, the fertility rate in the U.S. is now at about 2.1, the level necessary to ensure zero population growth (assuming approximately equal levels of in and out migration).[5]

Worldwide persistent low fertility since the 1970's[4,6,7] has led to global declines in the size of successive birth cohorts and a corresponding increase in the proportion of older to younger population.[6] Fertility rates have been declining long-term throughout most of the world – and in many countries below the 2.1 replacement rate that would keep populations from declining over time.[4,5,6,7] In many of these countries, population continues to grow and largely due to the age structure of the population (relatively high numbers of young people even if lessening percentages of young people, as the generations equalize) and immigration.[4, 6, 7] As mentioned above, the U.S. itself appears to be leveling off at about a fertility rate of 2.1, a level last seen only in the 1970's[4, 7] (**Figure 3**).

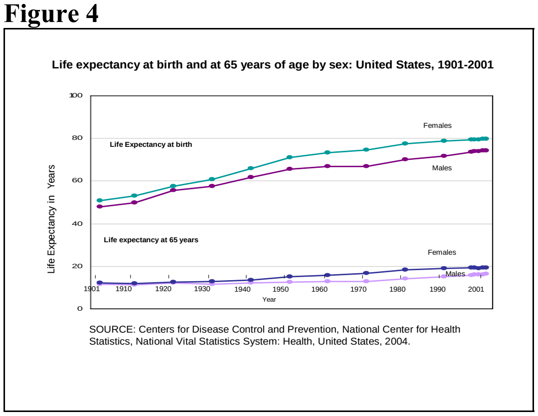
Figure 3



Increases in Life Expectancy

Life expectancy is a summary measure of mortality, representing the average number of years of life to be expected if mortality rates remained constant. In the U.S., life expectancy has been improving long-term - in fact, improving by about 30 years since 1900[8] (**Figure 4**). However, there are many countries with higher life expectancies than the U.S.[9] (**Figure 5**).

Figure 4



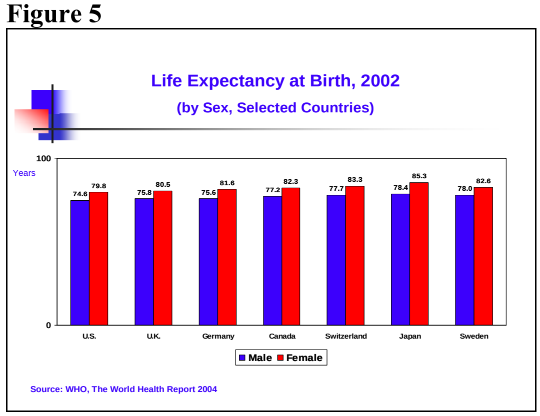
According to **Figure 4**, the overall trend in life expectancy for the U.S. was upward throughout the 20th Century. The gain for women exceeded that of men until the 1970's, primarily because of heart disease and lung cancer among men due to widespread adoption of cigarette smoking.[8] After the1970's, the gain for men exceeded that for women and the gender gap narrowed. From 1990 forward the trend reflected greater decreases in heart disease and cancer deaths among men and proportionately greater increases in respiratory disease related mortality among women.[8]

Improvements in life expectancy are shared by developed and developing countries alike. Improvements in life expectancy since about the mid-1800's are generally attributed to the interplay of many factors including innovations in medicine, sanitation, nutrition, as well as new modes of familial, social, economic, and political organization.[6]

Immigration

Immigration is another factor impacting U.S. population aging trends, but it is a lesser one than either fertility rate declines or improvements in mortality (increases in life expectancy). U.S. immigrants tend to be young adults with relatively high fertility rates, which might reduce the proportions of elderly in the U.S. A decline in immigration of mostly young adults after World War I contributed to population aging thereafter. On the other hand, the immigration rise that began after World War II will gradually slow but not stop long-term U.S. population aging trends, as long-term U.S. aging trends (long-term declines in fertility and long-term increases in life expectancy) are so well-established and strong.[4]

Figure 5



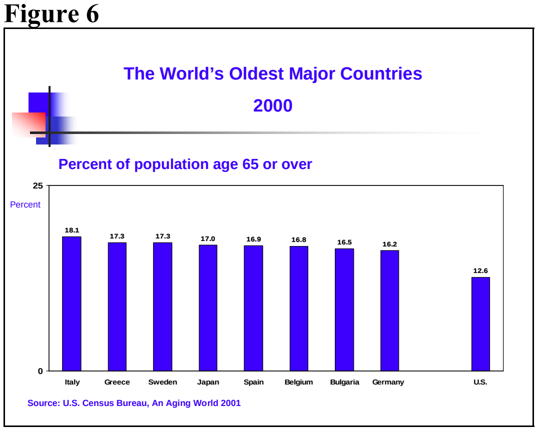
The World's Oldest Countries

Italy is now the world's oldest country, with about 18.1 percent of its population age 65 or older (**Figure 6**). However, the U.S. is still very young by developed world standards, with only about 12.6 percent of its population being age 65 or older. But when baby boomers begin to turn 65 in 2011, the percent elderly in the U.S. will rise markedly for some time to come.[6, 10]

U.S. and Global Aging and Long-Term Care

Global demographic trends and the health and long-term care needs of elderly persons are cross-national stories, perhaps requiring cross-national solutions. Italy, Greece, Sweden, Japan, Spain, etc., are the world's "oldest" nations (the U.S. is only 38th among the major countries)[11]; but all must grapple with issues of access, cost, and quality in health and long-term care, with potential for cross-national solutions (**Figure 6**). The problems are clearly not peculiar to but shared by the U.S.; and, in the U.S., the VA as the largest health care delivery system in the nation will be impacted.

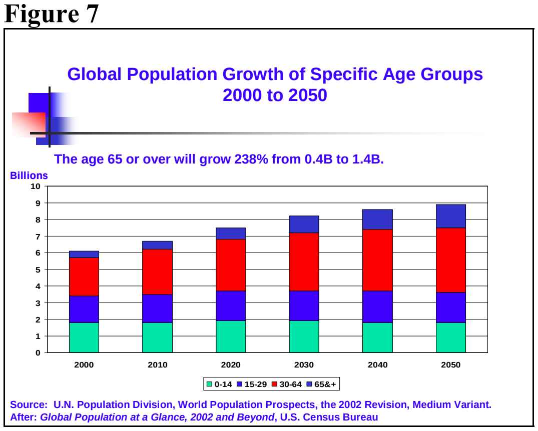
Figure 6



The 65 or Older; the 85 or Older

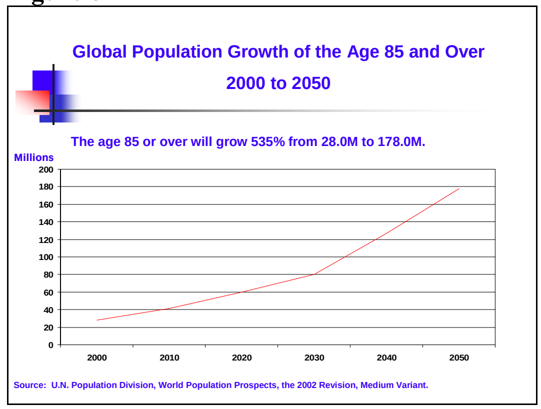
Globally, the age 65 or older will grow some 238%, from 0.4 billion in 2000 to 1.4 billion in 2050[1, 12] (**Figure 7**).

Figure 7



The age 85 or older will grow some 535%, from 28.0 million in 2000 to 178.0 million in 2050[1, 12] (**Figure 8**).

Figure 8

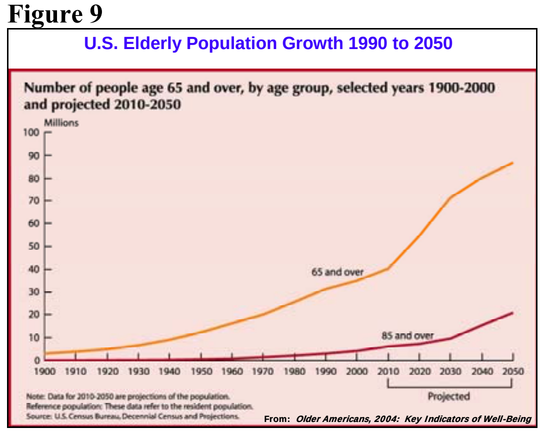


The age 85 or older is the population at risk for use of long-term care services including nursing home care and also community based home care. Increasing numbers of the "oldest old" is a cross-national phenomenon and, as we shall see, a particular issue for the U.S., the VA, and VA health care.

U.S. Elderly Population Growth, 1990-2050

In the U.S., the age 65 or older population grew from 3.1 million in 1900 to 35.0 million in 2000 and will grow to 86.7 million in 2050 (**Figure 9**). The "oldest old" (85 or older) grew from about 0.1 million in 1900 to 4.2 million in 2000 and will grow to 20.9 million in 2050.

Figure 9



Baby boomers (born between 1946 and 1964) begin turning age 65 in 2011, and U.S. elderly population growth rates will decline after 2030 when the last baby boomers join the ranks of the elderly. In 2030, about 20% of the U.S. population will be age 65 or older.[6, 10]

200 Year U.S. Aging Trend

The Baby Boom is not the reason for the aging of the U.S. population[4]. The U.S. has been growing older for the last 200 years. U.S. population aging is the result of long-term trends in increasing longevity (improving mortality rates), 1900-Present, and declining fertility, 1800-Present, reducing the relative numbers of young people and increasing the relative numbers of older people in the United States.[4]

Indeed, a permanent change is taking place in the American demographic profile. From 1880 to 2080 the relative numbers of older people increase and the relative numbers of younger people decrease, as the population is becoming more evenly distributed across the generations (equalization of the generations).[4] In this timeframe, the Baby Boom is just a relatively short-lived fertility boom within a very long-term downward trend in U.S. fertility rates. The Baby Boom does explain the rapidity of aging in the U.S. in the coming decades, as boomers only start to turn age 65 or older in 2011.[4, 10] The impact of baby boomers upon longer term U.S. aging trends declines from about 2030 on, but the U.S. population will continue to age well beyond that point, albeit a little more slowly.

Recall from above that the U.S. is still relatively young among the world's major countries, with only about 12.6 percent of the population age 65 or older in 2000. Baby boomers begin to turn 65 or older in 2011, and the percent elderly in the U.S. will increase dramatically to reach 20 percent by 2030.[4, 10]

Veteran Population

Related to these global and U.S. national aging trends, about two-thirds of all elderly males in the U.S. (**Figure 10**) are veterans and either eligible for or potentially eligible for VA health care.

War periods are key drivers of veteran population numbers. However, most Vietnam Era veterans are baby boomers and, to VA, the boomer phenomenon means that Vietnam Era veterans are aging. In particular, aging Vietnam Era veterans will cause a new peak in the age 65 or older veteran population in 2013. About one-third of all age 85 or males in the U.S. are veterans and there will be a new peak in the age 85 or older veteran population in 2012. Between 2000 and 2020, the total veteran population will decrease by 32%, or from 26.5 million to 18.1 million (**Figure 11**).

The number of veterans age 65 or older peaked in the year 2000 at 10.0 million, primarily due to the aging and mortality of the World War II cohort and Korean War veterans, but it will peak again in the year 2013 at 9.3 million, primarily due to the aging of the Vietnam Era cohort following behind World War II and Korean War veterans.

Figure 10

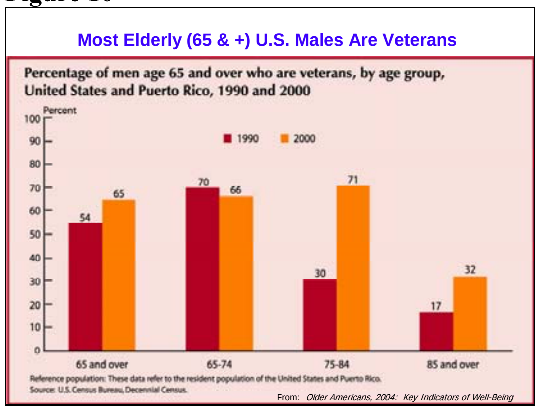
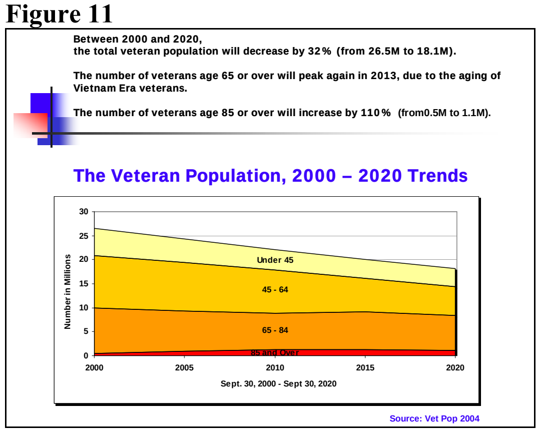


Figure 11

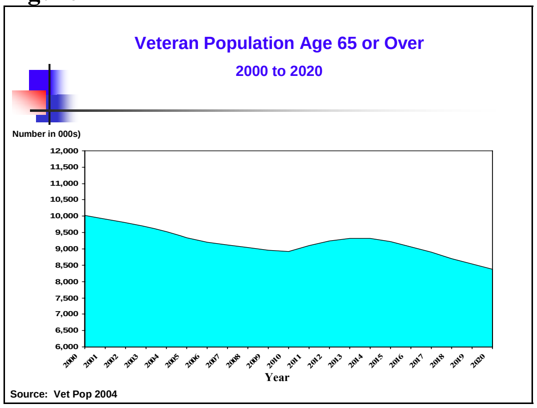


Veterans 65 or Older, or 85 or Older

From 2000 to 2020, the age 65 or older veteran population is projected to decrease by 16%, or from 10.0 million to 8.4 million (**Figure 12**).

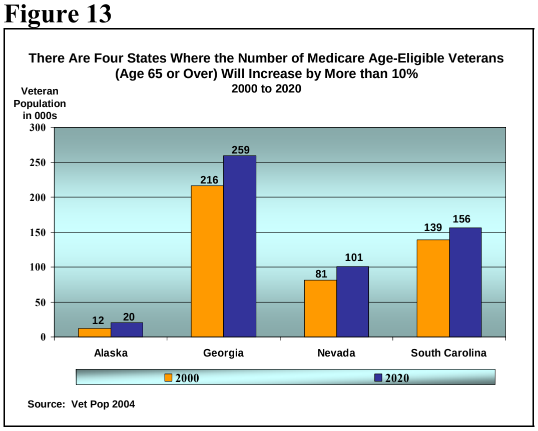
The new peak in the age 65 or older veteran population in 2013 is primarily due to the aging of the Vietnam Era cohort. The Vietnam Era cohort has special health care needs that will be different than those of their predecessors, i.e., veterans of World War II or the Korean War, but also different from the newer Gulf War Era veterans. VA must be prepared to deal with the aging of Vietnam Era veterans and their potential impact upon the VA health care system, in terms of planning for the appropriate programs and services.

Figure 12



Despite the overall decline between 2000 and 2020 in the veteran population age 65 or older, there are various regions of the country where there will actually be growth in the age 65 or older veteran population. In particular, the elderly veteran population will grow by 10% or more in the States of Alaska, Georgia, Nevada, and South Carolina (**Figure 13**).

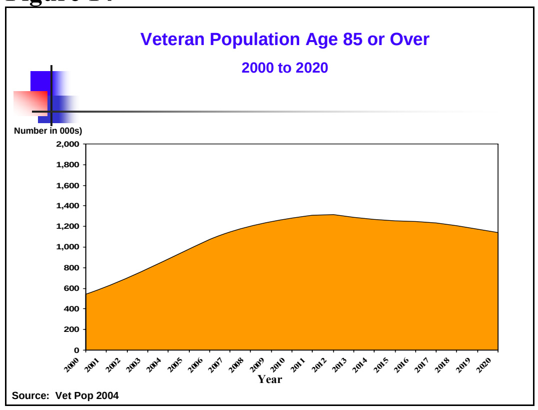
Figure 13



Between 2000 and 2020, the veteran population age 85 or older will increase by 110%, from 0.5 million to 1.1 million (**Figure 14**). The age 85 or older veteran population will peak in 2012 at 1.3 million, representing an increase of 143% over the total of 0.5 million veterans age 85 or older in the year 2000.

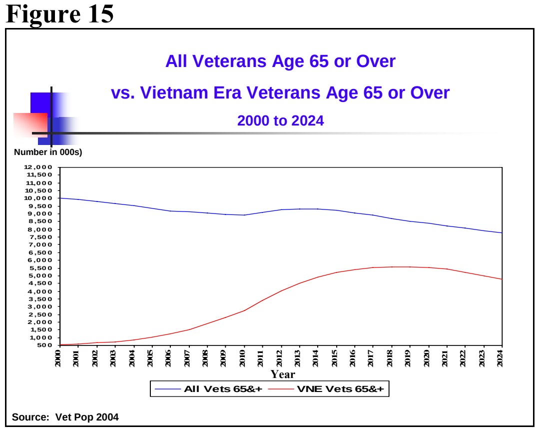
Although not readily seen in **Figure 14**, between 2000 and 2020, the age 85 or older veteran population will increase in every State in the nation. Of course, the age 85 or older population is the population at risk for use of long-term care services such as nursing home and/or home and community based services. For VA, this underscores the need to recognize gaps between demand and capacity in long-term care all across the country.

Figure 14



Further evidence of the impact of aging Vietnam Era veterans shows that in the year 2000, Vietnam Era veterans represented 5.5% of all age 65 or older veterans. In 2014, Vietnam Era veterans will comprise 52.7% of all age 65 or older veterans. In 2024, more than 6 in 10 (61.3% of all) elderly veterans will be Vietnam Era veterans (**Figure 15**).

Figure 15



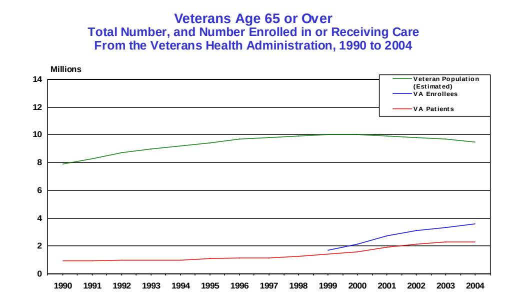
The Role of VA Health Care

Within the Department of Veterans Affairs, the Veterans Health Administration administers the largest health care delivery system in the United States, with 21 Veterans Integrated Services Networks, 157 hospitals, 869 outpatient clinics including 703 Community Based Outpatient Clinics, 134 VA Nursing Homes, 206 Readjustment Counseling Centers, and 42 Residential Re-habilitation Treatment Centers.

And VA health care continues to transition in terms of modalities of care.[13] Once a largely inpatient system, VA over the past decade has increasingly focused on outpatient care.[13]

As a consequence of eligibility reform legislation passed in 1996 and implemented in 1998, VA instituted an enrollment system for VA health care, as well as a hierarchical system of prioritizing veterans in terms of their eligibility for care. As of this writing, there are 7.5 million veterans, more than a quarter of the total veteran population, enrolled for health care in the VA system. About 5 million (two-thirds) of the total are actual users of VA health care programs and services. Furthermore, about half of all enrolled veterans are age 65 or over (**Figure 16**), underscoring the fact that elder care and long-term care are prime issues for VA.[10] The U.S. and veteran population trends discussed in this paper underscore the fact that elder care and long-term care will continue to be prime issues for VA for many years to come.

Figure 16



Note: Veterans Health Administration (VHA) enrollees are veterans who have signed-up to receive health care from VHA, and VHA patients are enrollees who have received care in each year through VHA.
Reference populations: These data refer to the total veteran population, VHA enrollment population, and VHA patient population.
Sources: Veterans Actuarial Model (Vet Pop 2001 Adj, and Vet Pop 2004) and Veterans Health Administration Enrollment and Patient files.

See: ***Older Americans, 2004: Key Indicators of Well-Being***

Health Care Challenges

What does the future portend? There are many possible scenarios, but one scenario may be trends in the actual vs. preventable causes of mortality in the United States. Modifiable health behaviors **(Figure 17)** are the leading preventable causes of mortality in the U.S.[13, 14] Over time, health behaviors in the U.S. population have led to increased prevalence of obesity and diabetes, which may soon overtake smoking as the leading cause of preventable mortality in the U.S.[14]

The numbers of overweight/obesity related deaths (poor diet and physical activity) increased dramatically from 1990 to 2000, faster than smoking related deaths. In the year 2000, data show (**Figure 17**) that smoking related deaths and deaths due to overweight/obesity are very nearly on par, and according to conservative measures.[13, 14]

Related to this, cardiovascular disease accounted for 38.0 percent, or 1 of every 2.6, deaths in the United States in 2002.[15] That is, of over 2,400,000 deaths in the U.S. from all causes in 2002, cardiovascular disease was a primary or contributing cause in about 1,400,000 of those deaths. Tobacco use, overweight/obesity, and diabetes are well-recognized risk factors in cardiovascular disease.[15]

Figure 17

Preventable Causes of Premature Mortality in U.S.

	Proportion of Total (1990)	Proportion of Total (2000)
Tobacco	19%	18
Diet / Activity	14%	17%
Alcohol	5%	4%
Infectious	4%	3%
Toxic	3%	2%
MVA	2%	2%
Firearms	2%	1%
Sexual Behavior	1%	1%
Illicit Drugs	1%	1%

Source: Mokdad, et al, JAMA, Mar 10, 2004, Vol. 291, No. 10
See, also, Dr. Jonathan Perlin, Townhall Presentation, July 2004.

Conclusion

Global, U.S., and veteran population related trends in aging show that aging, elder care, and long-term care are primary challenges VA is facing and will continue to have to deal with for many years to come. An example of a particular aging related area of concern to VA is the diagnosis and treatment of Dementia including Alzheimer's Disease.

Furthermore, due to trends in either preventable or actual causes of mortality in the U.S., overweight and obesity, diabetes, heart disease, and pulmonary disease are particular disease categories and risk factors that will command much attention from VA well into the future.

However, emphases on healthier lifestyles in concert with trends in increasing life expectancy, and reported improvements in chronic disability rates among older adults,[16, 17] and other factors, suggest that the future may not be all that bad. People the world over have the power and ability to choose the better future they want. VA, in particular, is increasingly proactive in fostering change for the better, through performance and quality monitoring, information technology, care coordination, disease management programs, and smoking cessation programs.

Also, VA has long been in the forefront of research on aging, research that impacts U.S. national health care, not just veterans' health care. In 2004 alone, VA researchers were involved in 472 aging research projects – 143 funded by VA and the rest by other sources including the National Institutes of Health. Research areas include geriatric assessment, dementia, rehabilitation of stroke patients, telemedicine, age-related muscle loss, and insulin metabolism.[18]

Indeed, just as VA has shifted its focus over the past decade from inpatient to outpatient modalities of care, VA will, in the next decade, become a more patient-centric system. In such a system, models of care cross-generational, gender, and geographic lines. The primary mechanisms for this next step in VA's evolution are new emphases on care coordination and electronic health records. By these means, VA care will move outside of clinic and hospital walls into the homes and communities of veterans, and with the fundamental goal of achieving the best possible outcomes.[19, 20]

References

1. U.S. Bureau of the Census. March 2004. "*Global Population at a Glance*". [Available at www.census.gov/ipc/wwww/wp02.html].

2. U.S. Census Bureau, *Statistical Abstract of the United States: 2004-2005* (at www.census.gov).

3. U.S. Census Bureau, 2004, "*U.S. Interim Projections by Age, Sex, Race, and Hispanic Origin*" (at www.census.gov).

4. Munnell, Alicia H. 2004. "*Population Aging: It's Not Just the Baby Boom*". Issue in Brief 16 (April). Chestnut Hill, MA. Center for Retirement Research at Boston College. [Available at www.bc.edu/crr].

5. Carnell, Brian. Total Fertility Rates (at http://www.overpopulation.com/faq/basic_information/total_fertility_rate/) .

6. Kinsella, Kevin and Victoria A. Velkoff, U.S. Census Bureau, series P95/01-1, "*An Aging World: 2001*", U.S. Government Printing Office, Washington, DC, 2001.

7. Sleebos, Joelle. "*Low Fertility Rates in OECD Countries*". OECD. 2003.

8. *Health United States, 2004*. Centers for Disease Control and Prevention, National Center for Health Statistics, National Vital Statistics System. 2004.

9. *The World Health Report, 2004*. World Health Organization. 2004.

10. Federal Interagency Forum on Aging-Related Statistics, *Older Americans, 2004:Key Indicators of Well-Being,* Federal Interagency Forum on Aging-Related Statistics, Washington DC, US Government Printing Office, November 2004.

11. Kevin Kinsella and David R. Phillips. Population Bulletin, *Global Aging: The Challenge of Success*, Population Reference Bureau. March 2005.

12. World Population Prospects, the 2002 Revision, Medium Variant. U.N. Population Division. United Nations.

13. Dr. Jonathan Perlin, Townhall Presentation, Office of the Under Secretary for Health U.S. Department of Veterans Affairs, July 2004.

14. Mokdad, Ali, H., et al. "*Actual Causes of Death in the United States, 2000*". JAMA. Vol. 291, No. 10, March 10, 2004.

15. American Heart Association, *Heart Disease and Stroke Statistics – 2005 Update*. Dallas, Texas: American Heart Association; 2005.

16. Freedman, Vicki A., et al. "*Recent Trends in Disability and Functioning Among Older Adults in the United States, A Systematic Review*", JAMA; Vol.288; No. 24; December 25, 2002.

17. Fries, James F. *Reducing Disability in Older Age*. Editorial in JAMA; Vol.288; No. 24; December 25, 2002.

18. *Geriatric Research in VA*, Fact Sheet. Office of Public Affairs, Department of Veterans Affairs, April 2005.

19. Perlin, Jonathan B. *Veteran Health Care: Meeting Tomorrow's Challenges*. Commentary, in FORUM – translating research into quality health care for veterans. VA Health Services Research & Development Service, June 2005.

20. Fihn, Stephan D. VA Research: *Responding to New Challenges*. Response to Commentary, in FORUM – translating research into quality health care for veterans. VA Health Services Research & Development Service, June 2005.

21. Robert B. Friedland and Laura Summer. "*Demography Is Not Destiny*". Center on an Aging Society, Georgetown University, Washington, DC, March 2005.

22. Gibson, Mary Jo, et al. AARP Public Policy Institute, "*Long-Term Care in Developed Nations: A Brief Overview*", October 2003, Washington DC.

23. Bogaert, Henri. "*The Budgetary Challenges Facing the EU*". Presentation at the Conference on the Economic and Budgetary Implications of Global Aging, Brussels, Belgium, March 4-5, 2003.

24. Center on an Aging Society, Georgetown University. "*A Decade of Informal Caregiving*" Washington , DC.

25. "*The World Health Report 2004*", World Health Organization.

26. "*A Changing Political Landscape As One Generation Replaces Another*", AARP, 2004, Washington, DC.

27. AARP Global Aging Program, Spring 2004, "*Global Report on Aging*". Washington, DC.

28. AARP Public Policy Institute, "*Global Aging Achieving Its Potential*", Washington, DC.

29. Paul Schreyer and Francette Koechlin, "*Purchasing Power Parities*", OECD Statistics Brief, March 2003, No. 3. OECD, Paris, France.

Contact: donald.stockford@va.gov

This page intentionally left blank.

Proposed Forecasting Methodology for Pharmacy Residency Training

Dilpreet K. Singh, M.S., M.P.A., Lori Golterman, Pharm. D.,
Gloria J. Holland, Ph.D., Linda D. Johnson, Ph.D., R.N., & Evert H. Melander, M.B.A.
Office of Academic Affiliations
Department of Veterans Affairs

Introduction

The Department of Veterans Affairs (VA) is the largest single provider of health professions training in the world. VA's Goal is to be a preferred training site and an employer of choice. Seventy percent of all physicians and a significant percent of all other health professionals including pharmacy residents in the United States receive part of their training at VA.

VA is supported by a congressionally appropriated budget, thus the available resources determine to a large degree the number of pharmacy resident positions system-wide. Facility-specific distribution of pharmacy resident positions is done by the Veterans Health Administration (VHA), Office of Academic Affiliations (OAA). This paper discusses use of results of the Learners' Perceptions Survey in determining facility-specific allocation of pharmacy trainees.

Background

VA provides health care for over 4.5 million of the nation's veterans through a network of hospitals, outpatient clinics, and nursing homes. As one of the four statutory missions, "To educate for VA and for the Nation," VA conducts an education and training program for health professions trainees through partnership with affiliated U.S. academic institutions. Each year, over 87,000 medical and associated health students, physician residents and fellows receive some or all of their clinical training at VA facilities. Of these, approximately 31,000 are physician residents/ dentists, 17,000 are medical/dental students, and 39,000 are associated health trainees. Pharmacy residents are associated health trainees.

Scope of VA Clinical Training Programs

Following are the two major VA clinical training programs:

a. **Education of Physicians and Dentists**: VA funding of approximately $432 million supports over 9,300 medical/dental resident positions each year. VA physician faculty treat veterans, supervise students and physician residents, and conduct research.

b. **Associated Health Education Programs**: Through affiliations with individual health professions schools and colleges, clinical traineeships and fellowships are provided to trainees in more than 40 professions, including nurses, pharmacists, audiologists, dietitians, social workers, psychologists, physical therapists, optometrists, podiatrists, physician assistants, respiratory therapists, and speech language pathologists. Over 36,000 associated health students receive training in VA facilities each year and provide a valuable recruitment source for new employees. The greatest majority of associated health trainees receive clinical experiences on a without compensation (WOC) basis. A student funding support of approximately $46 million is provided each year to almost 3,100 trainees.

VA Pharmacy Training Program

The VA pharmacy program is the largest training program for advanced clinical pharmacy training in the country. VA trains approximately 300 residents annually in 75 American Society of Health-System Pharmacists (ASHP) accredited residency programs. The training includes such areas as: Pharmacy Practice, Research, and Specialty Residencies (e.g., Geriatrics, Ambulatory Care, Infectious Disease, Psychiatry, Intensive Care, etc.). VA also funds the advanced training of pharmacy fellows. OAA allocates federal funds to a limited number of pharmacy residency and fellowship programs annually. The resident/fellow must be a United States citizen to be eligible for this funding.

A majority of pharmacy residents have a doctorate of pharmacy although Bachelor's and Master's degree students are also accepted into the training program. They are typically at VA for one year of training. VA allocates approximately $10 million/year to support pharmacy resident positions. In 2004, VA paid 258 residents with an annual stipend rate of $33,000 per resident.

However, there is a great demand for pharmacists with at least one year of residency. On an average, every year facilities request about 20 percent more pharmacy resident positions than allocated. If funds become available, there is always a dilemma on how best to allocate pharmacy resident positions to sites that will provide the highest quality training and maximize recruitment opportunities.

Pharmacy Residency Allocation Methodology

In the past, the number of pharmacy resident positions allocated to facilities has been based primarily on the number of positions allocated in prior years. Using this as a baseline, a criteria-based methodology has been developed consisting of both qualitative standards and quantitative measures for allocating pharmacy resident positions.

a. **Qualitative Standards:** The qualitative standards are considered in residency allocations to help determine which facilities should be allocated additional residency slots. All pharmacy residency programs must be accredited by the American Society of Health-System Pharmacists prior to being funded. A strong consideration is given to the facility, the multi-disciplinary practice sites within the medical center and the pharmacy faculty that is available to support the residency program. The VA Medical Centers that are associated with colleges of pharmacies provide additional opportunities for training pharmacy residents and students.

The facility must meet the Profession-Specific Standards of Excellence in Clinical and Inter-Professional Education/Training. The Standard of Excellence report is submitted annually by each facility. Six major areas of inquiry are as follows: (1) Does the VA residency program meet or exceed elements or standards for accreditation and adhere to VA's goal of providing patient focused multidisciplinary education and training? (2) Does the educational infrastructure (e.g., facility staff and material resources, clinical education coordinator, etc.) at local facilities and Veterans Integrated Services Networks (VISNs) support excellence in clinical education and training? (3) Does the residency program contribute to patient-focused care that reflects VA's health care priorities (primary care, geriatrics, mental health, rehabilitation) and special emphasis programs (Spinal Cord Injury or Dysfunction, AIDs, PTSD, Women's Health, Persian Gulf Syndrome, homelessness, etc.)? (4) Is the residency program at the facility affiliated with academic programs, if so, is the relationship enhanced through ongoing collaborative activities? (5) Does the multi-disciplinary education address knowledge, skills, and attitudes appropriate for successful collaboration and teamwork in clinical settings? and (6) Are the results of evaluations used to plan and implement program improvements that promote high quality education experiences for trainees?

b. **Quantitative Measures/Learners' Perceptions Survey:** The focus of this paper is to use quantitative outcomes of the Learners' Perceptions (LP) Survey as one indicator for allocating pharmacy resident positions. A brief description of the survey is given below:

The LP Survey of VA physician residents and associated health trainees has been conducted annually since 2001. The Government Performance and Results Act (GPRA) of 1993 required agencies to establish measurable performance goals and develop tools to measure progress toward organizational goals. In support of GPRA, VHA was charged with development of a national performance measure for its' teaching mission. The LP survey provides quantitative indicators to help highlight strengths and opportunities for improvement in VA clinical training experience.

A survey questionnaire was developed based on a literature search and focus group studies of students and faculty. The questionnaire was pilot tested at 22 sites before its system-wide implementation. It is a scientifically valid tool to collect perceptions of clinical trainees registered at the VA facilities. All registered trainees are contacted by surface mail or e-mail to complete the LP Survey. Starting in 2004, the LP survey has been administered through a web-based system. Response to the survey is voluntary and it requires approximately 15 minutes to complete.

For the 2004 LP Survey, from all 160 VA facilities that had clinical trainees, 41,092 trainees were surveyed. The completed survey questionnaires were received from 8,869 trainees with a 22 percent response rate system-wide; a rate comparable to other web-based surveys. However, over 90% of the pharmacy residents responded to the LP Survey questionnaire.

LP Survey Measures

Three of the LP Survey measures that are applicable to allocating pharmacy resident positions are described below. They are: Performance measure, five satisfaction domains, and would trainees consider employment with VA.

(a) **Performance Measure:** The performance measure provides an overall perceptions score of all residents' clinical experience in the VA. To obtain the performance measure score, trainees are asked, **"**On a scale of 0 to100, where 100 is a perfect score and 70 is a passing score, what numerical score would you give your most recent VA clinical training experience?" VA's performance measure goal is that by 2006 all physician residents and other trainees will give a score of 85. In 2004, all trainees gave a score of 84, whereas pharmacy residents gave a score of 88.

(b) **Five Domains**: The LP Survey also measures satisfaction of trainees for five major domains, i.e., Clinical Faculty/Preceptors, Personal Experience, Working Environment, Learning Environment, and Physical Environment.

Based on the perceptions of pharmacy residents, Clinical Faculty/Preceptors is statistically the most important domain that impacts trainee perceptions. Every domain is further rated for its elements, e.g., for Clinical Faculty/Preceptors domain, the two most important elements are teaching ability and availability of faculty. These domains are listed in priority order with Clinical Faculty/Preceptors being statistically the most important and Physical Environment the least. The most important elements are also listed under each domain.

- Clinical Faculty/Preceptors: Teaching ability, accessibility/availability, evidence-based clinical practice, being role models, and mentoring by faculty.
- Personal Experience: Enhancement of your clinical knowledge and skills, ownership/ personal responsibility for patients' care, quality of care your patients receive, enjoyment of your work, and personal support from colleagues.
- Working Environment: Faculty/preceptor morale, Internet access, workspace, Computerized Patient Record System (CPRS), and computer access.
- Learning Environment: Preparation for clinical practice, time for learning, degree of supervision, interdisciplinary approach, and spectrum of patient problems.
- Physical Environment: Facility cleanliness/ housekeeping, heating and air conditioning, parking, facility maintenance/upkeep, and availability of phones.

(c) **Employment with VA:** Since one of VA's goals is to be a future employer of choice, two employment questions were included in the LP Survey. "Before this clinical training experience, how likely were you to consider a future employment opportunity at a VA medical facility?" Possible responses included: very likely, somewhat likely, had not thought about it, somewhat unlikely, and very unlikely. Thirty-nine percent of pharmacy residents responded very or somewhat likely.

To assess whether training in the VA makes a difference in considering employment with VA, the question, "As a result of this clinical training experience, how likely would you be to consider a future employment opportunity at a VA medical facility?" is included. Possible responses included: a lot more likely, somewhat more likely, no difference, somewhat less likely, or a lot less likely. The percent of trainees considering VA employment almost doubled after receiving training at the VA as compared to before receiving VA training. Seventy seven percent of the pharmacy residents were a lot more likely, or somewhat more likely to consider employment with VA after completing their training experience at VA.

This shows that overall training at VA has a great impact on pharmacy residents considering future employment with VA and may be an important consideration in allocating pharmacy residency positions.

Application of the LP Survey Data for Allocating Pharmacy Trainee Positions

Facility-specific high performance measure scores may be used as one of the quantitative indicators for allocating pharmacy resident positions. The bar chart (Chart 1), given below, provides the facility-specific pharmacy residents' performance measure scores. Since the number of pharmacy residents for each facility is very small, an average of four years

(2001-2004) of the LP Survey performance measure score was calculated, and it ranged from 81-95. The bars show the number of facilities with their performance score, for example, 8 facilities received an average performance measure score of 89. A difference of 5 points in the performance measure score is considered statistically significant.

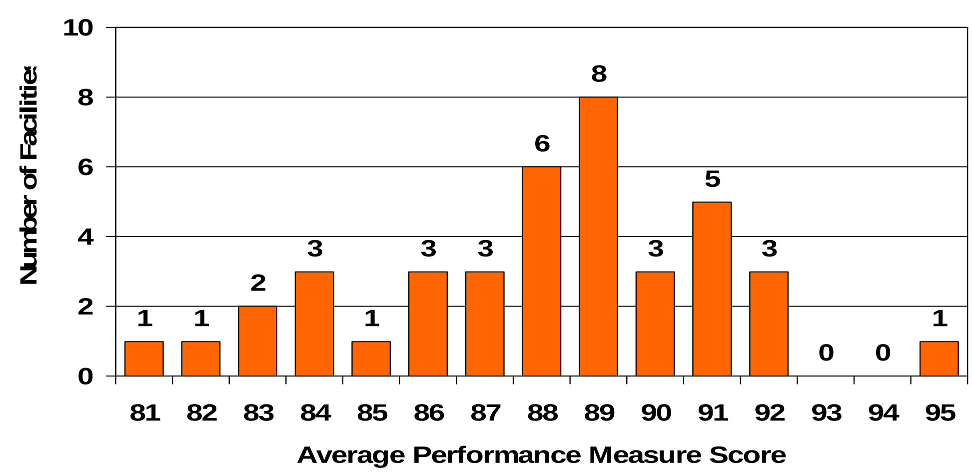
High scoring facilities (a performance score of 88 or above) should be given priority for allocating pharmacy residency positions. If a facility has not been able to fill allocated pharmacy resident positions over time (in the past three years), it would not be eligible for additional positions. Since most of the facilities have similar performance scores, satisfaction with specific domains and their elements, and the percent of pharmacy residents considering VA post residency employment may provide valuable data.

Summary/Conclusion

In summary, the LP survey is an excellent data source about the VA pharmacy residency programs. It not only provides overall perceptions of trainees but also provides data on satisfaction with each domain and its elements, such as teaching ability of faculty, availability of faculty, etc. However, results of the LP survey must be used in conjunction with the facility qualitative standards. In the long-term, a composite index should be developed to evaluate facilities based on relative weights of various qualitative and quantitative standards.

An annual review of trends of performance measure and domains will help maintain the value of this tool as being timely and reflective of the current recruitment market. By applying these trends to practice, the VA pharmacy residency program will remain competitive by meeting the needs of students pursuing residencies in VA and hopefully enhance recruitment in VA.

Chart 1. Pharmacy Residents
Average Performance Measure Score 2001-2004



Projections and Estimates in Support of Federal Tax Administration

Session Chair: Russell Geiman, Internal Revenue Service, U.S. Department of the Treasury (russell.geiman@irs.gov)

Forecasting the Weekly Volume of Individual Income Tax Return Filings

Katy Yeh, Internal Revenue Service, U.S. Department of the Treasury

The Internal Revenue Service (IRS) received 132 million individual income tax returns in 2004, the vast majority of which were filed between January and April. To ensure the volume of returns is processed timely each filing season, IRS needs accurate forecasts of the number of returns to be filed, and when. Consequently, individual return volumes are forecasted for each week throughout the filing season. Filing patterns differ for taxpayers filing returns electronically versus on paper. They also differ by the major return types—Forms 1040, 1040A, and 1040EZ. This paper looks at the data and methodology IRS staff uses in deriving these weekly forecasts.

Tracking and Estimating the Direct Revenue Effects of IRS Enforcement Actions

Bill Gammon, Internal Revenue Service, U.S. Department of the Treasury

The Enforcement Revenue Information System (ERIS) combines information about Internal Revenue Service (IRS) tax enforcement cases from several sources to track the movement of cases and the amount of direct revenue collected. It provides decision makers with critical information about case flows and revenue streams corresponding to the traditional IRS enforcement activities such as nonfiler investigations and auditing of returns. ERIS is also used to estimate the revenue to be realized from proposed IRS enforcement initiatives in the budget—providing the "return" part of a "return on investment" dimension that is quite valuable to policy makers. This paper summarizes the information available from ERIS and its use in IRS enforcement revenue estimation.

Geographically Optimizing IRS Tax Education Outreach Staffing

Jay Schnapp and Ronald Deaett, Internal Revenue Service, U.S. Department of the Treasury

The Internal Revenue Service (IRS) seeks to ensure tax compliance through a balanced approach involving taxpayer service as well as enforcement. The job of the IRS Stakeholder, Partnership, Education and Communication (SPEC) function is to conduct educational outreach which enables taxpayers to fully understand and meet their tax obligations "upfront"—thereby reducing the number of contacts with IRS. To accomplish its mission, SPEC has locations nationwide and develops partnerships with key external stakeholders who can help inform the public. However, a major challenge for IRS is determining where best to locate SPEC employees. The model we present in this paper optimizes the geographic allocation of SPEC personnel based on eight factors. The model assigns each IRS geographic territory a score that determines where the next staff year should go.

The Effects of State Mandates on Federal Electronically Filed Returns

Michelle Chu and Melissa Kovalick, Internal Revenue Service, U.S. Department of the Treasury

This paper examines the effects of state electronic filing mandates on the number of federal individual income tax "e-file" returns. State e-file mandates are gaining in popularity. The goal of this article is to articulate a forecasting model that effectively predicts the marginal volume impact of state criteria on the federal e-file program. Specifically, the analysis explores various relevant contributing factors such as the effects of participation in the joint Fed/State return filing programs, and the existence and severity of varying penalties imposed by the states. This paper attempts to determine which state criteria have the greatest marginal effect on the volume of federal returns filed electronically.

This page intentionally left blank.

Forecasting the Weekly Volume of Individual Income Tax Return Filings

Katy Yeh, Internal Revenue Service
U.S. Department of the Treasury

Background

In 2004, Internal Revenue Services (IRS) received 132 million total individual returns, 92% of which is received by the end of April. To ensure that the massive volume of tax returns is processed timely, IRS needs to plan enough staff and equipment to receive and process these returns. During the filing season IRS hires many seasonal workers. When to bring in the extra staff depends in part on the amount of paper returns to be received. In addition to paper returns, many returns are filed electronically, e.g. in 2004, 61.5 million (47%) of total individual returns were e-filed. IRS needs computer equipment to receive the growing electronic returns, especially at peak periods. Therefore, part of the plan for a successful and smooth filing season relies on accurate forecasts of the number of returns to be filed each week by type of return. The forecasts of the weekly volume of returns are not only used for scheduling staffing, they are also used to monitor the progress of the filing season.

The US individual income tax return is mainly composed of Forms 1040, 1040A, and 1040EZ. Each form serves a unique segment of the tax paying public. Form 1040EZ is the simplest form and is designed for filers who are single or married couple filing jointly, do not claim any dependents, source of income is limited to wages and salaries under $100,000, and makes no adjustment to income. Form 1040A allows filers to claim more source of income, deductions, and tax credits, but few other unique "schedules." Form 1040 is the most complicated return of the three. Taxpayer can use it to file schedules such as Schedule A (itemized deduction), Schedule B (interests and ordinary dividend), Schedule C (profit or loss from business), etc. There are also some international returns, Form 1040NR (U.S. Nonresident Alien Income Tax Return), Form 1040PR, for residents of Puerto Rico, and Form 1040SS (U.S. Self-Employment Tax Return) for residents of US territories: Guam, American Samoa, CNMI, Virgin Islands, and Puerto Rico.

Return Processing and Data Issues

In 2004, IRS had seven campuses that processed individual returns: Andover, Atlanta, Austin, Fresno, Kansas City, Memphis, and Philadelphia. In addition to processing paper returns, five of these campuses - Andover, Austin, Kansas City, Memphis, and Philadelphia campuses also processed electronic returns. International returns were processed at the Philadelphia campus.

As returns are received by campuses, they are counted and put into processing pipeline. Reports are then generated showing the number of returns received and processed. The report we use in forecasting the receipt volume is the "IIRAPHQ" report, which roughly stands for Individual Income tax returns Received And Processed report for Head Quarters analysts. It shows the weekly counts of paper and e-file returns received in each campus as well as cumulative return volume through that week. In 2004 IRS received an average of 3.4 million paper returns per week during months of February and March. A total of 27.9 million returns were received in the first three weeks of April. In the week of April 15 alone IRS received 13.5 million paper returns. Imagine multiple tractor-trailers each carrying 40,000 pounds of mail to the campuses. Because of the massive number of paper returns received by the campuses, it is impossible to count returns one by one when they first arrive at the loading dock. So paper return volumes are sometimes "estimated" for purpose of the weekly IIRAPHQ report.

These estimated return volumes sometimes can cause problems in the data. In general, when paper returns are received, they are extracted and batched. Once extracted, returns are sorted by form (1040, 1040A, etc.) and by with and without remittance (i.e. whether they had a check attached). Paper returns without remittance and all electronically filed returns are classified as Other-Than-Full-Paid (OTFP) returns. With remittance paper returns are further sorted into Part-Paid (a subset of OTFP) and Full-Paid (FP) returns. These sorted returns are then grouped into a batch. Information for each batch is entered into the inventory control system, Batch/Block Tracking System (BBTS). When there's a large amount of inventory, receipts cannot be extracted and/or batched the same day. In this case, return counts are estimated by weight, volume, "eye-balled" based on experience, or estimated from the weekly projection using historical information.

As the batches move down the processing pipeline, however, BBTS updates and makes adjustment to previously estimated volume. The adjustments to previous weeks' counts are reflected in the cumulative receipt volume in the IIRAPHQ report. So there is inherent noise in the data. Furthermore, the IIRAPHQ report is issued every week, if there's an adjustment made to receipt volume it does not show to which week the adjustment was made. In our forecasting method, we use the cumulative receipt volume. In theory, the cumulative receipt volume should increase as filing season progresses. However, because adjustment are made each week, depending on the magnitude of the adjustment, sometimes the IIRAPHQ report will show a decrease in cumulative return volume. This is usually the case with paper FP returns. Therefore, it is very difficult to get precise counts of paper returns, especially for weeks when large adjustments were made.

Besides the IIRAPHQ report, however, there is another source that provides paper return counts called the Production Information and Monitoring System (PIMS). PIMS can track return receipts and make volume adjustments on a daily basis. Since return volumes are updated daily, depending on when the data query is run, receipt volume changes. To overcome this problem, data from PIMS is extracted after June when most of the returns have been processed. As more returns get processed, receipt volume will not change much, if at all, in the earlier weeks. For this reason, we rely mainly on the PIMS data in our forecasting models to give us more accurate counts of paper returns. E-filed returns do not have this problem. They are received electronically and require no estimation. There is generally no error in e-file return count from IIRAPHQ report.

Filing Pattern

Taxpayers who file returns on paper have different filing patterns than those who file electronically. E-filers have a tendency to file earlier in the year as shown in Graph 1 in the appendix. In contrast most paper filers generally wait until the final weeks in April. Early filers usually submit their returns around the end of January after receiving the necessary document to file their returns such as Forms W-2 (Wage and Tax Statement), 1099-INT (Interest Income), etc. In 2004 most of the returns filed between January and March were e-filed. By April 9, cumulative return volume for e-file was 51.7 million and 37.7 million were paper. Two weeks later, the number of paper return jumped to 60.4 million, an increase of 22.7 million, while e-file returns increased only 7.8 million to 59.6 million. In addition to filing options, each paper form has a separate share of the taxpayer population and a different filing pattern. Of the 70.7 million paper returns received in 2004, 50.9 million were Forms 1040/NR/PR/SS, 11 million were Form 1040A, and the remaining 8.6 million were Form 1040EZ. Relatively speaking the filing patterns for the latter two "simpler" forms tend to be bit "earlier" than for the forms 1040/NR/PR/SS.

Electronic returns are broken out by how returns are transmitted to IRS. TeleFile returns are returns filed over telephone. This electronic filing medium has been decreasing over the past several years; only 2.9 % (3.8 million) of total individual returns were telefiled last year. Qualified tax payers can also file returns online using tax preparation software or electronic filing services provided by Free File Alliance, a group of tax software companies. Practitioner e-file returns are basically returns prepared by paid tax professionals who send returns electronically to IRS. Due to promotion of electronic filing, online and practitioner returns have been growing; online returns increased 2.6 million from 2003 to 14.6 million in 2004, and practitioner returns grew 6.2 million to 43.2 million. Consequently, total paper return filings have been declining in recent years.

Converting Data to Cumulative Percentages

To forecast the weekly volume of individual returns, instead of forecasting the nominal return volume, we actually forecast the cumulative filing percent, which the cumulative week receipt volume divided by the total receipt volume for the entire year. In general, there's a downward trend in these cumulative filing percentages across forms, filing method, campuses, and at the US level. Graph 2 displays the cumulative filing percent for total electronically filed returns.

Because the IIRAPHQ report tallies returns by week on Fridays, not by calendar date, year-to-year data are lined up by the closest week-ending date. However, the table below shows how this week-ending measurement approach actually shifts calendar dates over the years and results in a loss of one or two calendar days compared to the corresponding weeks of prior years. Since the cumulative filing counts through a given week generally include one or two fewer filing days compared to the year before, the effect is a downward trend in the cumulative percent curve. And it is this steady downward pattern that is easier for the statistical models to forecast.

Comparable Week			Cumulative Filing Days		
2003	2004*	2005	2003	2004*	2005
Jan 3	Jan 2		3	2	
Jan 10	Jan 9	Jan 7	10	9	7
Jan 17	Jan 16	Jan 14	17	16	14
⋮	⋮	⋮	⋮	⋮	⋮
Apr 11	Apr 9	Apr 8	101	100	98
Apr 18	Apr 16	Apr 15	108	107	105

* Leap year

Another filing pattern displayed in the cumulative filing percent is a cyclic pattern that is most pronounced in the weeks around April 15 filing deadline. Every year the day of the week on which April 15 falls is different. And the day in which April 15 falls on has a dramatic impact on the volumes of paper returns received that week. If April 15 falls in the early part of the week, IRS would receive the bulk of returns filed on or few days before the 15. If April 15 falls in the later part of the week, however, the bulk of the returns filed on or few days before the 15 will take few days to be delivered by the US postal service. By the time IRS receives those returns, they are counted as receipt volume for the week after April 15.

Also, when April 15 falls on Saturday or Sunday, the filing deadline becomes the next Monday. For example, in 1999 April 15 fell on Thursday. In 2000, however, April 15 fell on Saturday. So the actual filing date is April 17. Because the data are lined up by the week of filing deadline, the comparable week of April 16, 1999 to 2000 is the week ending April 21, 2000, which results in a "level shift" in the data for 1999 and 2000. Graph 3 shows the cumulative filing percent of paper returns for years 1994 through 2004. The cumulative filing percent for week ending 4/16/99 is 71.5%, and the cumulative filing percent for 4/21/00 is 84.4%. As April 15 cycles through the week over the years, the cumulative filing percent for the week of April 15 exhibits a rather dramatic cyclic pattern of growth and decline. Similar cyclic pattern for the electronic returns occur around the first week of February (see Graph 2).

Forecast Methodology: Data Series Construction

Although the main filing season runs from January through April 15, the campuses plan all the way through June for the period. Therefore, we forecast the weekly receipt volume through the last week of June. As part of our methodology we actually prepare two sets of forecasts, initially one at the US level, and another at the campus level. The campus level forecasts use both the PIMS data and the IIRAPHQ reports. The PIMS data were available in more recent year. The data used in the campus level forecasts were IIRAPHQ reports from 1994 to 2002 and the 2003 PIMS data. There were 1,236 weekly series to forecast as summarized below.

7 campuses x 3 forms x 25 weeks = 525 Paper (OTFP)
7 campuses x 3 forms x 16 weeks = 336 Paper (FP)
5 campuses x 3 forms x 25 weeks = <u>375 E-file</u> Returns
1,236 series

IRS has 7 campuses and three major forms (Forms 1040/NR/PR/SS, Form 1040A, and Form 1040EZ). The number of Forms 1040NR/PR/SS is small (700,000 returns in 2004) and is combined with Form 1040. There are 25 applicable weeks from January through June for IRS processing purpose. Also, Full-Paid (FP) returns are defined by IRS as paper returns received on or before April 22 with balance due and check enclosed. So there are only 16 applicable weeks for FP returns. The campus level forecasts are then summed to the derive one set of US weekly projections.

However, a second set of US level forecasts is also developed based on the IIRAPHQ report and have a total of 171 series. In this US level projection effort, we forecast weekly receipt volumes from January through the first week of May and the last week of June, about 19 weeks, for the following 9 aggregate of returns: US total, total FP, total paper, total paper Forms 1040/NR/PR/SS, total paper Form 1040A, total paper Form 1040EZ, Standard Electronically filed returns (Std ELF), online, and TeleFile.

Table 1 in the appendix shows how the campus level returns are summed to the US total, and the comparisons that are made to the second set of US level forecasts developed independently.

At the time when we forecasted the weekly volume for filing season 2004, the data available were weekly volumes from years 1994 to 2002 and partial year (January – August) data for 2003. As a result, we had to estimate the 2003 end of year return volumes first using the latest filing figures in the IIRAPHQ report and then derive estimated cumulative filing percents for 2003.

Forecast Methodology: Trend Methods

The models used in forecasting these nearly 1,400 unique data series were exponential smoothing models: simple exponential, Linear Holt, damped trend, Double Brown, linear regression and random walk with drift. Exponential smoothing models tend to produce forecasts that continue the most recent trend. We use SAS, Statistical Analysis Software, to fit the six statistical models and select models with the smallest root mean square error. We then evaluate these initial forecasts by looking at the year-to-year percent change and the weekly receipt percent. The year-to-year percent change is the difference of the two adjacent year's cumulative filing percent. For example, the year-to-year percent change for 2004 is the cumulative filing percent of 2004 minus the cumulative filing percent of 2003 for the comparable week. The weekly receipt percent is the proportion of total return volume received in that week. For example, the weekly receipt percent for the week of 3/12/04 is the cumulative filing percent of 3/12/04 minus the cumulative filing percent of 3/05/04.

There are several things to watch out for on these initial forecasts. One of them is negative forecasts. The cumulative filing percents are small for weeks in January. For paper FP returns, they are less than 1%. Small cumulative filing percents combined with downward filing pattern sometimes result in models producing forecast of negative value. Another thing to check for is non-cumulativeness. Because we forecast <u>cumulative</u> filing percent, the forecasts should increase as filing season progresses. Non-cumulative forecasts happen when the cumulative filing percents are very close together. In 2004, IRS received about 97% of the total e-file returns by the end of April. The weekly receipt percent after May is less than 0.5%. Models such as Double Brown and Random Walk with Drift would forecast steeper drop in cumulative filing percent than other models. Non-cumulative cumulative filing percent implies negative weekly receipt percent. Like the negative cumulative filing percent forecast, this cannot happen in reality. In certain cases, the statistical models cannot provide reasonable forecasts. We take the average or differences of most recent years' cumulative filing percents as the forecast or use the naïve forecast, the most recent actual cumulative filing percent.

In addition, forecasts for the weeks around April 15 required special attention. Approximately 30% of the total individual returns are received in the three weeks window period around April 15. Because of the cyclic filing pattern around this filing due date, the cumulative filing percent for these three weeks changed more dramatically than other weeks in the filing season. Looking at the year-to-year percent change in the 2004 data is even more problematic because of leap year. For example, even though April 15 fell on Thursday in both 1999 and 2004, it was on Wednesday in 1998 versus on Tuesday in 2003. Thus the week of April 15 in 2004 include only one post-due-date filing day, Friday April 16, whereas in 2003 there were three post-due-date filing days (see Table 2). Since leap year occurs every 4 years and there are 7 days in a week, in order to get a true comparison, we would have to compare the year-to-year change of 2004 to the year-to-year change of 1976. It takes 28[a] years to complete the cycle. However, the data we have only go back to 1994. Therefore, we look at the value of historical cumulative filing percent where April 15 fell. The cumulative filing percent for US total paper for the weeks of 4/9, 4/16, and 4/23 in 2004 were: 53.3%, 72.3% and 85.4% respectively. The cumulative filing percent for the same weeks in 1999 were 52.1%, 71.5% and 84.8%, respectively. This serves as some guidance on where the forecast should be.

To derive the cumulative filing volume for each week we multiply the cumulative filing percent to the projected end-of-year return volume. The cumulative filing volume from PIMS data is generally higher than the corresponding cumulative filing volume in the IIRAPHQ report, hence the sum of the campus forecasts are usually higher than the US level forecasts. If both the projected weekly return volumes for the US level and campus level look reasonable and differ by less than 5%, we used the campus level forecasts. Otherwise we took the average of the two forecasts as the new US level forecasts and make adjustments to the campus forecasts so that the adjusted campus forecasts will sum to the new US

[a] In this case, period is the smallest number that is multiple of both 4 and 7, 4 x 7 = 28.

level forecasts. The process was repeated in checking the adjusted cumulative filing percents, the adjusted weekly receipt percents, make further adjustment if necessary until all the forecasts make sense. The final forecasts were somewhere between the initial US level forecasts and the campus level forecasts.

Forecast Accuracy

It is always exciting, and a sense of relief, when our forecasts are on target and anxious when our forecasts are off target. Forecast accuracy in our weekly projections varied for different return type, by forms, campus, etc. Table 3 shows how well our projections perform in the 2004 filing season at the US level for the US total individual returns, total paper, and total electronically filed returns. Except for the weeks early in the filing season, January through mid-February, and the week of April 15, our forecast errors were less than 1% for US total returns, errors of 1% to 2% for total paper, and total e-filed returns had error of less than 1% for the last two weeks of February and most of March and about 2% for weeks after April. The forecast error of -7.29% and -13.84% for the week ending April 16 for US total and total paper respectively is due to our failure to adequately capture the cyclic filing pattern. Noise within the data play a part in the accuracy as well. Projection error for e-file returns is less than the projection error of paper returns.

In addition to the noise in the data, there are certain elements built into these projections that contribute to its accuracy and inaccuracy. As Table 3 shows, projection error is higher for weeks in January than all other weeks in the filing season. Few tax returns are filed in January and the small cumulative filing percents make it difficult to model and forecast.

The projection of the cumulative weekly receipt volume is actually forecast of forecast. The cumulative weekly filing percent is itself a forecast, as is the end of year total return volume we apply our percentages against. All forecasts come with errors. However, errors can either cancel out or accumulate. If the projected end of year total return volume is close to the actual volume, whereas the forecasted cumulative filing percent is off, the resulting weekly receipt volume forecast will be off and vice versa. If both the projected end of year total return volume and the cumulative filing percent are off, depending on the magnitude of errors, as long as the forecast errors cancel, the cumulative receipt volume forecast might still be accurate. Table 4 illustrates this.

Even though the end of year return forecast of 59.8 million for total e-file is 2.8 % lower than the 61.5 million received, since the forecasted cumulative filing percent is higher than the actual cumulative filing percent with forecasting error by about the same magnitude as the forecasting error of the end of year return volume, the projected cumulative receipt volumes for weeks from February 20 to March 19 are quite accurate with error of less than 1%. For weeks after April 23, where the cumulative filing percent forecast is close to the actual cumulative filing percent, the cumulative receipt volume forecasts are off the mark by about 2%.

There is another possible error introduced by using estimated 2003 cumulative filing percent. Forecasts produced by exponential smoothing models rely on the most recent data. If the estimated 2003 cumulative filing percent is higher/lower than the actual cumulative filing percent, then the models tend to forecast higher/lower cumulative filing percent for 2004. This is important because we forecast weekly return volumes throughout the filing season. The forecasted cumulative filing percents differ from the actual values by certain percentage consistently. In the case of total e-file returns, the cumulative filing percent forecasts are about 1% higher than the actual values for the weeks from mid-February through early April.

Conclusion

The weekly return volume forecasts are very important to IRS staff responsible for processing tax returns. Over-projection of return volumes can lead to a waste of resources, and under-projection of return volumes could lead to shortage of staff/equipment that could cause backlog of work and delay in processing. This paper examined the data, projection methodology and accuracy of the weekly level individual return filings developed for the 2004 filing season.

Acknowledgements

My sincere thanks to Francis Sharp who provided the IIRAPHQ reports, PIMS data and information on campus processing operations. Without his help, writing this paper would not have been possible.

Note: The views expressed in this paper represent the opinions and conclusions of the author and do not necessarily represent those of the Internal Revenue Service.

References

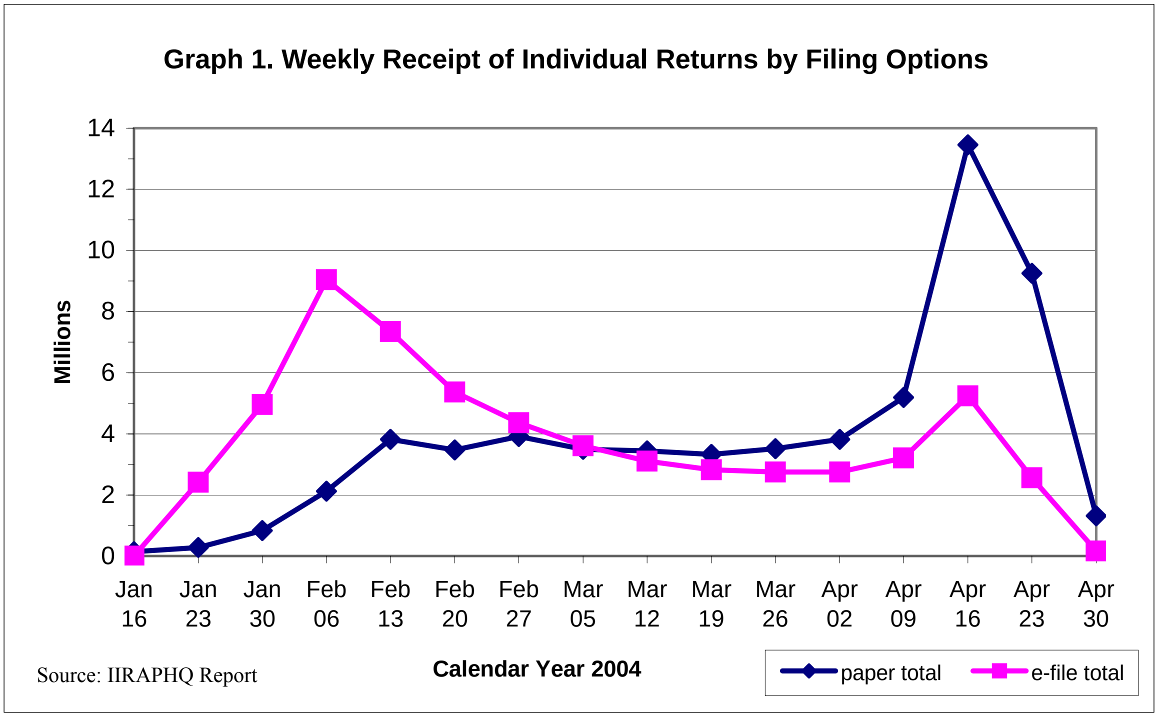
Makridakis, Spyros, Wheelwright, Steven and Hyndman, Rob. (1998). Forecasting: Methods and Applications. New York: John Wiley and Sons, Inc.

SAS/ETS User's Guide, Version 8 (1999). Vol. 1. Cary: SAS Institute, Inc.

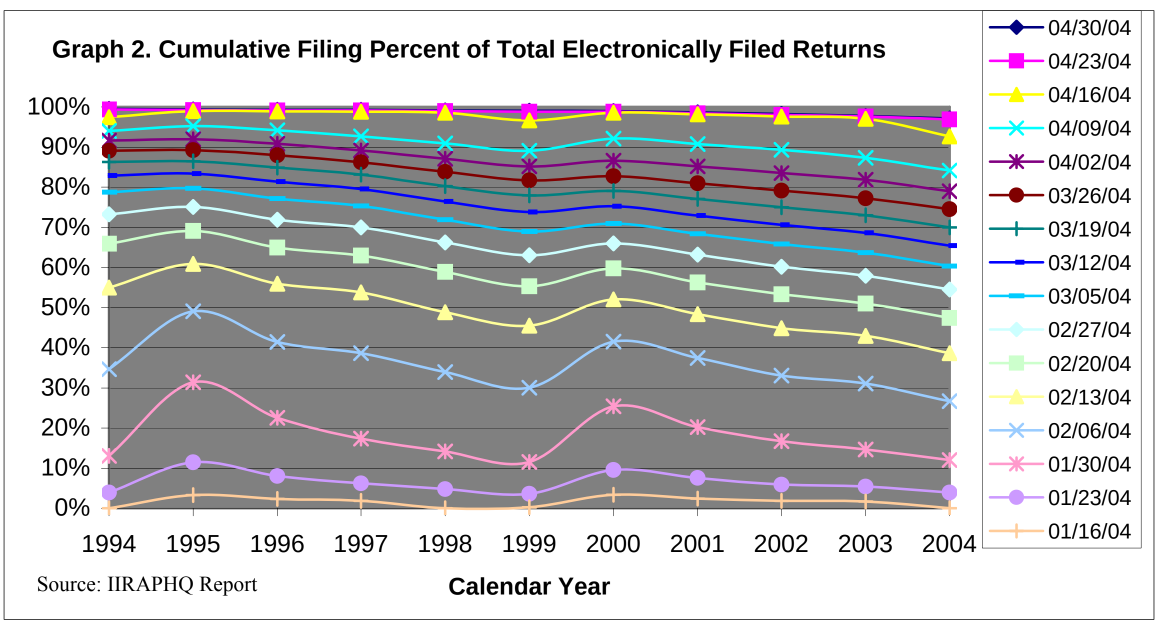
"Topic 154 - 2003 Forms W-2 and Form 1099–R (What to Do if Not Received)." 2004. Internal Revenue Service. December 2004 http://www.irs.gov/taxtopics/tc154.html.

Woodfield, T. J. (2002) Introduction to Time Series Forecasting Using SAS/ETS Software Course Notes. Cary: SAS Institute, Inc.

Appendix

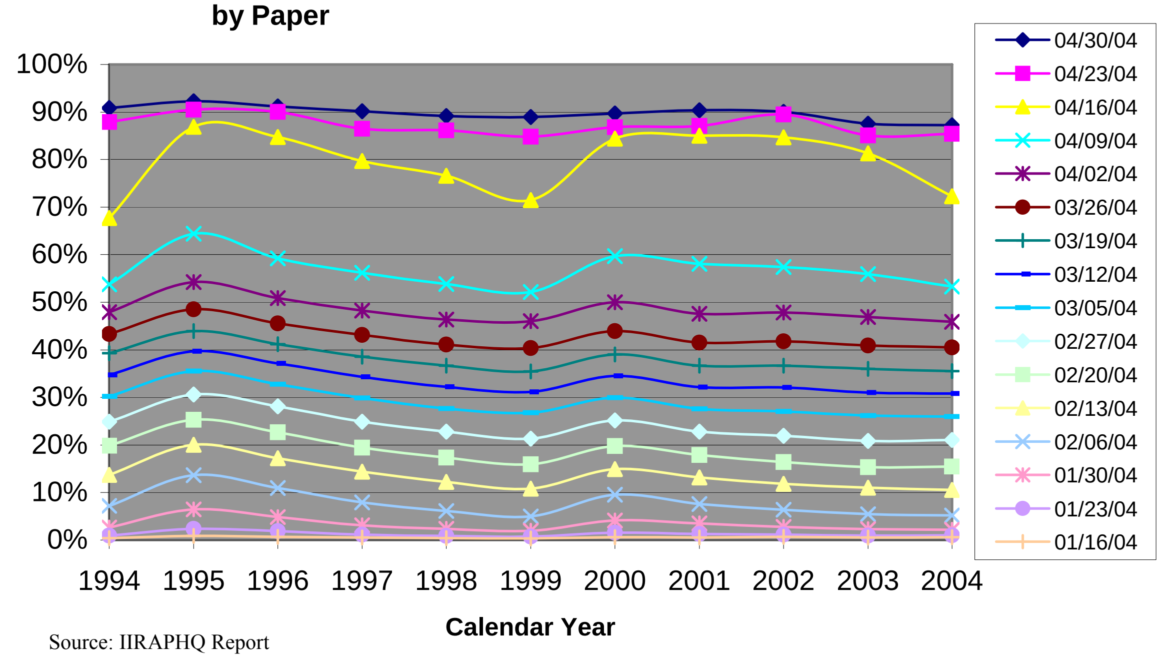


Graph 1. Weekly Receipt of Individual Returns by Filing Options
Millions
14
12
10
8
6
4
2
0
Jan 16
Jan 23
Jan 30
Feb 06
Feb 13
Feb 20
Feb 27
Mar 05
Mar 12
Mar 19
Mar 26
Apr 02
Apr 09
Apr 16
Apr 23
Apr 30
Source: IIRAPHQ Report
Calendar Year 2004
paper total
e-file total



Graph 2. Cumulative Filing Percent of Total Electronically Filed Returns
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
04/30/04
04/23/04
04/16/04
04/09/04
04/02/04
03/26/04
03/19/04
03/12/04
03/05/04
02/27/04
02/20/04
02/13/04
02/06/04
01/30/04
01/23/04
01/16/04
Source: IIRAPHQ Report
Calendar Year

Graph 3. Cumulative Filing Percent of Individual Returns Filed by Paper



Source: IIRAPHQ Report

Table 1

Campus Level

	Forms			
Paper	1040/NR/PR/SS	1040A	1040EZ	Total
Full-Paid	Returns processed by 7 campuses: Andover, Atlanta, Austin, Fresno, Kansas City, Memphis, and Philadelphia			Total FP
Other-Than-Full-Paid				
Total	1040/NR/PR/SS	1040A	1040EZ	Total Paper

	Forms			
E-file	1040	1040A	1040EZ	Total
Standard Electronic Filing	Returns processed by 5 campuses: Andover, Austin, Kansas City, Memphis, and Philadelphia			Std ELF

US level

US Total = total paper + total e-file
Other-Than-Full-Paid = US total – Full Paid
Full-Paid
Total Paper
Paper Form 1040/NR/PR/SS
Paper Form 1040A
Paper Form 1040EZ
Total E-file = Std ELF + TeleFile
Standard Electronically Filed (Std ELF)
Online
Practitioner = Std ELF - Online
TeleFile

Table 2. Year-to-Year Cumulative Filing Percent Change for US Total Paper Returns

Week Ending	1995 Sat=Mon	1996* Mon	1997 Tues	1998 Wed	1999 Th	2000* Sat=Mon	2001 Sun=Mon	2002 Mon	2003 Tue	2004* Th
04/09/04	10.63%	-5.16%	-3.04%	-2.33%	-1.72%	7.58%	-1.64%	-0.63%	-1.51%	-2.65%
04/16/04	19.19%	-2.15%	-5.03%	-3.14%	-5.05%	12.91%	0.60%	-0.36%	-3.36%	-8.99%
04/23/04	2.63%	-0.45%	-3.56%	-0.31%	-1.37%	2.02%	0.24%	2.46%	-4.44%	0.31%

Cumulative Filing Percent for US Total Paper Returns

Week Ending	1994 Fri	1995 Sat=Mon	1996* Mon	1997 Tues	1998 Wed	1999 Th	2000* Sat=Mon	2001 Sun=Mon	2002 Mon	2003 Tue	2004* Th
04/09/04	53.74%	64.38%	59.22%	56.18%	53.85%	52.13%	59.71%	58.07%	57.44%	55.92%	53.27%
04/16/04	67.69%	86.88%	84.73%	79.69%	76.55%	71.50%	84.40%	85.01%	84.65%	81.29%	72.30%
04/23/04	87.86%	90.48%	90.04%	86.48%	86.17%	84.79%	86.81%	87.06%	89.51%	85.07%	85.38%

* Leap year

Table 3. Projection Error of Cumulative Weekly Volume of Individual Returns for 2004 (volume in thousands)

Week	US Total			Total Paper			Total E-file		
Ending	Actual	Proj'd	Error	Actual	Proj'd	Error	Actual	Proj'd	Error
Jan 09	264	253	4.17%	264	253	4.17%	-	-	N/A
Jan 16	411	1,281	-211.68%	411	399	2.92%	-	882	N/A
Jan 23	3,114	3,625	-16.41%	694	671	3.31%	2,420	2,953	-22.02%
Jan 30	8,920	9,410	-5.49%	1,536	1,585	-3.19%	7,384	7,825	-5.97%
Feb 06	20,092	21,480	-6.91%	3,663	3,911	-6.77%	16,429	17,568	-6.93%
Feb 13	31,263	32,261	-3.19%	7,486	7,647	-2.15%	23,777	24,614	-3.52%
Feb 20	40,110	40,032	0.19%	10,965	10,722	2.22%	29,145	29,310	-0.57%
Feb 27	48,390	48,149	0.50%	14,879	14,617	1.76%	33,511	33,533	-0.07%
Mar 05	55,493	55,620	-0.23%	18,369	18,591	-1.21%	37,124	37,030	0.25%
Mar 12	62,038	61,988	0.08%	21,810	22,070	-1.19%	40,228	39,918	0.77%
Mar 19	68,182	67,962	0.32%	25,134	25,326	-0.76%	43,048	42,636	0.96%
Mar 26	74,453	74,476	-0.03%	28,653	29,211	-1.95%	45,800	45,265	1.17%
Apr 02	81,021	81,139	-0.15%	32,467	33,033	-1.74%	48,554	48,106	0.92%
Apr 09	89,427	90,514	-1.22%	37,657	39,322	-4.42%	51,770	51,191	1.12%
Apr 16	108,116	115,996	-7.29%	51,108	58,181	-13.84%	57,008	57,815	-1.42%
Apr 23	119,935	120,003	-0.06%	60,359	61,754	-2.31%	59,576	58,249	2.23%
Apr 30	121,421	121,345	0.06%	61,676	62,957	-2.08%	59,745	58,388	2.27%
May 07	121,953	122,114	-0.13%	62,086	63,508	-2.29%	59,867	58,605	2.11%
Jun 25	124,395	124,682	-0.23%	64,016	65,581	-2.44%	60,379	59,101	2.12%
Dec 31	132,200	131,597	0.46%	70,693	71,794	-1.56%	61,506	59,803	2.77%

Source: Weekly Tracking Report 2004

Table 4. Projection Error of Calendar Year 2004 U.S. Total E-filed Individual Returns

Week	Cumulative Return Volume*				Cumulative Filing Percent			
Ending	Actual	Proj'd	Diff	Error	Actual	Proj'd	Diff	Error
Jan 09	0	0	0	N/A	0.00%	0.00%	0.00%	N/A
Jan 16	0	882	-882	N/A	0.00%	1.47%	-1.47%	N/A
Jan 23	2,420	2,953	-533	-22.02%	3.93%	4.94%	-1.00%	-25.50%
Jan 30	7,384	7,825	-441	-5.97%	12.01%	13.08%	-1.08%	-8.99%
Feb 06	16,429	17,568	-1,139	-6.93%	26.71%	29.38%	-2.67%	-9.98%
Feb 13	23,777	24,614	-837	-3.52%	38.66%	41.16%	-2.50%	-6.47%
Feb 20	29,145	29,310	-165	-0.57%	47.39%	49.01%	-1.63%	-3.43%
Feb 27	33,511	33,533	-22	-0.07%	54.48%	56.07%	-1.59%	-2.92%
Mar 05	37,124	37,030	94	0.25%	60.36%	61.92%	-1.56%	-2.59%
Mar 12	40,228	39,918	310	0.77%	65.41%	66.75%	-1.34%	-2.06%
Mar 19	43,048	42,636	412	0.96%	69.99%	71.29%	-1.30%	-1.86%
Mar 26	45,800	45,265	535	1.17%	74.46%	75.69%	-1.23%	-1.65%
Apr 02	48,554	48,106	448	0.92%	78.94%	80.44%	-1.50%	-1.90%
Apr 09	51,770	51,191	579	1.12%	84.17%	85.60%	-1.43%	-1.70%
Apr 16	57,008	57,815	-807	-1.42%	92.69%	96.68%	-3.99%	-4.30%
Apr 23	59,576	58,249	1,327	2.23%	96.86%	97.40%	-0.54%	-0.56%
Apr 30	59,745	58,388	1,357	2.27%	97.14%	97.63%	-0.50%	-0.51%
May 07	59,867	58,605	1,262	2.11%	97.34%	98.00%	-0.66%	-0.68%
Jun 25	60,379	59,101	1,278	2.12%	98.17%	98.83%	-0.66%	-0.67%
Dec 31	61,506	59,803	1,703	2.77%	100.00%	100.00%	0.00%	0.00%

* in thousands

Tracking and Estimating the Direct Revenue Effects of IRS Enforcement Actions

Bill Gammon and Peter Rose
Internal Revenue Service, U.S. Department of the Treasury

This paper is intended to describe how we in the Internal Revenue Service track direct enforcement revenue and how we use the data to estimate the effects of IRS enforcement actions. It will also define what we (IRS) mean when we use the term enforcement revenue, along with a brief historical context for it.

Direct Enforcement Revenue – a Definition

In its simplest form, enforcement revenue is defined as direct revenue generated as a result of an enforcement action. Another somewhat simplistic definition is that revenue generated when we expend enforcement resources. A more formal definition would be the one used by GAO – "…enforcement revenue includes the direct revenue resulting from enforcement actions, such as audits, delinquent return investigations, or efforts to collect delinquent tax debts."[1]

For these purposes, indirect revenue refers to "…any revenue that might result indirectly from those enforcement actions, such as might occur if voluntary compliance increased as a result of an increase in IRS' enforcement presence."[2]

At this point we should differentiate for the reader enforcement revenue, and the estimation thereof, from "tax collections." Tax collections are all monies received from the levy and collection of taxes - $1.738 trillion (net of refunds) for FY '04[3]**.** Enforcement revenue, which totaled $43.1 billion in FY '04,[4] is essentially a subset of net tax collections.

Enforcement actions on which we expend enforcement resources include the commonly considered things like audit determinations and payment demand notices, but a number of other actions also trigger enforcement revenue. The computer matching program used to verify reported wages, interest, dividends and distributions that identifies under reporting results in enforcement revenue, as does the collecting of penalties associated with taxes. Offsetting an outstanding liability by the reduction of a subsequent refund is considered enforcement revenue as well. For example, a taxpayer owes $500 for tax year 2003. When the taxpayer files the 2004 return claiming a $1,000 refund, $500 of that refund would be used to offset the outstanding liability from 2003, satisfying the debt and generating enforcement revenue.

Another way to explain enforcement revenue is with examples of what it is not. Withholding, estimated tax payments and amounts sent in (remitted) when the return is filed are not considered enforcement revenue because they were voluntarily made by the taxpayer, or on the taxpayer's behalf, without any actions on IRS' part.

The last source of revenue mentioned above – offsets – introduces the issue of timing for enforcement revenue. IRS enforcement actions and tax payments frequently cross fiscal years. For example, the audit of a large multi-national business can take several years to complete. The internal appeal process can add a number of years, and finally a court review can add even more time. All the while, if there is a portion of the additional tax that is not being contested, the taxpayer might be making payments on it to stop the accrual of interest.

Several issues bear on how and when we account for the revenue in such a case. First, we must establish taxpayer's legal liability for the tax. In the multi-national example above, the taxpayer agrees to the portion not being contested, but the final amount at issue cannot be determined until the courts rule. No legal liability can be determined until that time. Once the amount is determined, an "assessment"[5] is posted to the account, and the case will be closed, with a formal notice of the amount due sent to the taxpayer. At this point, the legal liability has been established,

[1] See Tax Administration: Assessment of IRS' Report on Its Fiscal Year 1995 Compliance Initiatives (GAO/GGD-97-158, August, 1997)
[2] Ibid.
[3] IRS Data Book for 2004, Table 1, http://www.irs.gov/pub/irs-soi/04db01co.xls
[4] Prepared Remarks of IRS Commissioner Mark W. Everson, November 18, 2004, http://www.irs.gov/newsroom/article/0,,id=131282,00.html

[5] An assessment is the posting to the taxpayer's account of an amount determined to be owed.

meaning the government has a legal, defensible claim to the money. Any monies received relating to the assessment, even though they may have been received in the intervening years, are reported as enforcement revenue in the year of the assessment. Any monies received in years after the year of the assessment and closing would be reported in those subsequent years. This can be several years after the actual enforcement action, in this case the audit, has been completed. The decision on the timing to report the revenue hinges on the establishment of an unfettered right to the monies, in accordance with generally accepted accounting principles.

Conspicuous by its absence from the list of activities generating enforcement revenue is our Criminal Investigation (C.I.) function. While the prosecution of criminal tax charges usually involves some amount of tax – there has to be a quantifiable harm to the government in order to bring a case to trial – any monies resulting from the action are attributed to the enforcement area generating them: examination, collection, under reporter, or Appeals.

The connection drawn between enforcement resources and revenue above should not be lost on the reader. As the federal agency charged with administering the Internal Revenue Code, the primary responsibility of the Service is to address compliance with the tax laws. Revenue resulting from those actions is a result of compliance responsibility, but is nonetheless important, especially from a return on investment (ROI) perspective. Significant resources are involved in enforcing the tax laws. For FY '04, approximately 40% of the IRS budget was dedicated to tax law enforcement.[6]

Direct Enforcement Revenue – a Historic Context

During the early 1980's it became apparent to IRS executives and especially to members of our oversight Committees that we needed a consistent definition of enforcement revenue and a reliable means of identifying and tracking it. Budget hearing testimony would see executives from different enforcement functions providing dramatically different and confusing numbers. Examination officials would say that IRS had made recommendations amounting to $40 - $45 billion a year; Collection officials would say the Service had collected $20 - $25 billion per year; Appeals executives would indicate that their staffs had handled $45 - $50 billion worth of unagreed issues each year.

Each was technically correct, but all left stakeholders asking for better, less confusing information. So, in the latter part of the decade the Commissioner asked that a system be put in place to identify and track direct enforcement revenue that would allow IRS to say with certainty what it was for any given year. This led to the development of the Enforcement Revenue Information System (ERIS – nee Enforcement Management Information System – EMIS).

ERIS came on line with examination and collection information in 1992. Information reporting (under reporting) data was included in 1994 and math error adjustments were incorporated in 1999.

Tracking Direct Enforcement Revenue – the Enforcement Revenue Information System (ERIS)

The system that IRS uses to track direct enforcement revenue is ERIS, a large data warehouse that combines disparate information from a number of different IRS data sources to identify and capture the results of enforcement activities. It then uses business rules to do three seemingly simple things:

1. determine the amount and timing of direct enforcement revenue,
2. attribute that revenue to the activity that generated it, and
3. avoid double counting.

To determine the amount of direct enforcement revenue we first identify the case as an enforcement case. This is done by analyzing the codes used when transactions are posted to the Masterfile. The Masterfile is the IRS database that stores various types of official taxpayer account information. It includes individual, business, employee plans and exempt organizations data. When any one of a number of transaction codes is present, the case is flagged as an enforcement case and all information about that module is extracted and sent to ERIS.

At the Masterfile level we aggregate data about taxpayers in several distinct ways, but two of the

[6] IRS Budget in Brief, http://cfo.fin.irs.gov/SPB/BudgetFormulation/docs/bib_05.pdf

major ways are referred to as tax modules and tax entities. A module is the data for one taxpayer for one type of return for a specific tax period. For example, your individual income tax return (Form 1040) for last year (2004) creates a tax module of data.

The second is as a tax entity. This could be as uncomplicated as your individual income tax return data for all available tax years, not just one year, or as complicated as all information for all types of tax for all available years for one taxpayer. This level of detail can get complicated quickly because a taxpayer can be liable for several different types of tax (income, employment, excise, etc.) for any number of years. For example, if you had a small business – a sole proprietorship - you would file a Form 1040 Schedule C – Profit or Loss From Business as part of your return. If this business had employees, you would also be required to file employment tax forms 940 - Employer's Annual Federal Unemployment (FUTA) Tax Return, and 941 - Employer's Quarterly Federal Tax Return. If the business was a small trucking firm you might also be liable for a Form 2290 - Heavy Vehicle Use Tax Return, an excise tax. Your tax entity information would include data on all types of tax for all years, significantly more information than just your tax module.

ERIS receives and analyzes data at the module level of detail because it is the smallest chunk of information. Modules can be aggregated to entities, but it can be very difficult to break an entity into the component modules in our data systems.

Once a module is identified as an enforcement module, any subsequent activity on it causes it to be a part of our monthly extract so it can be re-analyzed. Because of certain operations of tax law, such as net operating loss and tax credit carryovers (back and forward) and court litigation, it can be a number of years between activities on a module. To determine enforcement revenue accurately, we keep prior year information indefinitely. This multi-year effect and our need to re-analyze a module with new information, creates what we refer to as a dynamic system. The implications are that once we get information on an enforcement module, we have to keep it forever, so our dataset must, by necessity, grow ever larger. There are obvious storage and processing time implications associated with this that we are now analyzing. We are exploring near-line archiving of older, inactive modules to mitigate the impact, but no decision has been made as of this date.

Although some Masterfile transaction codes are only informational, most of them represent either debits (increases) or credits (decreases) to the taxpayer's account. Determining the amount of enforcement revenue associated with each module is done by adding up these debits and credits and analyzing when they happened in relation to other activity on the module, based on business rules.

Timing direct enforcement revenue uses the closing date of the enforcement activity. Different functions define a closed action differently, so, although ERIS receives information on enforcement actions as soon as the case is flagged, it does not report revenue out until the function making the assessment posts it as a closed case.

The next step is to attribute the revenue to the function (Examination, Appeals, Collection, and Information Reporting) that generated it. One may ask why is it important to do this since our external stakeholders are mostly interested in knowing only what the total enforcement revenue for the year is. The first reason is because being able to attribute the revenue functionally allows us to estimate the potential yield associated with hiring initiatives that are functionally-specific. We'll say more about this below under Revenue Estimation Methodologies. A second reason is that, even though we do not use enforcement revenue as a performance metric for employees, case-specific enforcement revenue is a critical piece of information used in at least one of our workload selection models under development. Being able to optimize the yield associated with various workload mixes is only good management in these ROI-centric times. Lastly, any time you have this type of information, its human nature for people to want to know what their piece of the pie amounts to.

The general business rule for attributing enforcement revenue is that the function generating the assessment gets credit for any monies received because of it.

Finally, there is a need to avoid double counting enforcement revenue. Double counting can occur because the same enforcement case will most likely be worked by more than one enforcement function. For example, if Examination completes an audit and the taxpayer owes an additional $100 in tax but doesn't pay at that time, the collection of that tax would be handled by Collection. If we were to add the $100 from Examination activity to the $100 from Collection activity, enforcement revenue might appear to be $200, when it is only $100. The rules of attribution explained above address most of these possible double count issues, but there are also a number of rules to cover some very unusual circumstances.

ERIS currently runs on an IBM-390 mainframe at our Detroit Computing Center. Monthly extracts of the Masterfile data are combined with extracts from five other systems to populate our data warehouse. A common key structure composed of the Taxpayer Identification Number (TIN – a 9 character field), the Masterfile Tax Account Code (type of tax – income, employment, excise, etc. – MFT, a 2 digit field) and the Tax Period (a 6 digit field in YYYYMM format) is available in each data source and, collectively, defines a tax module (see above). The other five systems are:

1. IMAC - Individual Midwest Automated Compliance system – containing key information and as many as 728 elements of what was reported on individual tax returns as filed.
2. BMAC - Business Midwest Automated Compliance system – containing key information and data on 15 different business tax returns (i.e., Corporation – Form 1120, Partnership – Form 1065, Gift – Form 709, Estate – Form 1041, etc.) as filed.
3. AIMS – Audit Information Management System – containing key information and case inventory information on returns being audited.
4. IRPCA - Information Reporting Program Case Analysis system – containing key information and case inventory information on the matching of information reporting documents to filed returns.
5. ICS – Integrated Collection System – containing key information and case inventory information on Collection cases.

A large portion of monthly processing time is devoted to formatting the data from the different systems to look like our most common record, an individual Masterfile record. When the formatting is done, it has to be sorted and analyzed to produce our monthly output; currently over 4.1 billion records with almost 1.2 terabytes of data. However, in order to accommodate the 3 business rules mentioned above, the processing works through much of this data as many as four times in one full cycle. See Attachment A for a high-level schematic of the data sources and process used.

The monthly output consists of both hardcopy reports and selected electronic data subsets. Both contain summary information that eliminates specific taxpayer identification. Some focus on overall direct enforcement revenue and others provide function-specific views (Examination, Appeals, Collection, Information Reporting/Under Reporting) of enforcement revenue and activity. See Attachments B through F for some examples of the type of information available.

Revenue Estimation Methodologies

IRS has been estimating the direct enforcement revenue expected to be generated by hiring initiatives for over 20 years, dating back to the 1985 Initiative. A methodology for estimating revenue from hiring initiatives using ERIS data was developed as a justification for funding the 1995 Initiative. This methodology developed as a prelude to the 1995 Initiative forms the basis of the current methodology. GAO reviewed it, saying it "represented a significant improvement over past methodologies."[7]

The current methodology used for hiring initiative revenue estimates relies heavily on two pieces of information. They are the anticipated number of FTE hires for a particular type of enforcement work and the historic yield rates for that particular type of work. Yield rates are developed with operational data (e.g. audit recommendations, assessments of additional tax, collections of tax, FTEs, etc.) as the starting point. All of the enforcement dollars provided by the Operating Divisions (Wage & Investment, Small Business/ Self-/Employed, Large & Mid-sized Business, Tax Exempt & Government Entities) are converted to collected dollars and then tied back to ERIS data - Total Enforcement Revenue Collected. These dollars collected are then used to compute historic yield rates by function within an Operating Division.

In general, the enforcement revenue estimates assume that the average yield (revenue collected) per FTE is the same for new hires as for current employees. The revenue for new hires is then

[7] Tax Administration: Assessment of IRS' Report on its Fiscal Year 1995 Compliance Initiatives (GAO/GGD-97-158, August, 1997)

discounted based on several factors including the following: assumed lower productivity of new employees (less experience, less training), lower average value of cases available for new employees, time new employees spend in training, lost revenue associated with taking experienced personnel off casework to instruct the new hires and the 'enter on duty date' of the new hires. For example, the 1995 enforcement initiative budget submission included a funding request of $145 M (to hire an additional 4,660 revenue producing FTE – primarily revenue agents and officers) was estimated to produce $9.6 billion over 5 years.[8]

Discussions regarding updating the estimation methodology started during the FY06 Budget Process. They included access to and better understanding of data available at higher levels of ERIS summary databases (DB2 tables) and various studies of ERIS data including the flow of collected dollars. Together, these have resulted in a re-thinking of the revenue estimation model, and the resulting new methodology for estimating revenue from hiring initiatives is under development and will be used when estimating revenue for future hiring initiatives. It relies on ERIS as the sole source of revenue data used in the estimation process and considers collection streams in each fiscal year of a hiring initiative in arriving at estimated revenue collected. The new model will still consider the factors used to discount revenue in the old model.

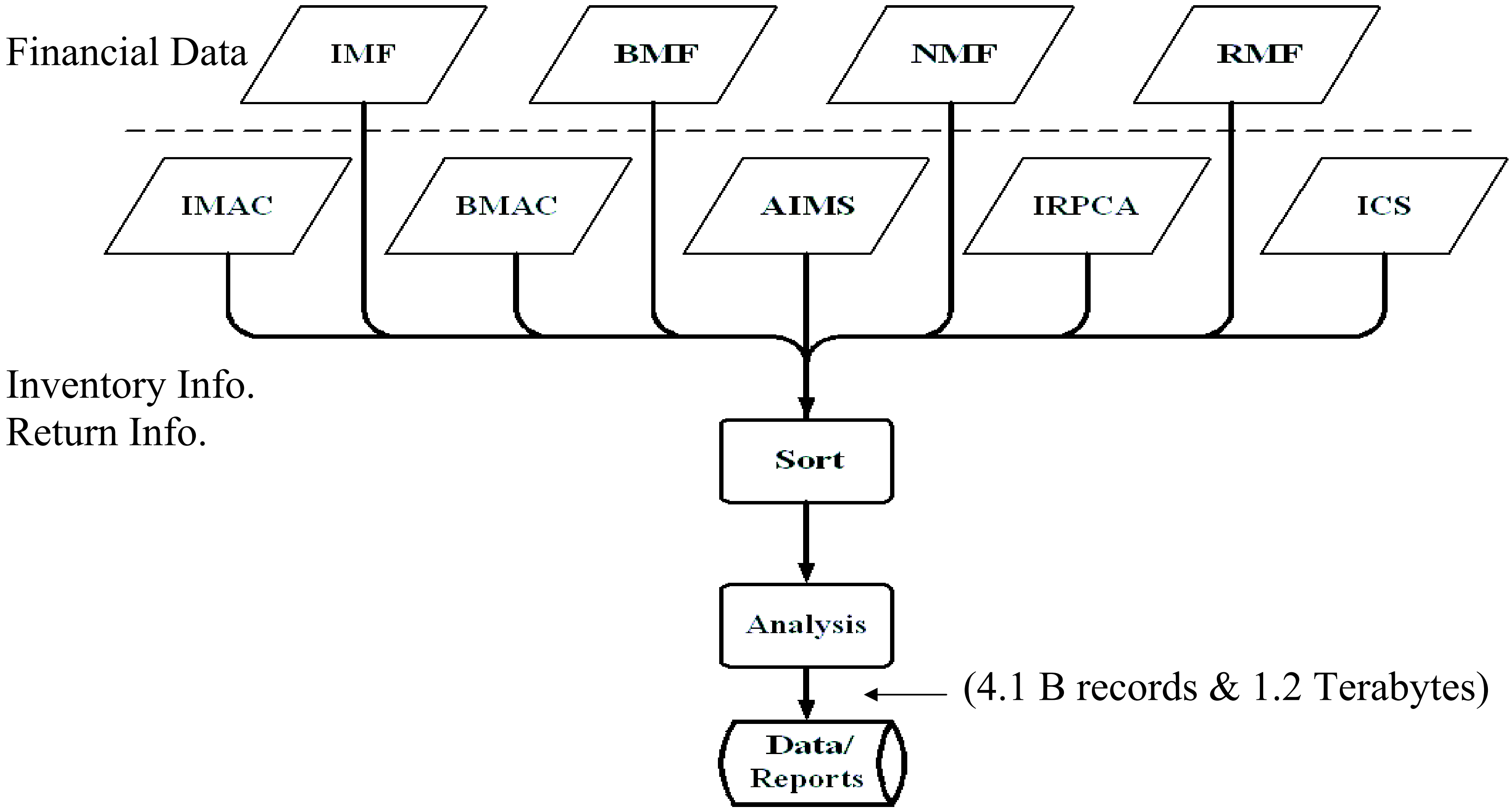
Revenue Projections

In addition to initiative estimations, ERIS revenue information is used to project direct enforcement revenue twice a year for IRS internal use by the Commissioner. It is done early in the fiscal year and just after the mid-year data is available. The method used for revenue projections involves deriving yield per FTE data using Total Enforcement Revenue Collected and FTE levels from the most recently completed fiscal year. The computed yield per FTE data is used in conjunction with projected FTE levels (for enforcement functions only) in future years to project enforcement revenue collected for those future years. For example, both the early and mid-year projections for FY '03 generated projections of about $37 billion, and the actual enforcement revenue for the year turned out to be $37.6 billion.

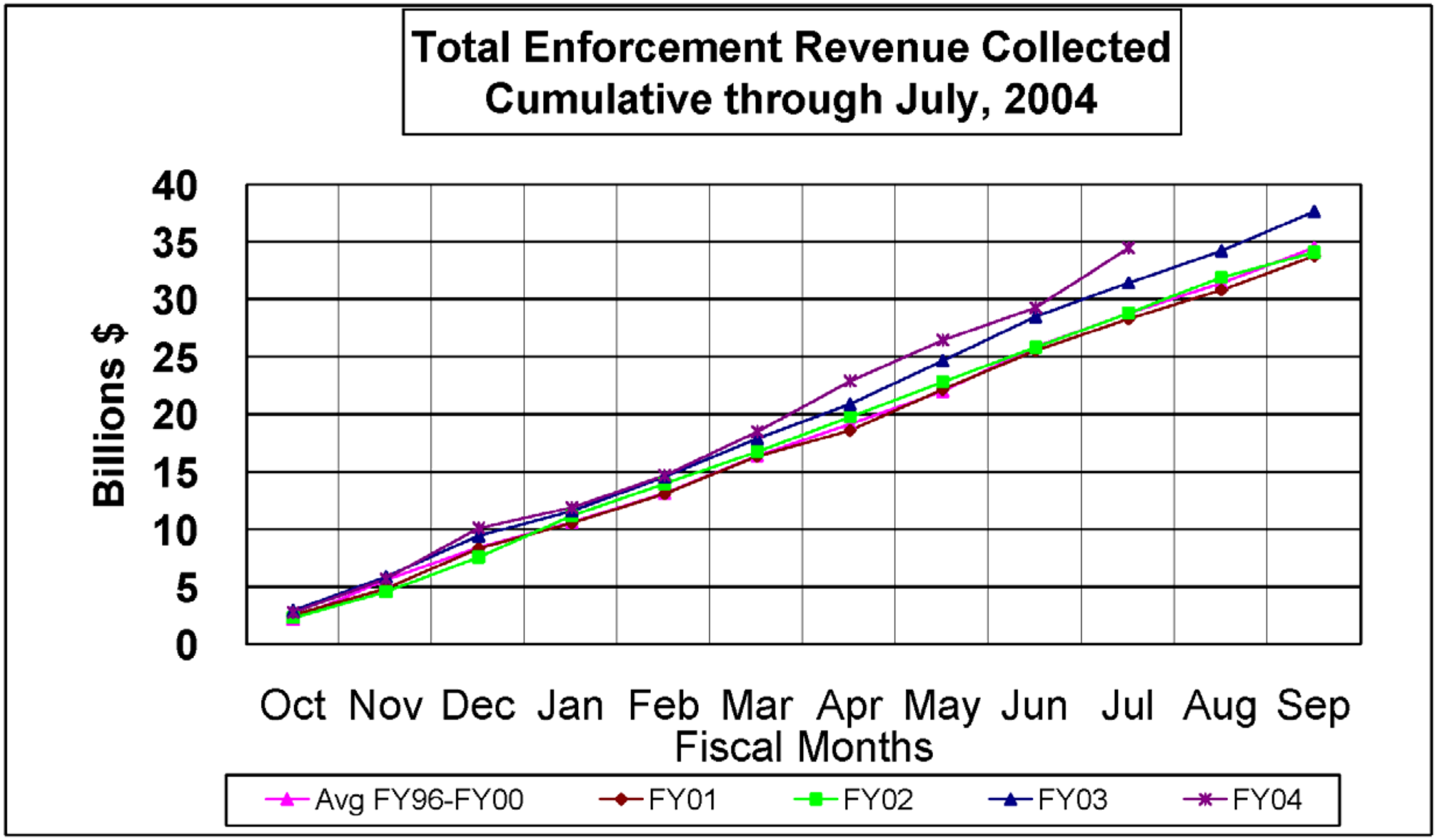
Note:

The views expressed in this article represent the observations and conclusions of the authors. They do not necessarily represent the opinions of the Internal Revenue Service.

[8] Ibid.

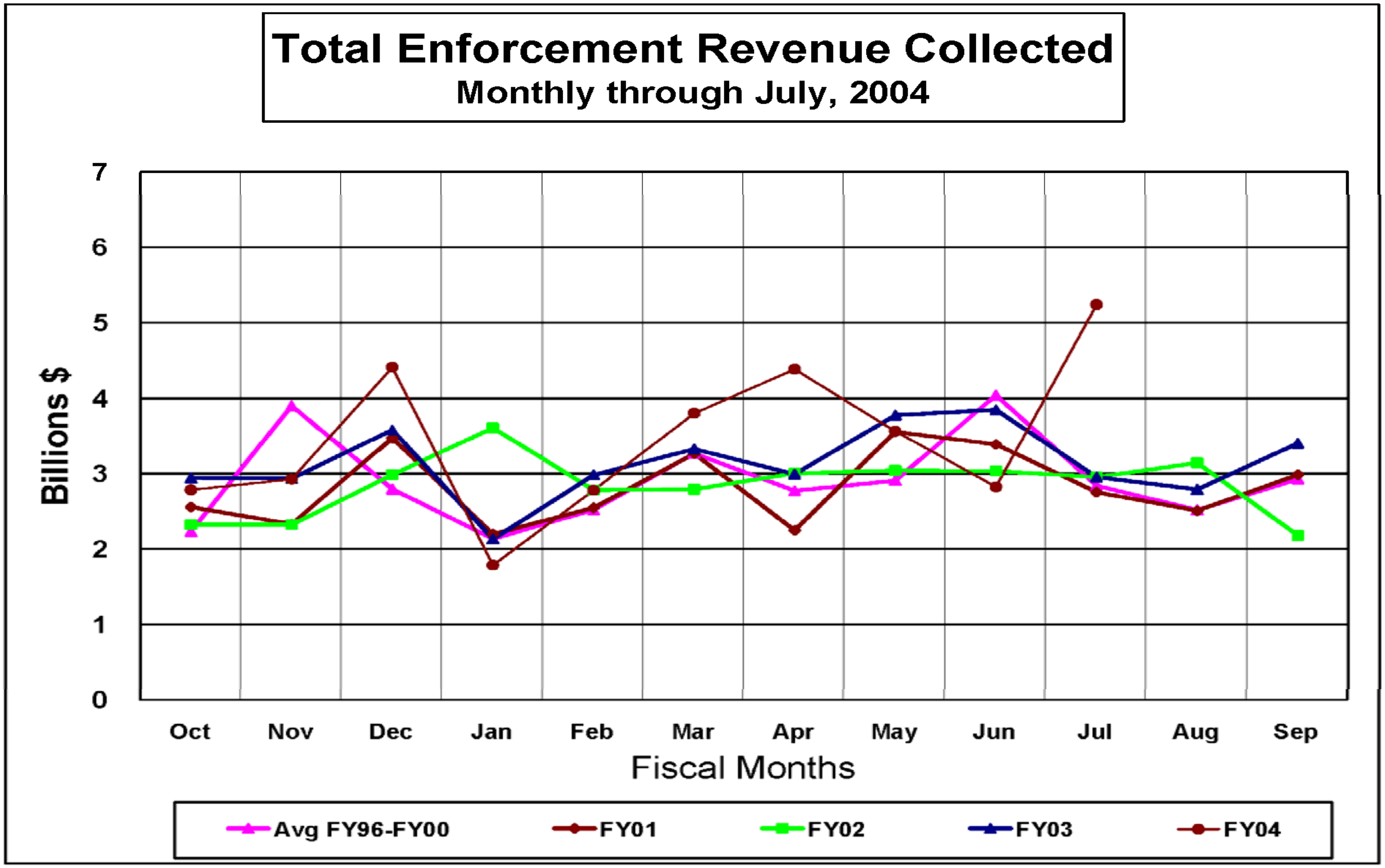


Attachment A



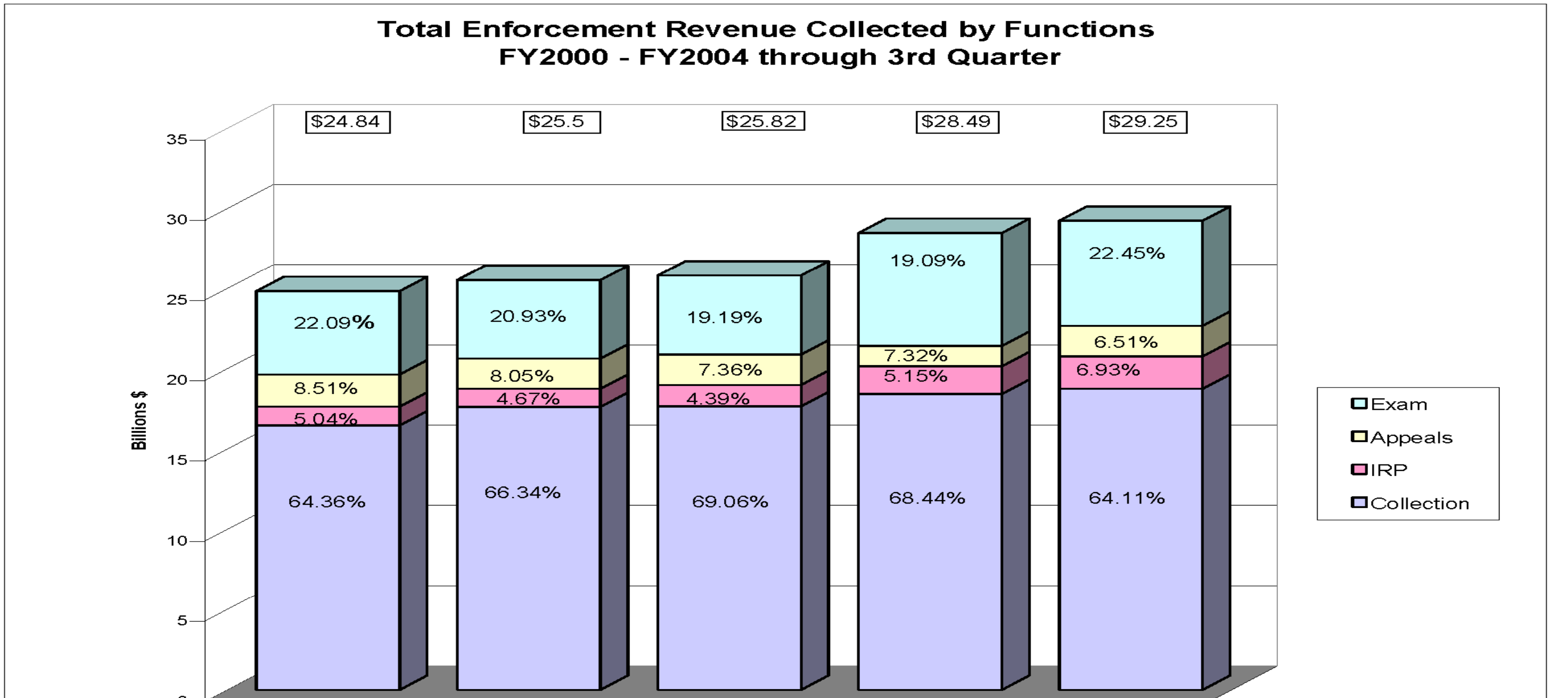
Source: Enforcement Revenue Data
Contact: 202.874.0499

Attachment B



Source: Enforcement Revenue Data
Contact: 202.874.0499

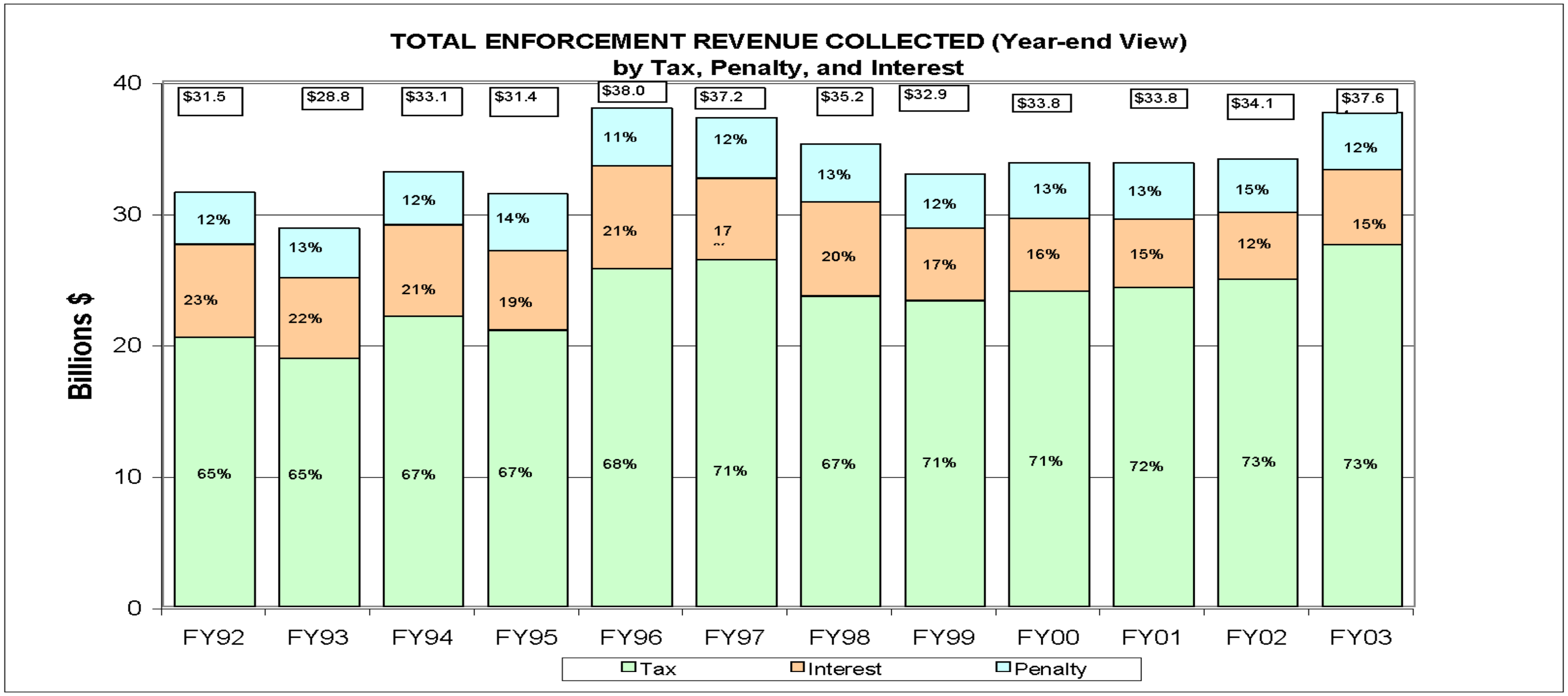
Attachment C



	FY00	FY01	FY02	FY03	FY04
Exam	$5.20	$4.90	$4.93	$7.06	$6.57
Appeals	$2.00	$1.88	$1.89	$1.27	$1.91
IRP	$1.16	$1.12	$1.33	$1.72	$2.03
Collection	$16.48	$17.63	$17.67	$18.44	$18.75

Source:Enforcement Revenue Data
Contact: 202.874.0499

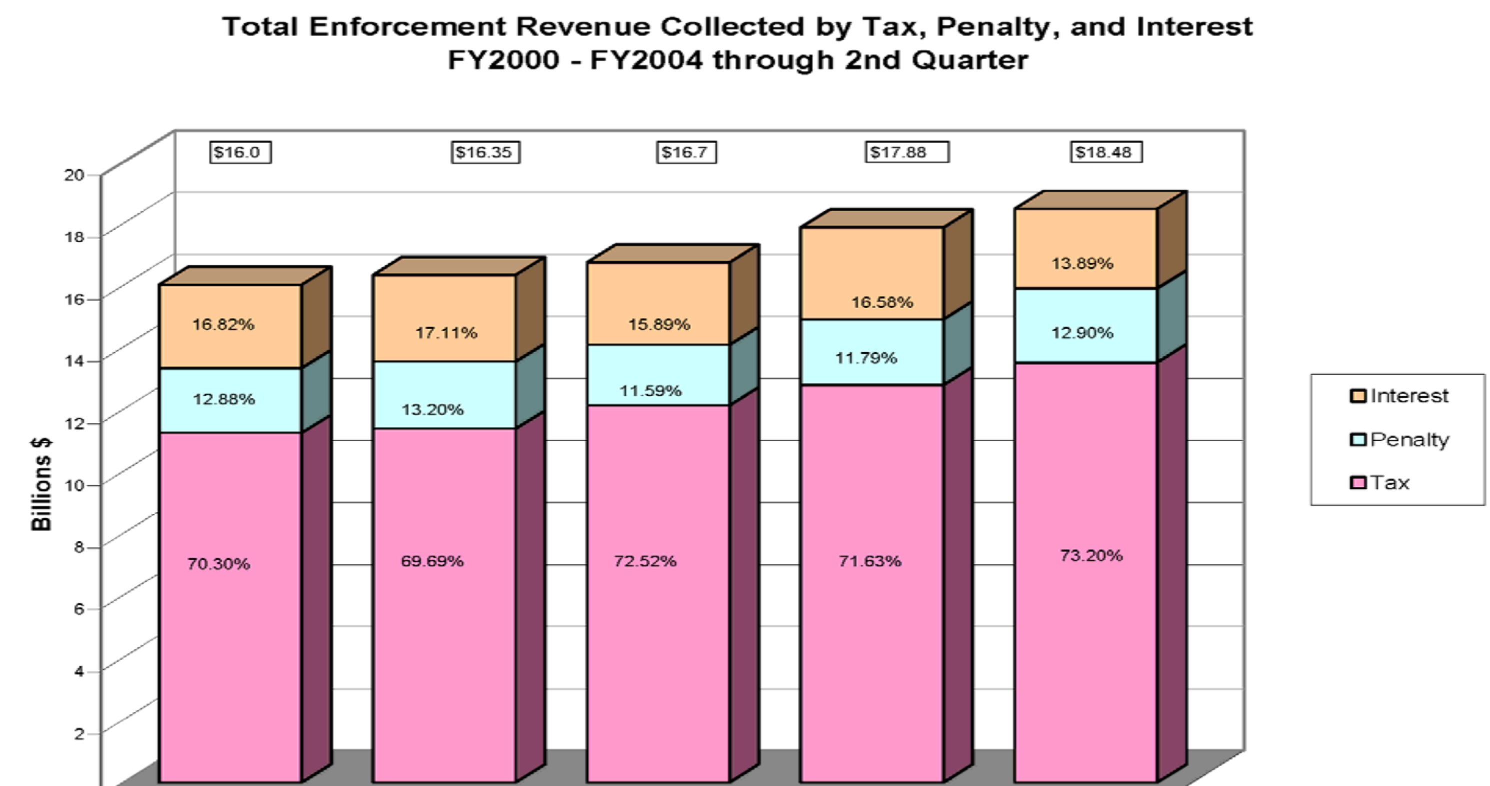
Attachment D



	FY92	FY93	FY94	FY95	FY96	FY97	FY98	FY99	FY00	FY01	FY02	FY03
Interest	$7.132	$6.182	$7.018	$6.092	$7.844	$6.194	$7.168	$5.525	$5.520	$5.193	$5.100	$5.744
Penalty	$3.963	$3.795	$4.036	$4.330	$4.419	$4.633	$4.402	$4.126	$4.269	$4.286	$4.100	$4.355
Tax	$20.444	$18.841	$22.042	$21.009	$25.701	$26.418	$23.630	$23.287	$24.011	$24.300	$24.900	$27.535

Source: Enforcement Revenue Data
Contact: 202.874.0437

Attachment E



	FY00	FY01	FY02	FY03	FY04
Interest	2.70	2.80	2.66	2.96	2.57
Penalty	2.07	2.16	1.94	2.11	2.39
Tax	11.27	11.39	12.15	12.81	13.53

Note: Due to rounding, percentage may not sum to 100%.

Source: Enforcement Revenue Data
Contact: 202.874.0499

Attachment F

This page intentionally left blank

Geographically Optimizing IRS Tax Education Outreach Staffing

Jay Schnapp and Ronald Deaett
Internal Revenue Service, U.S. Department of the Treasury

The views expressed in this article represent the opinions and conclusions of the authors; they do not represent the opinion of the Internal Revenue Service.

Introduction

The Internal Revenue Service (IRS) seeks to ensure tax compliance through a balanced approach involving taxpayer service as well as enforcement. The job of the IRS Stakeholder, Partnership, Education and Communication (SPEC) function is to conduct educational outreach that enables taxpayers to understand and meet their tax obligations "upfront," thereby reducing the number of IRS contacts downstream. To accomplish its mission, SPEC develops partnerships with key external stakeholders (such as trade associations) who can help inform the public. However, a major challenge for IRS is determining where, geographically, to allocate SPEC employees. The model we present in this paper provides a method that allows SPEC to distribute their staff based on eight factors. The model assigns scores to each IRS geographic territory to determine where the next staff year should go.

Background

In 2001, the IRS underwent a major reorganization. The agency transformed from an organization arranged along geographic lines to one organized along taxpayer segments. This reorganization grouped the major functions of the IRS under four main business operating divisions, listed below.

Wage and Investment – Serving individual taxpayers with income primarily from wages and investments

Small Business/Self Employed – Serving individual taxpayers filing returns with farming or small business income and other entities (corporations, partnerships) with assets less than $10 million

Large and Midsize Business – Serving large entities (corporations, partnerships) with assets over $10 million

Tax Exempt/Government Entities – Serving non-profit entities, pension plans and governments

Each of these organizations provides services to taxpayers by addressing each of three main return filing phases:

Pre-Filing (Outreach, Education, Tax Assistance)

Filing (Submission Processing/Returns Processing)

Post-Filing (Compliance [e.g. Examination and Collection])

When addressing taxpayer compliance issues, the most costly and time-consuming is the post-filing phase. It is much less costly to the IRS for SPEC to ensure taxpayers file correctly through pre-filing outreach than it is for the Examination and Collection functions to correct it after the fact.

Prior to 2001, each of the 63 IRS geographic districts had a Taxpayer Education Coordinator to address pre-filing issues. This person provided presentations to the public and external organizations and training to individuals to create neighborhood "Volunteer Income Tax Assistance (VITA)" sites. (Educational institutions sponsored VITA sites for their local communities, as well.)

The reorganization in 2001 created SPEC as the successor to the Taxpayer Education Coordinator. IRS planned to expand staffing from 100 to 800, nationwide, and to be a more pro-active force in the communities. The SPEC organization changed education and outreach by moving from a retail approach (assisting taxpayers one by one) to a wholesale approach (recruiting municipalities, anti-poverty agencies, other non-profits and corporations) to form coalitions to provide tax assistance. SPEC was organized into seven different Area offices nationwide. Aligned under these SPEC Area offices were forty-eight Territory offices that handle implementation of the outreach and education efforts.

Subsequently, budget realities have curtailed the expected hiring of staff for the SPEC organization at around 400 technical employees, instead of 800. This resulted in unbalanced staff in each SPEC post of duty and inefficiencies in handling the SPEC workload across the country.

SPEC recognized this staffing issue and engaged IRS's Wage and Investment Research division to develop a tool they could use to 1) measure their current workload and 2) determine where best to allocate their next available staff year.

Model Development

Ideally, SPEC would like to measure their activities' effect on voluntary compliance, and distribute staffing in a way that would maximize impact. Several difficulties arise with the practical application of this approach. It is difficult to measure voluntary compliance; and hence it is difficult to quantify the impact of the activities that SPEC undertakes on voluntary compliance. (although other studies are in process to accomplish this task) Additionally, moving staff that is currently in place to other locations is cost prohibitive. The result is SPEC had no data-driven approach to decide the staffing level - going forward - in each SPEC Territory. SPEC management recognized this and selected eight factors to use as proxies for workload to help them make decisions based on the data available.

Data

The eight SPEC-identified factors are presented below. We divided them into three categories: internal measures, population/returns, and geography. Internal measures include items that SPEC can affect directly through outreach and coalition-building work. Population and returns contains three major customer groups SPEC is charged with serving: the elderly, the limited English proficient, and the low income. Geography is the third group, for which size (in square miles for each SPEC Territory) is used. The list below shows the workload factors, by category. Appendix A presents the information in more detail.

Internal Measures

E-Filed Returns Goal
Volunteer Prepared Returns Workload
Paid Preparers V-Coding
Total Weighted Partners

Population and Return Filing

Total Elderly Population
Total Households with Limited English Proficiency (LEP)
Low Income Returns

Geography

Total Square Miles

During project development, SPEC stipulated the following model requirements:

Model should use all eight workload factors.

Model users must have the ability to change the weight assigned to each of the eight workload factors (to handle changing priorities).

Rankings of SPEC Territories should be relative to one another and not just ordinal in nature.

No "black box" approach -- i.e., the model clearly should demonstrate the method used to determine placement of the next staff year. Additionally, users of the model and those affected by the model output must understand the reasoning and method used in the model.

After considering and discussing several alternative methods with SPEC, including multiple linear regression and optimization, the best representation of SPEC needs was a scoring system to rank SPEC Territories based on current staff and the eight workload factors.

Scoring Method

We gathered data for the eight workload factors, and developed a scoring system. The scoring system consists of one overall score for each SPEC Territory based on the eight factors.

Calculating the Individual Workload Factor Scores

Calculating the score for each of the eight workload factors began by determining a coverage ratio for each SPEC Territory. The coverage ratio represents the amount of work required of each current SPEC technical staff member, by factor. The coverage ratio calculation is determined by dividing the raw number measure for each factor by the current technical staff in place at each SPEC Territory.

$$\textit{Coverage Ratio}_f = \textit{Raw Number}_f \,/\, \textit{Current Staff}$$

Where:

Raw Number = Return Count, Paid Preparer Count, Square Miles, Weighted Partners or Population (depending on the workload factor)

f = workload factor

The model ranks the SPEC Territories from highest to lowest. Each SPEC Territory receives a score between 0 and 100. The individual workload factor scores are calculated by two different methods: one-sided and two-sided.

One-Sided Score

The one-sided score applies to the following six workload factors: E-Filing Returns Goal, Paid Preparers V-Coding, Total Elderly Population, Total Population with Limited English Proficiency, Low Income Returns, and Total Square Miles.

For these six workload factors, each SPEC Territory receives a rating of between 0 and 100. The highest coverage ratio receives a score of 100 and the lowest a score of 0. The model assigns a relative score to the SPEC Territories with a coverage ratio between the highest and lowest, based on the distance they are from the minimum value. The calculation is as follows:

Coverage Ratio $Score_f = (CR_f - CR_f(minimum))/CR_f(range)$

Where:

CR = coverage ratio

f = workload factor

This method of calculating the score meets the SPEC requirement. That is, SPEC Territories, with about the same coverage ratio for the workload factor, receive about the same score.[9]

Two-Sided Score

The two-sided score applies to the remaining two workload factors: Volunteer Prepared Return Workload and Total Partners. SPEC management requested this scoring method. The logic behind the decision was that if a SPEC Territory had extreme values, the SPEC Territory needed extra staff to either 1) support the additional workload (high value), or 2) increase the activity in this factor (low value).

For these two workload factors, each SPEC Territory receives a rating of between 0 and 100. The median value of the coverage ratio receives a score of 0[10]**.** The highest and lowest values receive a score of 100. The SPEC Territories between the highest and the median are assigned a relative score based on the distance they are from the median. The model assigns a relative score to the SPEC Territories between median and lowest coverage ratio based on the distance they are from the minimum value. The calculation is below.

Values above the median

Coverage Ratio $Score_f = (CR_f - CR_f(median))/(CR_f(maximum) - CR_f(median))$

Values below the median

Coverage Ratio $Score_f = (CR_f(median) - CR_f))/(CR_f(median) - CR_f(minimum))$

Where:

CR = coverage ratio

f = workload factor

Calculating the Overall Score

The model calculates the overall score by adding together a weighted score from each of the individual eight workload factors.

Overall Score $= SUM \{(WGT_{f1\text{-}8} * S_{f1\text{-}8})\}$

Where:

WGT = workload factor weight

S = Coverage Ratio Score

f = workload factor

Weights

The customer provides the weights for each of the eight workload factors at the time the model is run. The user can change weights to provide flexibility in using the model as SPEC priorities change over time. However, the model provides a default weighting system, as listed in Table 1.

Table 1 – Default Weights

Workload Category	Default Weights
E-Filing Returns Goal	5%
Volunteered Prepared Return Workload	10%
Paid Preparers V-Coding	5%
Total Partners	15%
Total Elderly Population	15%
Total Population with Limited English Proficiency	10%
Low Income Returns	35%
Total Square Miles	5%

Results

The result of this work is a practical, user-friendly tool. The model provides an overall score for each SPEC Territory, based on the eight weighted workload

[9] The model originally used a simple ranking method as a scoring system, with the highest value receiving a score of 48 and the lowest a score of 1. We discarded this method as it produced too wide a range of scores for SPEC Territories that have about the same value for a coverage ratio for each workload factor.

[10] Due to technical limitations in the final programming, the median value gets a score that is only close to zero.

factors. It provides a mechanism for determining which SPEC Territory is most in need of additional staff. The SPEC Territory that receives the highest overall score is the location that is most in need of additional staff.

Sample reports produced by the model are in Appendix B. The reports list the SPEC Territory with the highest overall score first. They show the eight workload factor scores and the applicable coverage ratios. The particular sample report shows that based on technical staffing currently in the model, Chicago is the territory most in need of staff.

Final Product

We packaged the final model into a customized Microsoft Access database application. The customized application allows for simple push-button access to the model. This lets SPEC management run reports, add and subtract staffing to SPEC Territories, weight the eight workload factors, and determine the next SPEC Territory most in need of the staff year.

The customized Microsoft Access application makes the model dynamic. The interface allows SPEC management to update the SPEC Territory staffing figures used in the model when needed. Once the user updates the staff in the model, the user may run the model again; and it will calculate the new overall score for each SPEC Territory based on the new staffing allocation.

Appendix C contains sample screen prints of the application interface.

Conclusion

The model developed here supports the IRS SPEC organization by providing a data-driven, workload-based approach to determining the Territory most in need of additional staffing. The model is a tool for SPEC management. SPEC management uses the tool, in conjunction with other factors, to help make decisions about where staff should be located. Other factors include IRS Commissioner emphasis, Treasury Directives, and political considerations.

The model is in use by SPEC and is flexible enough to meet most of the staffing needs. SPEC management recently used the model to help with the reorganization of SPEC into a smaller number of Areas (4) and Territories (42).

Wage and Investment Research continues to support SPEC and the model by updating data and geographic locations.

Acknowledgements

Special thanks go to Javier Framinan, Sheldon Schwartz, and John Yakovich for contributions and input on the project and paper.

Appendix A – Workload Factors, Definitions and Data Sources

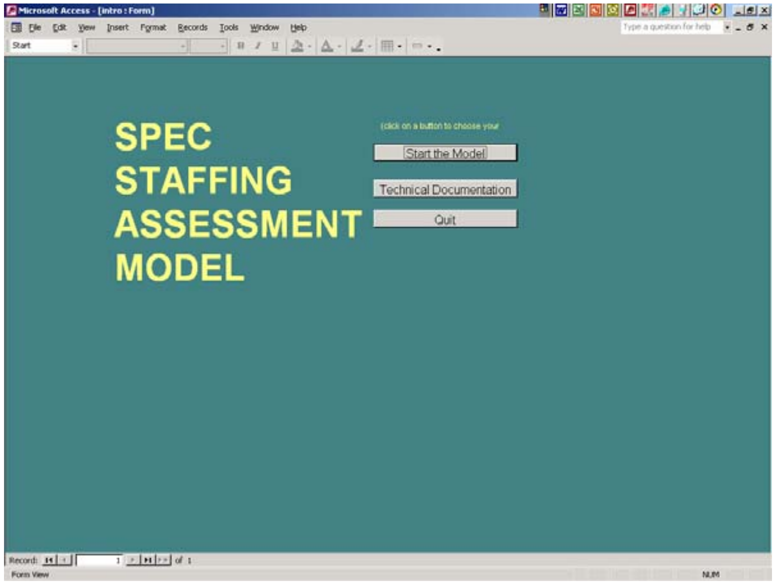
Workload Factor	Category	Measure	Data Source	Definition
E-Filing Returns Goal [a]	Internal	# of Returns	Tax Year 2002 Returns	Calculated as the following: 80% of the number of Form1040 returns filed in the SPEC Territory LESS the actual number of E-Filed returns in the SPEC Territory
Volunteer Prepared Returns Workload [b]	Internal	Population	Tax Year 2002 Returns & Census 2000	Calculated as the following: 7.5% of the sum of (Households with a resident aged less than 65 with income less than $35K + Households with resident aged 65 and over + Population in military quarters) LESS the actual number of volunteer prepared returns in the SPEC Territory
Paid Preparers V-Coding [c]	Internal	# of Paid Preparers	Tax Year 2002 Returns	Paid Preparers filing 36% or more of their total prepared returns on paper
Total Partners [d]	Internal	Weighted # of Partners	SPEC Time and Reporting System (STARS)	Calculated as the sum of the following: Sum of (Number of coalitions * weight of 3 + Number of partners who are coalition members * weight of 2 + Number of partners not coalition members * weight of 1)
Total Elderly Population	Population	Population	Census 2003 Population Estimate	Population aged 60 plus
Total Households with Limited English Proficiency	Population	Households	Census 2000	Households where no person over age 14 speaks English "very well" (linguistically isolated)
Low Income Returns	Population	# of Returns	Tax Year 2002 Returns	Returns with income less than the Earned Income Tax Credit limit ($34,178 for Tax Year 2002)
Total Square Miles	Geography	Square Miles	Census 2000	Total Square Miles of land in the SPEC Territory

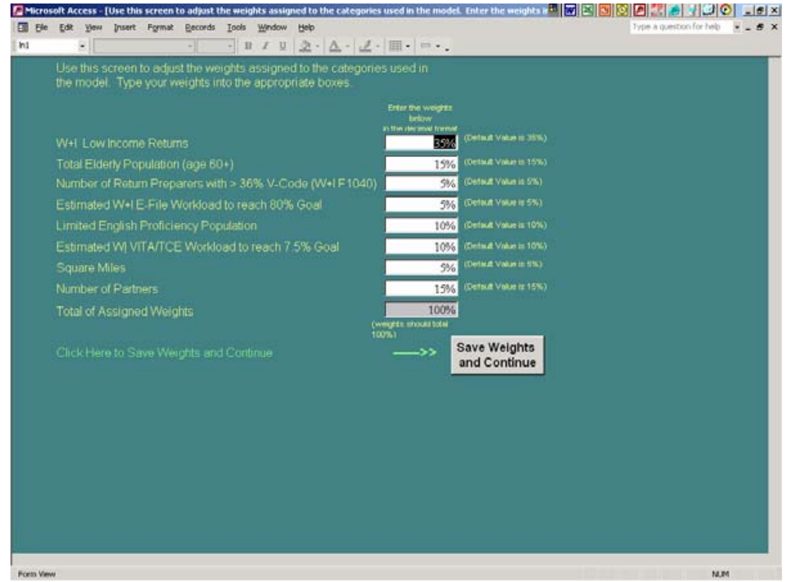
a = The 80% figure is a Congressionally mandated E-Filing goal for the IRS.
b = The 7.5% percentage was used so that all SPEC Territories would be below the result, so that final calculations show a deficit that SPEC can strive to close. The percentage listed does not represent an actual goal for SPEC. Persons in the military are a large consumer of volunteer return preparation services.
c = Counting only paper returns that originally were prepared on a computer, printed and filed by mail rather than E-Filed.
d =Coalitions are a group of agencies, local government and/or non-profit organizations having a group identity and activity in at least one of three strategic areas: tax preparation, outreach, or asset-building. Partners are individual organizations that work independently in these three strategic areas.

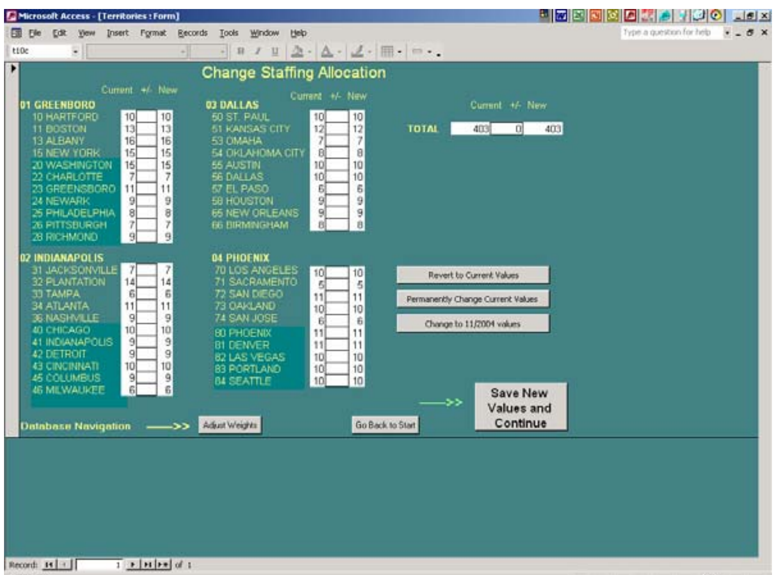
Appendix B – Sample SPEC Territory

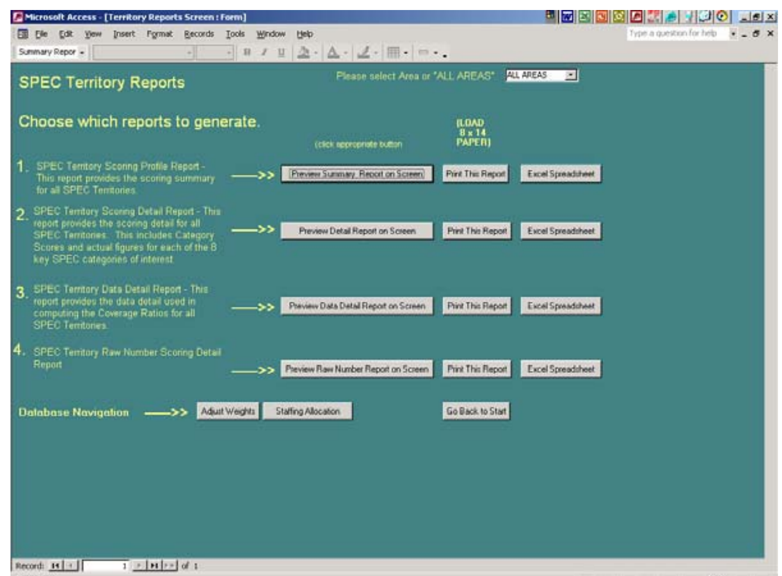
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	SPEC STAFFING ASSESSMENT MODEL - SPEC Territory Scoring Report																			
2					SPEC Market Segments															
3					Total W&I Low Income Returns		Total Elderly Population(Aged 60 Plus)		Number of Paid Preparers Y-Coding >-36% F1040 W&I		Estimated W&I E-File Workload to Reach 80% Goal		Total Population with Limited English Proficiency		Estimated VITA & TCE Workload to Reach 7.5% Goal		Total Square Miles		Total Weighted Partners	
4					(WEIGHT: 35%)		(WEIGHT: 15%)		(WEIGHT: 5%)		(WEIGHT: 5%)		(WEIGHT: 10%)		(WEIGHT: 10%)		(WEIGHT: 5%)		(WEIGHT: 15%)	
5					(Source: TT 2002 Tax Returns)		(Source: Census 2000)		(Source: Return Preparer		(Source: TT 2002 Tax Returns)		(Source: Census 2000)		(Source: TT2000 Tax Returns)		(Source: Census 2000)		(Source: SPEC PAS System	
6	Territory	Technical Staff	Grades 12/13 Staff	Combined Coverage Score	Coverage Ratio	Coverage Ratio Score	Coverage Ratio	Coverage Ratio Score	Coverage Ratio	Coverage Ratio Score	Coverage Ratio	Coverage Ratio Score	Coverage Ratio	Coverage Ratio Score	Coverage Ratio	Coverage Ratio Score	Coverage Ratio	Coverage Ratio Score	Coverage Ratio	Coverage Ratio Score
7	(a)	(b)	(c)	(d)	(f)	(g)	(h)	(i)	(j)	(k)	(l)	(m)	(n)	(o)	(p)	(q)	(r)	(s)	(t)	(u)
8	CHICAGO, IL	10	2 / 3	67.09	246,414	100.00	199,761	87.67	217	77.52	141,771	89.12	21,673	27.33	9,758	54.99	5,558	8.42	10.0	13.04
9	DETROIT, MI	9	3 / 3	65.65	226,589	88.40	181,983	76.51	193	66.63	125,051	75.95	7,162	7.18	12,113	94.72	6,312	9.59	8.7	36.23
10	PHILADELPHIA, PA	8	2 / 2	65.54	241,983	97.41	205,603	91.34	205	72.21	155,589	100.00	9,550	10.50	10,611	69.38	2,578	3.79	11.6	6.36
11	TAMPA, FL	6	1 / 2	65.30	228,334	89.42	219,401	100.00	130	38.22	99,705	56.00	12,897	15.14	9,196	45.50	1,921	2.77	18.2	53.94
12	COLUMBUS, OH	9	2 / 3	59.93	231,525	91.29	176,445	73.04	176	59.26	129,410	79.39	5,233	4.50	10,082	60.45	3,942	5.91	13.8	22.02
13	PITTSBURG, PA	7	0 / 1	59.12	200,268	73.00	186,625	79.43	149	46.68	111,747	65.48	2,493	0.69	10,496	67.44	7,174	10.93	18.7	57.92
14	LOS ANGELES, CA	10	0 / 5	57.64	198,121	71.74	151,379	57.30	267	100.00	135,530	84.21	50,433	67.27	10,511	67.69	1,195	1.64	11.8	7.64
15	CHARLOTTE, NC	7	2 / 2	56.87	218,333	83.57	167,682	67.54	135	40.36	76,129	37.43	5,308	4.60	10,972	75.48	6,374	9.69	15.4	34.03
16	MILWAUKEE, WI	6	2 / 1	53.17	194,091	69.38	154,440	59.22	74	13.09	80,825	41.13	5,012	4.19	8,059	26.31	9,052	13.85	23.2	90.30
17	SEATTLE, WA	10	2 / 5	53.07	159,973	49.42	119,299	37.17	106	27.41	95,153	52.41	10,923	12.40	3,033	83.60	64,492	100.00	6.3	77.39
18	BIRMINGHAM, AL	8	2 / 0	51.33	198,970	72.24	156,377	60.44	112	30.06	66,821	30.10	2,270	0.38	12,425	100.00	12,206	18.75	13.5	20.00
19	NEW YORK, NY	15	3 / 4	49.10	141,664	38.71	133,578	46.13	181	61.44	117,469	69.98	33,217	43.36	8,528	34.22	141	0.00	5.3	95.36
20	SAN JOSE, CA	6	2 / 2	47.65	158,099	48.33	119,225	37.12	174	58.23	108,501	62.92	26,374	33.86	6,555	0.93	8,210	12.54	24.5	100.00
21	NEWARK, NJ	9	2 / 3	45.68	180,266	61.30	163,838	65.12	246	90.46	146,920	93.17	21,014	26.42	7,868	23.08	824	1.06	11.0	1.82
22	PLANTATION, FL	14	2 / 3	45.45	117,567	24.61	158,171	61.57	83	16.88	50,738	17.44	73,997	100.00	11,168	78.77	1,316	1.83	7.7	52.80
23	NEW ORLEANS, LA	9	0 / 2	40.64	169,739	55.14	133,411	46.03	97	23.50	59,143	24.06	4,481	3.46	9,566	51.75	10,626	16.29	8.6	38.16
24	SACRAMENTO, CA	5	1 / 1	34.93	130,157	31.98	113,045	33.24	120	33.78	78,598	39.38	11,836	13.67	5,999	12.09	8,922	13.65	21.6	78.91
25	SAN DIEGO, CA	11	0 / 4	34.82	153,962	45.90	119,217	37.12	183	62.14	99,600	55.91	21,632	27.28	5,773	17.53	3,311	4.93	13.1	17.02
26	NASHVILLE, TN	9	1 / 2	32.76	138,660	36.95	108,308	30.27	63	7.85	47,692	15.04	2,410	0.58	7,641	19.26	4,580	6.90	6.2	78.74
27	INDIANAPOLIS, IN	9	2 / 2	32.56	147,192	41.94	112,129	32.67	101	25.40	65,441	29.02	3,262	1.76	6,843	5.78	3,985	5.97	7.2	61.35
28	RICHMOND, VA	9	3 / 1	32.19	159,972	49.42	124,411	40.38	110	29.32	91,551	49.58	6,230	5.88	6,122	9.13	4,399	6.62	13.6	20.40
29	OMAHA, NE	7	1 / 1	31.78	132,428	33.30	122,363	39.09	90	20.11	42,436	10.90	3,772	2.47	7,593	18.43	18,963	29.25	19.1	61.04
30	PHOENIX, AZ	11	4 / 2	31.38	131,042	32.49	110,793	31.83	92	21.10	62,091	26.38	13,597	16.11	4,019	59.83	21,363	32.98	9.4	24.11
31	AUSTIN, TX	10	4 / 1	30.75	129,869	31.81	82,940	14.35	50	2.27	51,378	17.94	18,394	22.78	3,506	72.20	7,399	11.28	8.3	42.61
32	BOSTON, MA	13	3 / 3	30.46	146,510	41.54	128,407	42.88	144	44.79	97,775	54.48	10,086	11.24	7,310	13.67	4,378	6.59	12.3	11.33
33	HOUSTON, TX	9	1 / 3	29.72	119,425	25.70	75,299	9.55	72	11.96	62,980	27.08	15,913	19.33	5,319	28.48	2,316	3.38	6.0	82.61
34	CINCINNATI, OH	10	3 / 3	28.28	139,564	37.48	109,573	31.06	93	21.37	62,108	26.39	1,993	0.00	7,455	16.11	4,520	6.81	16.4	41.09
35	SAN FRANCISCO, CA	10	2 / 3	26.79	90,388	8.71	85,394	15.89	117	32.61	77,867	38.80	14,752	17.72	3,632	69.17	1,372	1.91	19.0	60.00
36	DALLAS, TX	10	2 / 2	26.29	143,931	40.04	102,424	26.57	79	15.36	66,938	30.20	13,844	16.46	7,231	12.33	5,292	8.00	9.7	18.26
37	WASHINGTON, DC	15	4 / 2	26.09	75,508	0.00	62,125	1.28	65	9.16	52,380	18.73	3,921	2.68	2,689	91.91	656	0.80	5.0	100.00
38	LAS VEGAS, NV	10	0 / 2	26.03	89,308	8.07	60,086	0.00	54	4.12	44,722	12.70	6,128	5.74	2,353	100.00	19,197	29.61	6.8	68.70
39	JACKSONVILLE, FL	7	0 / 3	24.89	129,510	31.60	110,730	31.79	58	5.99	49,633	16.57	2,976	1.37	3,735	66.68	3,914	5.86	10.4	5.59
40	ATLANTA, GA	11	2 / 2	24.69	150,763	44.03	102,043	26.34	86	18.64	51,207	17.81	6,464	6.21	6,445	1.33	5,264	7.96	12.9	15.70
41	ST. PAUL, MN	10	3 / 5	23.72	126,253	29.69	105,491	28.50	68	10.22	51,742	18.23	4,207	3.07	4,454	49.34	22,447	34.66	10.5	4.35
42	EL PASO, TX	6	0 / 1	22.63	84,732	5.40	61,063	0.61	45	0.00	28,590	0.00	11,160	12.73	3,850	63.92	19,008	29.32	6.3	76.81
43	OKLAHOMA CITY, OK	8	1 / 3	21.38	114,785	22.98	103,908	27.51	82	16.74	43,750	11.94	4,155	3.00	5,220	30.88	16,271	25.07	13.6	20.91
44	HARTFORD, CT	10	1 / 4	20.80	88,234	7.45	80,820	13.01	100	24.62	63,644	27.60	7,841	8.12	3,631	69.19	589	0.70	8.5	39.13
45	KANSAS CITY, MO	12	2 / 3	20.52	124,473	28.65	103,285	27.12	88	19.49	54,539	20.43	2,620	0.87	5,477	24.68	7,433	11.33	10.3	8.70
46	DENVER, CO	11	0 / 3	20.16	103,026	16.10	76,153	10.08	83	16.91	56,469	21.95	5,470	4.83	3,737	66.64	31,488	48.71	10.2	9.88
47	PORTLAND, OR	10	2 / 2	19.42	94,968	11.39	79,720	12.32	69	10.71	50,523	17.27	4,796	3.89	3,530	71.64	17,874	27.56	9.5	21.74
48	GREENSBORO, NC	11	3 / 3	17.04	103,481	16.37	77,542	10.96	73	12.48	42,952	11.31	3,597	2.23	4,396	50.74	3,109	4.61	13.5	19.67
49	ALBANY, NY	16	3 / 5	14.83	93,643	10.61	79,729	12.33	77	14.57	56,460	21.95	2,934	1.31	4,290	53.31	2,819	4.16	12.4	11.82
50																				
51	TOTALS	403	75 / 107																	

Appendix C – Sample Screen Shots of the Microsoft Access Model









This page intentionally left blank.

The Effects of State Mandates on Federal Electronically Filed Returns

Michelle Chu and Melissa Kovalick
Internal Revenue Service, U.S. Department of the Treasury

Background Information

The Restructuring and Reform Act of 1998 stated that the Internal Revenue Service (IRS) should set goals to have a minimum of eighty percent of all Federal tax and information returns filed electronically by calendar year (CY) 2007. The benefits of electronically filing tax returns are numerous; the IRS reduces operational expenses by reducing the need to manually transcribe returns data, the probability of transcription errors is eliminated, and tax law compliance is improved through elimination of computational mathematical errors.

Additionally, the availability of paperless filing corresponds with the increased trend in society to use electronic options in a growing number of areas, specifically the banking industry. While many factors affect the continued growth of electronic filing for Federal returns, this paper focuses on the corresponding effects of state-level electronic filing mandates (which require state individual income tax returns to be electronically filed by practitioners meeting specified qualifications) on the volume of Federal individual income tax returns that are electronically filed. The paper will only focus on the first year experiences of each of the five states that had electronic filing mandates as of calendar year (CY) 2004 due to data availability.

Overview of Individual Electronic Filing Program

According to the IRS website, the option of e-filing individual income tax returns as an alternative method to filing paper income tax returns began in 1986. In 1986, 25,000 individual returns were transmitted to IRS Cincinnati Campus via a modem connection. After almost two decades of incorporating technological advancements into the electronic filing programs, the IRS has received over 61 million electronically filed individual income tax returns in processing year 2004. These returns were filed using three electronic filing options: on-line (including the Free File program), practitioner filed, and the TeleFile program. Total individual electronically filed returns have reflected solid growth in recent years, with an annual average growth rate of over sixteen percent in the past five years. In CY 2004, the percent of total individual income tax returns that were electronically filed was 47 percent. The volume of individual returns filed electronically from CY 1986 to 2004 is presented in Graph 1.

In order to further encourage electronic filing of individual returns, the IRS has established and/or supported several initiatives. Recent noteworthy initiatives include the Free File program and "e-services." In conjunction with the private industry, a consortium of private sector companies was formed to provide free tax preparation and electronic filing options via the Internet. Beginning in CY 2003, certain qualifying taxpayers were able to access various consortium member companies' website and successfully prepare and electronically file their Federal individual returns free of charge. More than 2.7 million individual returns were electronically filed during the program's inception year. The Free File program resulted in approximately 3.5 million returns being electronically filed during the second year of its existence.

The "e-services" products are designed to offer a suite of web-based products to qualified Electronic Return Originators (EROs) by allowing tax professionals who electronically file a minimum of five tax returns to conduct business with IRS electronically 24 hours a day, seven days a week. "E-services" such as electronic account resolution, disclosure authorization, and automated transcript delivery system can be used by qualified EROs. IRS will continue enhancing the services offered by this program as a further incentive to file electronically.

Another factor that influences the growth in Federal electronic filing is the state-level initiatives mentioned above. Specifically, the IRS has experienced stronger adaptation rates in certain states that have mandated the electronic filing of state returns. Although the specifics vary by state, the basic premise of the mandates is that tax practitioners who file a certain number of state-level returns in a given year must electronically file all state returns in the next year. Penalties for non-compliance may or may not exist. This corresponds with the Federal electronically filed program because it is assumed that practitioners mandated to file state returns will be more likely to electronically file Federal tax forms as well. Although the mandates do share some basic characteristics, no two state mandates are exactly the same. As of CY 2004, the marginal effect of Federal electronic filing also varied among the five states that have imposed mandates.

A study was conducted by IRS Research staff in the Wage and Investment (W&I) Division which analyzed the relationships that may exist between state level electronic filing and Federal electronic filing programs (McMillian). Although the study was mostly focused on the affects of the joint Fed/State electronic filing program, some analysis was done on the states that have mandated electronic filing requirements. The study found that the states with mandated electronic filing requirements consistently have the highest electronic participation rates. This analysis suggested that this important topic should be studied further in order to reach a more conclusive and meaningful outcome.

Although there were only five states that had imposed mandates as of CY 2004, there are additional four states that are scheduled to enact mandates in CY 2005. Due to the cost savings associated with electronically filed returns, the success of one state will be instrumental in encouraging other states to join the program. As W&I Research staff indicated, it is imperative that the IRS study the underlying cause and effect relationships to more accurately estimate the effects of these mandates at the state level. This paper will explore the possible reasons for the varied influences of mandated states on the federal returns. It will also attempt to identify the characteristics that significantly contribute to the volume of Federal electronically filed returns, and to build models that can quantify the marginal effect of the state mandates.

State Mandates

The five mandated states that this paper will focus on are Minnesota, Wisconsin, California, Michigan, and Oklahoma. General descriptions of each state's mandate will show the similarities and differences among the mandates. The state of Minnesota was the first state to impose an electronic filing mandate on personal income tax returns. Practitioners who had prepared more than 100 individual income tax returns in CY 2000 were required to electronically file state individual income tax returns starting in CY 2001. There is a fine of $5 per return assessed for every return filed on paper. The state of Wisconsin followed in CY 2003 with a comparable electronic filing mandate. The specifics of the mandate are similar to that of Minnesota with the exception of any penalty for non-compliance.

Three of the five states with state mandates imposed their rules in CY 2004. The state of California started requiring electronic filing of individual income tax returns from practitioners who filed more than 100 returns in CY 2004. However, the state established a new precedent with a $50 penalty for each acceptable return filed using software that was not electronically filed. Although the penalty aspect of the mandate had to be delayed one year due to untimely legislative activities, the jump in Federal electronically filed returns from the state of California indicates that the effect of this unplanned delay was minimal with respect to compliance.

The state of Michigan imposed electronic filing mandates on practitioners who prepared more than 200 individual income tax returns. The electronic filing mandate in Oklahoma was imposed on practitioners who prepare more than fifty individual returns; these practitioners must now electronically file all individual income tax returns. Michigan and Oklahoma do not currently impose any penalties for non-compliance.

Assumptions and Regressors

In an effort to quantify the effects of state electronic filing mandates on Federal electronically filed individual income tax returns, there were numerous explanatory variables that were considered and included in the analysis during the experimental phase. Due to the fact that the state electronic filing mandates are a relatively new phenomenon, there were many uncertainties associated with the observations available. Because of this, there were also a number of assumptions generated which helped shape the final set of ten variables used for the analysis.

The "EPRate" variable represents the mandating states' total individual electronic participation rate at the year *t-1*. Year *t* represents the first year of the mandate and the participation rate is equal to the total Federal electronically filed returns as a percent of the total individual returns for the mandating states.

Another variable considered, labeled "Mandated", is the percentage of state practitioners covered by the various mandates. The data set used to generate these percentages excluded practitioners who filed less than ten returns. It is assumed that practitioners filing less than ten returns do not represent the generally accepted definition of a practitioner.

An additional factor that may contribute to the volume of Federal electronically filed returns is the existence of larger practitioners (who mainly prepare individual income tax returns) in the mandated states. In this instance, large practitioners are defined as those who file more than 200 tax returns in a given tax year. The underlying assumption is that larger practitioners are more established in the public community, have larger client bases, and are more likely to have the necessary infrastructure and financial resources to support newer technologies. These factors are expected to increase large practitioner's response to, and compliance with, electronically filed mandates.

Whether the state is a member of the joint Fed/State program (variable "FS") may also contribute to the state mandates, translating into higher volumes of Federal electronically filed returns. The joint Fed/State program allows the electronic filing of both Federal and state income tax returns at the same time. The Federal and state return data are placed in separate packets by the software. These packets are then transmitted in one "envelope" to the IRS and then forwarded to the participating state, which receives and processes the state electronic return. Membership in this program should serve to promote the states' acceptance of electronic filing due to the benefits of the joint Fed/State program participation, including faster processing and one confirmation for the receipt of returns at both levels.

The existence of a state-imposed penalty (variable "Penalty") for failure to comply with the mandate may also contribute to the increase in Federal electronically filed returns. In this case, the assumption is that when state practitioners must pay penalty fees for non-compliance, especially when the amount of the penalty is severe, such as the $50 per return penalty enforced by the state of California, practitioners are much more likely to comply with the mandate. As a result, more Federal returns will be electronically filed along with the state returns.

Since one of the primary benefits most often associated with electronic filing involves receiving refund checks more rapidly, an assumption was made regarding the amount of average state refunds (variable "Refund"). The expectation is that the higher the amount of the state refund, the greater the willingness will be to electronically file state returns. Similarly, a corresponding increase will be observed in the volume of Federal returns that are electronically filed.

Additional independent variables that were considered in the preliminary modeling analysis included the mandated states' levels of real per capita personal income (variable "PersonalInc") and resident employment (variable "Emp"), total population (variable "Pop"), and the amount of federal tax payments (variable "Fedtaxpymt") for the initial year each state had a mandate in effect. These variables were initially selected due to the anticipated casual relationships with the dependent variable. For example, it was expected that growth in real per capita personal income and resident employment would lead to increases in electronically filed returns.

Similarly, as levels of employment accelerate, personal income has a strong tendency to increase as well. From 1990 through 2003, personal income (measured in CY 2000 dollars) and total employment at the U.S. level reflected nearly perfect proportional increases; an increase of one unit of employment corresponded with an increase of almost exactly one additional unit of personal income, or $0.9923 *(Global Insight, Inc.)*. The total population levels in each state were also considered to account for the effects of the magnitude of the taxpayer pool for each state.

Finally, an assumption was made that the amount of federal tax payments would have a negative correlation with the number of electronically filed returns. This was based on the premise that taxpayers with balances due on their returns are less likely to electronically file as there is little perceived incentive to make an early payment.

For all of the explanatory variable data sets, with the exception of "FS" and "Penalty," values were measured via quantitative, nominal data. Data relating to taxation was extrapolated from internal data sources *(refer to "Data Sources")*, while economic and demographic variables were provided by Global Insight, Inc., a commercial database vendor. Binary "dummy variables" were used to represent enrollment in the joint Fed/State program and the existence (or lack) of penalties.

There are a few interesting state level observations to note from the input data set. Among the five states that currently impose mandates, the states of California and Minnesota had the lowest EPRates, at 32 percent and 30 percent respectively. These states are also the only two states that are currently not enrolled in the joint Fed/State program, and they are also the only states with penalties imposed. Interestingly, these states also have the highest values for state tax refund amounts and real per capita personal income.

Derivation of Y

This section will discuss the definition of the dependent variable, "Y." Since indicators do not currently exist to identify the Federal individual income tax returns electronically filed with the IRS solely as an indirect response to the state-level electronically filed mandates, "Y" had to be derived from a combination of returns data. One assumption made was that the marginal difference in the volume of returns projected prior to the knowledge and existence of state mandates, and the actual volume of returns processed for the mandated year, represents the volume of returns resulting from the mandate, with an adjustment made for the underlying annual growth rates. The methodologies used to generate these volumes involve the use of NHQ Office of Research Document 6187 (Calendar Year Projections of Individual Returns by Major Processing Categories) projections of Federal electronically filed returns by practitioners at the state level.

The projections of electronically filed returns from the states for year *t,* prior to the existence of mandates were generated. Then, the actual master file volume of returns after the states had mandated electronic filing was downloaded. To correct for the errors inherent in most projections work, the projection error rate was calculated for two years prior to the mandate, with the exception of Minnesota, where only one year of data was available. The average of the error rates was derived and applied to the original set of projections. The difference between the two set of numbers represents the assumed marginal effect (referred to as "Y_{unadj}") of state electronically filed mandates on the Federal electronically filed returns from the tax practitioner community.

Descriptive Statistics for Y

A preliminary analysis of the unadjusted data ("Y_{unadj}") for the marginal change in the electronically filed Federal returns (the dependent variable) indicated that the data did not fit the pattern of a normal distribution. Rather, the initial data set for "Y_{unadj}" was negatively skewed.

$$skewness = \frac{\sum_{i=1}^{N}(Y_i - \bar{Y})^3}{(N-1)s^3}$$

Skewness "Y_{unadj}" = - 0.73

In addition, the kurtosis value of 0.41 was further indication that the distribution was not normal. Since the kurtosis for a standard normal distribution is 3.0, excess kurtosis is defined as

$$kurtosis = \frac{\sum_{i=1}^{N}(Y_i - \bar{Y})^4}{(N-1)s^4} - 3$$

which indicates that the kurtosis is zero for the standard normal distribution *(NIST)*.

In order to normalize the data set for the dependent variable (which enhances subsequent statistical analyses), the values for "Y_{unadj}" were squared in order to derive the data set "Y_{adj}." While the adjusted data was still somewhat platykurtic (-0.50), the revised skewness (0.36) indicated a more normal distribution than that observed for "Y_{unadj}."

Correlation and Identification of Significant Independent Variables

For each independent variable considered, Pearson's *r* was calculated to determine the existence and/or magnitude of a linear relationship with "Y_{adj}." With the exception of the number of large practitioners and the volume of practitioners covered by the mandate, all other explanatory variables exhibited moderate to high levels of correlation (each measuring a minimum r-value of at least 0.55) with the dependent variable. However, the coefficients of correlation indicated that high levels of multicollinearity were present among many of the independent variables, particularly resident employment, federal tax payments, and total population. The calculated *r* denoted an almost perfect correlation among these three variables (r was greater than 0.99 in each inter-relation).

Due to the limited size of the population data, the number of independent variables had to be minimized in order to optimize the statistical significance of the model. Issues surrounding the restricted degrees of freedom presented one of the greatest constraints in the analysis. With only five states having mandatory electronically filed programs in existence as of 2004, the initial degrees of freedom were already limited to four. Since the addition of each independent variable would use another degree of freedom and thus increase the spread of the t-distribution, further testing was necessary to detect the variables that accounted for the most variation in "Y_{adj}."

A stepwise regression (Attachment 1) was run to determine which of the variables had the most significant impact on the dependent variable. At a p-level of 0.85, resident employment was identified as having a significant impact on "Y_{adj}." Resident employment accounted for more than three-fourths of the total variation in the univariate regression model (R-square was equal to 0.7618). However, stepwise regression has some well-documented limitations such as biases present in the R-square values and regression coefficients, along with problems when collinearity is present *(Harnell).* These factors led to the generation of alternative methods for identifying other possible explanatory variables since Pearson's r values indicated high levels of correlation among the explanatory variables. Additionally, t-test values suggested that several other explanatory variables, including the number of large practitioners (with a calculated t-value of 23.31 and a p-value of 0.0001) and electronic filing participation rate (with a calculated t-statistic of 12.51 and a p-value of 0.002), were also significant despite the results of the stepwise regression.

Using various combinations of explanatory variables, all possible models were generated via the SAS R-SQUARE procedure (Attachment 2). The results were sorted and ranked, initially based on the number of explanatory variables used (consideration of the degrees of freedom), then on the selected statistics of fit. The statistics of fit used to rank these models were R-square, adjusted R-square, and mean square error (MSE). The resulting regression models were further analyzed based on other statistical measurements including the t-test, calculated F-values, and the beta coefficients, in order to select the optimal model to explain the relationship among the dependent and independent variables, while eliminating model bias such as multicollinearity.

As previously mentioned, the limited population size significantly constrained the statistical analyses that could be performed on the regression models. Therefore, assessing and selecting one definitive model to predict the marginal change in Federal electronically filed returns did not appear to be a prudent option at this point in the development cycle of the state-mandate programs. For the final analysis, three separate models (selected based on their overall statistical significance) were examined for their respective goodness of fit and explanatory values. Each model had its specific advantages and disadvantages; however, since the restricted input data did not allow for verification and validity of the predictive nature of the equations, ranking the models was complex.

The first model (Attachment 3) used refund amount, federal tax payments, and population as independent variables. The resulting equation is expressed as:

$$MRtnsq = -0.17894 + 0.00045867 * Refund - 0.00001806 * Fedtaxpymt + 0.00005302 * Pop$$

Where:
MRtnsq = Y variable squared
Refund = Average refund amount from state returns filed electronically
Fedtaxpymt = Federal tax payments
Pop = State level population

The adjusted R-square was 0.9989, meaning that nearly all of the variation in the equation is accounted for by the explanatory variables. The F-value of 1178.50 also indicated that the model was significant, as did the p-value of 0.0214 at an alpha of 0.05. Furthermore, each individual variable also showed significance according to the calculated t- and p-values. Of the final three models considered, the goodness of fit statistics was the best for this equation. On an intuitive level, the predictive variables are logical since taxpayers who are owed refunds are likely to be attracted to electronic filing since refunds deposits will be received quicker (as opposed to paper processing). Similarly, taxpayers who owe the IRS money (as indicated by the "Fedtaxpymt" variable) are likely to behave in a contrary manner. Although the "Fedtaxpymt" and population variables appeared to exhibit multicollinearity (r was equal to

0.9965), there does not appear to be an innate, causal relationship between the variables; increases in population do not inherently affect income tax payments. In fact, the reverse may be true; as taxpayers claim more dependents, they may become less likely to owe Federal payments.

The second model (Attachment 4) explained the marginal increase (squared) in electronically filed returns via the state's electronic participation rate, the existence of penalties, and the level of real personal income. The equation generated is:

*MRtnsq = -1.71503 +1.89159 * EPRate + 0.15236 * Penalty + 0.03313 * PersonalInc*

Where:
MRtnsq = Y variable squared
EPRate = State-level Federal electronic participation rate at time t-1
Penalty = Existence of penalties
PersonalInc = Real per capita personal income

This model was also significant at an alpha level of 0.05, with a *p*-value of 0.0356 and an F-value of 426. Per the adjusted R-square statistic, the explanatory variables accounted for over 99 percent of the total variation in the model. Although all of the independent variables were significant to the equation, this model is slightly less preferable to the model described above due to some occurrences of multicollinearity. Personal income and electronic file participation have a strong negative correlation (*r* is equal to -0.89); this relationship appears to be somewhat causal since lower income taxpayers that electronically file their income tax returns are more likely to use practitioners. According to Statistics of Income's R:S-97 Report for filing year 2004, over 77 percent of taxpayers with Adjusted Gross Income (AGI) of less than $20,000 used practitioners to electronically file individual income tax returns. In addition, "EPRate" and "Penalty" reflect high levels of negative correlation with a calculated *r* of -0.971. This does not follow the logical assumption that states which enforce monetary penalties for non-compliance would have higher rates of electronic filing. However, this regression accounts for high levels of variation and can still be utilized as a predictor of marginal changes in electronic filing.

The final model (Attachment 5) was a bi-variate regression utilizing the state's electronic participation rate and the amount of refunds, described as:

*MRtnsq = -1.14515 + 2.07874 * EPRate + 0.00090317 * Refund*

Where:
MRtnsq = Y variable squared
EPRate = State-level Federal electronic participation rate at time t-1
Refund = Average refund amount from state returns filed electronically

Again, this model was significant at an alpha of 0.05 with a *p*-value of 0.0324 and an F-value of 29.9. The adjusted R-square was 0.9353. An advantage of this model is that there was an additional degree of freedom present in comparison to the first two models. However, one disadvantage is that with one less explanatory variable, it becomes more difficult to discern the marginal changes in Federal electronic filing volumes since additional independent variables account for more variation in the model (hence the higher R-square values for the previous models). Also, due to the unexplored nature of state mandates, it may be preferable to explore as many outside factors as possible in an attempt to discern the predicted impact these mandates will have on Federal electronic filing. Therefore, it is suggested that a combination of all three models be used to derive a valid range that will establish the most accurate impacts on electronic filing of Federal returns until the population size is augmented to allow for further analysis.

In order to further test the validity of the models discussed above, the estimated marginal impact of the Alabama's state electronic filing mandate were generated. Based on model one, the IRS can expect around 108,700 practitioner electronically filed returns from the state of Alabama as a result of the state level mandate in CY 2005. However, models two and three resulted in marginal affects of 324,800 and 360,500 returns, respectively. The range suggested by the models indicates that model one represents a more conservative estimate whereas models two and three imply a more aggressive outcome. Since there are several unknown factors, such as state level legislative and/or budgeting activities that may affect the final outcome of the returns, a more conservative estimate is suggested.

Conclusion

The analysis presented in this paper supports the initial findings of the 2004 IRS study by Research Staff in the W&I Division. This study adds another layer to the initial findings by evaluating ten variables that may help estimate the effects of mandates on Federal electronic filing. Basic assumptions were made around this topic and independent variables were selected based on the assumed relationships to the dependent variable as well as the availability of data. The stepwise regression procedure was initially generated in order to identify the best set of explanatory variables. However, due to the limitations inherent in stepwise regression procedures, a decision was made to use additional methods to select explanatory variables. Based on the SAS R-SQUARE option, a total of twelve models with three independent variables and eight bi-variate models were selected.

Using model statistics, three models were selected as the final set of models to be included in this study. The model experimentations showed that a model which includes the electronic participation rate, penalty, and personal income and another model that includes the state refund amount, Federal tax payments, and state population were representative of models with best model statistics. A significant level of model statistics was also found in a bi-variable model that included electronic participation rate and state refund amount.

Although the findings were significant, there are also limitations to this study that should be discussed. One of the limitations to this study is the method used to calculate the dependent variable. The assumptions used to derive the dependent variable presumed that the differences in the projected and actual volumes were solely representative of the effects of the electronic filing mandate. This may be a valid assumption but there is always a probability that other explanations for the marginal difference exist. Since complete information does not exist on the various state-level program activities, other possible reasons for the marginal difference cannot be completely dismissed.

In addition, another limitation to this study stems from the fact that there were only five observations at the time of this study. As a result of the limited number of observations available, the inherent problems related to the degrees of freedom presented constraints on model experimentation.

Moreover, additional explanatory variables should be explored in future studies of this topic. The explanatory variables can also be experimented in a transformed format. Transformed explanatory variables may enhance the model results. Future studies may also analyze the effects of state-level mandates in years following the first year of mandate. State mandates are usually planned such that the thresholds are gradually reduced so that a greater number of practitioners are covered under the mandate in the following years. Future analysis could be tailored around a new set of assumptions in an effort to quantify the effects for years following the first year of the mandate. Also, alternative methods of selecting independent variables could be used for model experimentation. Stepwise regression and SAS RSQUARE options were used for this study but variable selection based on additional procedures may result in alternative models with enhanced model statistics.

This study can also translate into analyzing the state level business electronic filing mandates. The practice of mandating electronic filing of business returns already exists for a number of states. Although the scope of this study was confined to the individual income tax return area, it would be interesting to study a similar effect on the business side.

It has become apparent that the state-level electronic filing mandates are gaining popularity in the electronic tax filing area. Studies show that these state-level mandates directly contribute to an increase in Federal electronically filed returns. The five states that have implemented state e-file mandates have contributed different degrees of electronic returns at the Federal level. The three states that have imposed electronic filing mandates in CY 2004 are assumed to have contributed a total of around 2.5 million Federal electronic income tax returns in 2004 alone. Additional four states (Alabama, New Jersey, Massachusetts, and Virginia) have imposed similar mandates in CY 2005. Many more states are in the process of considering some form of electronic filing mandates. All things considered, it is important for the IRS to effectively quantify the effects of these mandates. The IRS should consider assisting and encouraging states to mandate electronic filing of tax returns. This will allow the IRS to properly plan and budget for the future.

Works Referenced:

Harrell, Frank E. Regression Modeling Strategies. New York. Springer, 2004. "History of e-file". Internal Revenue Service. April 2005. http://www.irs.gov/efile/article/

IRS, RAS, Office of Research. *"Document 6187, Calendar Year Projections of Individual Returns by Major Processing Categories."* December 2004.

McMillian, Howard. *"State e-File Impact on Federal e-file."* Wage and Investment Research Group 1, Internal Revenue Service, 2004.

Mendenhall, W. and Sincich, T. *A Second Course in Business Statistics: Regression Analysis.* Macmillan, Inc, 1986.

"Minnesota Statutes 2004, 289A.08". Minnesota Office of the Revisor of Statues. February 2005. http://www.revisor.leg.state.mn.us

"NIST/SEMATECH e-Handbook of Statistical Methods", National Institute of Standards and Technology. http://www.itl.nist.gov/div898/handbook

"State E-File Mandates". NTS Services Corp. March 2005. http://www.ntslink.com/state_mandates.asp

"State Electronic Filing Programs for Individual Income Tax.". Federation of Tax Administrators. January 2005. http://www.taxadmin.org/fta/edi/PIT_snaps.html

Data Sources:

Global Insight, Inc.

IRS, Master File Reports. CY 1999-2004.

IRS, Electronic Tax Administration Marketing Database. CY 1999-2004.

IRS, State Income Tax Filing Programs. November 2004.

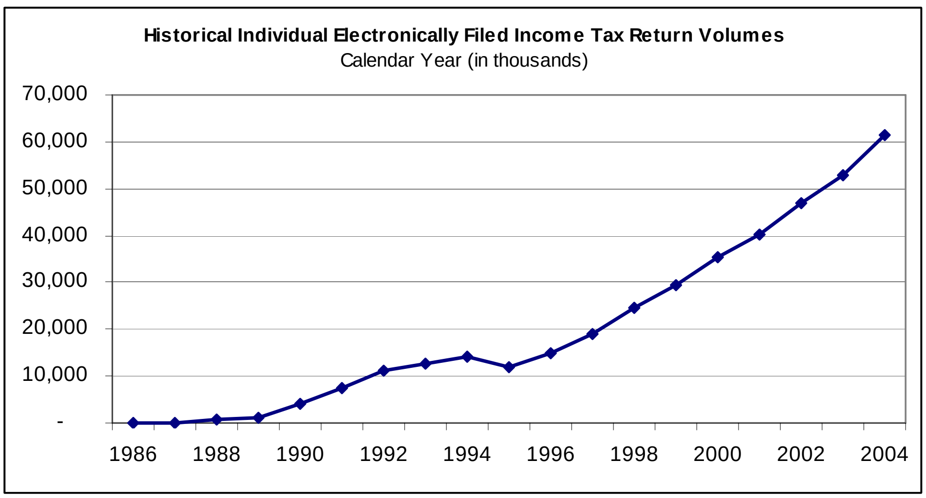
Statistics of Income R:S-97 Report, IRS Master File System, Filing Year 2004.

Acknowledgements

The authors would like to thank Terry Lanning, Melissa Wuebbels, and Kevin Mullen for providing us with critical information. We are grateful for your expertise.

Note: The views expressed in this article represent the opinions and conclusions of the authors. They do not represent the opinion of the Internal Revenue Service.

Graph 1



Note: Volumes represent the IRS master file data. The drop in CY 1995 is a result of IRS protection strategies instituted to combat refund fraud.

Attachment 1

The REG Procedure
Model: MODEL1

Dependent Variable: MRtnsq MRtnsq

Number of Observations Read 5

Number of Observations Used 5

Stepwise Selection: Step 1

Variable Emp Entered: R-Square = 0.7618

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	0.00646	0.00646	9.59	0.0534
Error	3	0.00202	0.00067298		
Corrected Total	4	0.00848			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	0.01865	0.01672	0.00083706	1.24	0.3460
Emp	0.00000652	0.00000210	0.00646	9.59	0.0534

Bounds on condition number: 1, 1

All variables left in the model are significant at the 0.1500 level.

No other variable met the 0.1500 significance level for entry into the model.

Attachment 2

Number in Model	Adjusted R-Square	R-Square	MSE	Variables in Model
2	0.9676	0.9352	0.00013722	EPRate Refund
2	0.8913	0.7826	0.00046070	PersonalInc Pop
2	0.8856	0.7711	0.00048494	PersonalInc Emp
2	0.8834	0.7668	0.00049414	LgPrac PersonalInc
2	0.8717	0.7434	0.00054378	PersonalInc Fedtaxpymt
2	0.8503	0.7006	0.00063433	Mandated Fedtaxpymt
2	0.8318	0.6636	0.00071273	Mandated Emp
2	0.8273	0.6546	0.00073177	Mandated Pop
3	0.9997	0.9989	0.00000240	Refund Fedtaxpymt Pop
3	0.9995	0.9978	0.00000460	EPRate Mandated Refund
3	0.9992	0.9969	0.00000663	EPRate Penalty PersonalInc
3	0.9992	0.9969	0.00000663	EPRate FS PersonalInc
3	0.9974	0.9898	0.00002165	EPRate Refund PersonalInc
3	0.9966	0.9864	0.00002875	LgPrac Penalty Emp
3	0.9966	0.9864	0.00002875	LgPrac FS Emp
3	0.9963	0.9852	0.00003128	EPRate Penalty Refund
3	0.9963	0.9852	0.00003128	EPRate FS Refund
3	0.9916	0.9665	0.00007098	LgPrac FS Fedtaxpymt
3	0.9916	0.9665	0.00007098	LgPrac Penalty Fedtaxpymt
3	0.9866	0.9463	0.00011370	LgPrac FS Pop

Attachment 3

Marginal Effects of State E-file Mandates

The REG Procedure
Model: MODEL1

Dependent Variable: MRtnsq MRtnsq

Number of Observations Read	5
Number of Observations Used	5

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	0.00847	0.00282	1178.50	0.0214
Error	1	0.00000240	0.00000240		
Corrected Total	4	0.00848			

Root MSE	0.00155	R-Square	0.9997
Dependent Mean	0.05596	Adj R-Sq	0.9989
Coeff Var	2.76628		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	-0.17894	0.00705	-25.38	0.0251
Refund	Refund	1	0.00045867	0.00001580	29.02	0.0219
Fedtaxpymt	Fedtaxpymt	1	-0.00001806	7.025987E-7	-25.70	0.0248
Pop	Pop	1	0.00005302	0.00000197	26.85	0.0237

Attachment 4

Marginal Effects of State E-file Mandates

The REG Procedure
Model: MODEL1

Dependent Variable: MRtnsq MRtnsq

Number of Observations Read	5
Number of Observations Used	5

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	0.00847	0.00282	426.02	0.0356
Error	1	0.00000663	0.00000663		
Corrected Total	4	0.00848			

Root MSE	0.00257	R-Square	0.9992
Dependent Mean	0.05596	Adj R-Sq	0.9969
Coeff Var	4.59978		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	-1.71503	0.06669	-25.72	0.0247
EPRate	EPRate	1	1.89159	0.09732	19.44	0.0327
Penalty	Penalty	1	0.15236	0.01012	15.06	0.0422
PersonalInc	PersonalInc	1	0.03313	0.00117	28.36	0.0224

Attachment 5

Marginal Effects of State E-file Mandates

The REG Procedure
Model: MODEL1

Dependent Variable: MRtnsq MRtnsq

Number of Observations Read 5

Number of Observations Used 5

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	0.00820	0.00410	29.88	0.0324
Error	2	0.00027443	0.00013722		
Corrected Total	4	0.00848			

Root MSE	0.01171	R-Square	0.9676
Dependent Mean	0.05596	Adj R-Sq	0.9352
Coeff Var	20.93137		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	-1.14515	0.21386	-5.35	0.0332
EPRate	EPRate	1	2.07874	0.39540	5.26	0.0343
Refund	Refund	1	0.00090317	0.00014188	6.37	0.0238

This page intentionally left blank.

Forecasting Key Aspects of Immigration

Session Chair: Signe Wetrogan, U.S. Census Bureau (signe.i.wetrogan@census.gov)
Discussant: Gregory Robinson

Improving the Measurement of Net International Migration for State and County Population Estimates

Katherine M. Condon (katherine.m.condon@census.gov), Administrative Records and Methodology Research Branch, and Sam Davis, Population Estimates Branch Population Division, U.S. Census Bureau

An important component of population change in the postcensal state and county estimates produced by the Population Estimates Program at the U.S. Census Bureau is net international migration. Based on internal evaluations, the distribution of international migrants at the subnational level is a problematic area for annual population estimates and is being investigated for improvement. In this paper we will present the approaches used for the 2001 and 2002 sets of county estimates and explore the effects of using each method by comparing the two sets of estimates and by comparing the differences in the distribution of net international migration. We will follow this with an examination of the relationship between the distribution of the change in the foreign-born population (which may include both net international movement and net internal movement of the foreign-born population) and the component of net international migration for population estimates at the subnational level. An example of the distributional differences between the two will be given. In conclusion, there will be an examination of the potential for improving the distribution of international migration at the subnational level for population estimates through utilization of two "new" sources of data: the American Community Survey (ACS) and Social Security Administration's 100% Numident file linked to IRS data (IRS-SSA). The ACS data could be used to give an indication of change in the county distribution of the foreign born since Census 2000. The IRS-SSA data are currently used for estimating net internal migration. Additionally, we believe IRS-SSA data could be extended to measure net internal migration of the foreign-born population.

Toward a Legal Immigrant Data Series for Population Projections

Frederick W. Hollmann, U.S. Census Bureau

This is a preliminary report on the results of an attempt to create a 31-year series of arrivals to the United States of people destined to become legal permanent residents. The activity was carried out in support of a program to produce U.S. population forecasts by applying time series methodology on the major components of change, including the components of net international migration. The sole data source employed in this activity was the immigrant micro-data file, formerly provided by the Immigration and Naturalization Service (INS), and currently (for FY 2002) by the Department of Homeland Security (DHS.) The emphasis in the activity was to establish an annual trend, not estimate an absolute level of legal immigration at some point, although evidence about level of immigration appears as a by-product.

Forecasting Age Distribution Curves

Tucker McElroy and William Bell, U.S. Census Bureau

The U.S. immigration data contain empirical age distributions for various categories of immigrants, and these curves evolve in time. Forecasts of the age curves must still be a distribution, and must reflect major dynamical shifts (such as modal movements) while suppressing noise. To address these issues, the authors have used logistic transforms, principal components analysis, cubic spline smoothers, and Bayesian methods. The combination of these methods provides satisfactory forecasts of the age distributions, and is an innovative use of several statistical techniques.

This page intentionally left blank.

Toward a Legal Immigrant Data Series for Population Projections

Frederick W. Hollmann
U.S. Census Bureau

The corporate decision of the Population Division to incorporate formal time-series analysis in our long-term population forecasts has created an unprecedented need for historical data series on components of population change. In the case of international migration (by many accounts, the most difficult component to forecast), we must assemble an annual series for net international migration by major component (legality, legal basis, foreign-born versus outlying areas, immigration versus emigration, etc.) cross-classified by some detail on country of birth. This level of detail should allow separate consideration of migratory determinants in forecasting migration, as well as characteristics pointing to race and Hispanic origin, in the forecasts. The annual series is a necessary prerequisite for a formal time series treatment that will allow the assessment of stochastic uncertainty. To produce a "definitive" annual international migration series is neither feasible nor entirely necessary for this purpose. The series should nevertheless be both plausible and consistent with our best interpretation of available data, in level, trend, and the degree of fluctuation, as all three of these affect the ultimate forecasts. Of the three major attributes, the overall level is least important, since we tend to identify the baseline level from official estimates that may arise from other sources.

Unfortunately, no easily interpretable data exist to produce an annual series for many of the components of international migration. Undocumented (unauthorized) migration (net of departures and legalizations), the migration of temporary workers (net of departures and transitions to legal permanence), migration of U.S. citizens from outlying areas, and the emigration of legal permanent residents (native and foreign-born) cannot readily be estimated as an annual series. Administrative data generally do not record these categories, and residual analyses of large-scale surveys cannot distinguish them from the flow of legal permanent residents.

With this brief report, we focus on that portion of immigration that is addressed directly by administrative data, namely legal immigrants who obtain legal permanent residency in the U.S. at some point. Because immigration does not always correspond to arrival as "usual residents" in the sense of the census definition, the evidence provided by administrative data is fragmentary, although some of the fragments are quite large. We focus on a single data source that addresses a large component of international migration, namely public-use immigrant data once provided by the Immigration and Naturalization Service (INS), now under the aegis of the Department of Homeland Security (DHS). This source provides comprehensive individual-level data on persons either immigrating directly as legal permanent residents (LPR) of the United States via application to the Department of State or immigrating via adjustment to legal permanent resident status from temporary (or unauthorized) U.S. residency, via application to the INS (now DHS). While the question that instigated this report was directed to the intercensal 1990s, the data series we discuss is for the period from 1972 to 2002, although we focus on the 1990s in the discussion of our findings. Because the large portion consisting of "adjustees" (those already resident in the US before legal immigration) did not arrive at the time of immigration, the largest single interpretive effort was to impute a complete series tied to date of arrival, rather than date of legal immigration.

Types of legal immigration

In this report, we consider four major types of legal immigration: employment-provision immigrants, family-provision immigrants, an aggregate of special immigrant classes, and refugees. Employment immigrants are composed of people who immigrate under a provision of immigration law that relates specifically to their employment, along with their dependents. Family-provision immigrants are those who immigrate under a provision related to their relationship to a U.S. citizen or legal resident of the United States. Special immigrant classes refer to a large variety of special provisions allowing people to immigrate. A recent example, in the 1990s, is the "diversity lottery" according to which immigration applicants from countries underrepresented among other immigrants are selected by random draw. Refugees include those admitted to legal permanent residence under various refugee provisions, either because they were previously admitted by the Department of State as refugees, or because they entered as parolees under special "refugee-like" provisions (from Cuba, in most cases.)

At the end of the report, we will consider, speculatively, the character of the intercensal 1990s series for those components not covered by legal immigration at all.

The need to estimate legal immigration by year of arrival, rather than year of admission

Because our objective is to establish a time series of legal immigration for use in forecasting, it is essential that the time series be representative of actual arrival as U.S. residents, as opposed to their acquisition of the legal status of "immigrant". A large portion (in recent years, more than half) of people who are admitted each year as legal immigrants were residing in the United States—in most cases, legally—before their admission.

While this distinction may appear to be a technical "fine point" motivated by the need to produce a time series that reflected fluctuations as well as level, the distinction is important for the level of immigration as well. In the 1990s, official estimates of the U.S. population produced by the U.S. Census Bureau adopted a "proxy rule" that assumed that the number of new arrivals of present or future legal immigrants in a given year was equal to the number of people becoming legal immigrants in the same year. Because immigration was generally on the rise during this period, this assumption understated the number of arrivals of present or future legal immigrants.

This distinction mandates a clarification of terminology. The definition of words such as "arrival," "admission," and "immigration" are a potential source of confusion, because they tend to be used in different ways in publications of different agencies. Throughout this report, we use the terms "admission" and "immigration" in the way current at INS and DHS. Both terms describe the registered event whereby a person becomes a legal permanent resident of the United States. This is also concurrent with the awarding of a "green card," or work permit. We use "arrival" to mean physical arrival in the United States (irrespective of the permanence of the move) and "census arrival" to mean arrival in the United States as a "usual resident" who, per official intent, would respond to a decennial census as a U.S. resident.

Interpreting the data, part 1: missing date of arrival

A coded variable known as "non-immigrant year of arrival" indicates the calendar year in which a person adjusting to LPR status last arrived in the United States. Hence, it is theoretically possible to simply separate the adjustee portion of the immigrants (via class-of-admission codes) from the new arrivals, and resort the adjustees by year of arrival. For the "adjustee" portion of the legal immigrants, date of arrival normally precedes date of immigration. Hence, it must be assumed—and indeed, the data show—that many people arriving late in the time series will be absent from the immigrant-based data.

Before we address this problem, two other problems intervene. First, non-immigrant year of arrival, because it represents last arrival in the U.S., does not generally match arrival to census-defined "usual" residency, even though it is closer than date of immigration. Second, a significant number of adjustee cases have missing codes for this variable, and the number becomes alarmingly large for fiscal years of immigration beginning 1998. A smaller number of cases have codes that are inconsistent with the respondent's age.

Our strategy for addressing the first problem was to cautiously disregard it, and assume that year of last arrival represents year of arrival to census-defined residency. If the trend in arrivals is "flat," this assumption produces little bias, since "census arrivals" missed in the data are offset by last arrivals, albeit for a different cohort of migrants. If arrivals are increasing steadily (closer to what is assumed to have occurred,) the assumption tends to bias the number of arrivals (level of the trend) downward, but has little effect on the direction of the trend. If the trend in arrivals increases at an increasing or declining rate, the assumption can produce bias in the direction as well as the level of the trend, and this may have been an issue for employment-based immigrants that "exploded" in the late 1990s.

Addressing the second problem of uncoded or inconsistent year of arrival was feasible, but required making some complex assumptions. The strategy was to define a universally coded variable within whose values we could assume that *duration of residence* (as opposed to year of arrival) in the U.S. between arrival and immigration would be uniform across coded and uncoded cases. We did this separately for refugee adjustees and the much larger class of non-refugee adjustees. For non-refugees, the chosen variable was a rather messy composite of class of admission and country of birth, tabulated by year of immigrant admission. We identified three major categories of class of admission, corresponding to those discussed above, 1) employment provisions, 2) family provisions (e.g., close relatives of U.S. citizens or permanent residents), and 3) special provisions. We defined 19 categories of country of birth within which we could see some distinctiveness in immigration trends, with a separate category for each of the major source countries, and regional categories for the others. Within each admission year, we examined the number of cases with coded and uncoded year of arrival for each of the 57 (=3x19) potential groups defined by these two variables, to determine whether the coded cases were sufficiently numerous, and the uncoded cases sufficiently few, to effect an imputation. Some aggregation of categories within year of admission was necessary. In some cases, especially in the years after 1997, we aggregated across year of admission, because the known cases, even for liberally defined country/class categories were too few to effectively impute duration of residence to the unknowns. For this reason, it was imperative that we were imputing duration of residence, rather than year of arrival. We randomly assigned duration-of residence and duration-of-life (age) values in days, consistent with year of arrival and date of adjustment for known-arrival cases, and a value of age in years or date of birth for all cases. We carried out the actual imputation by iteratively assigning to each unknown-arrival case a duration-of-residence value based on the distribution of known cases, while censoring to ensure that duration of residence never exceeded duration of life. The imputed duration value was then converted to a date of arrival for all cases by subtracting from date of adjustment.

For refugees, the process was similar, with a few differences. There was no need to disaggregate the data by class of admission. Because refugee arrivals depend very much on country-specific events and policies, we took a list of 33 country groups, most of which were single countries that have been sources of refugees admitted to the US. Because refugee movements are generally tied to dated events, we could not appropriately impute duration of residence before admission, rather we imputed date of arrival, censoring on both date of birth and date of admission. For each admission year, we randomly imputed unknown arrival date within the distribution of known arrival date. If an imputed date was either earlier than date of birth or later than date of admission, we censored the imputation and repeated the process. A very small percentage of cases could not be imputed in this way because they were either very young, or an inadequate number of known cases was observed for the range of eligible dates. These were ultimately imputed by a rectangular assumption on the eligible range.

Interpreting the data, part 2: arrivals not yet adjusted

This leaves the problem of estimating the annual number of legal immigrants arrived before the end of FY 2002, but not yet adjusted to legal residency, hence not in the data series. We simulated these cases using a combination of two approaches. The first approach relies on projection of future adjustments. The second approach, applicable only to employment-based immigrants, links recent arrivals who are future legal permanent residents to an external data source.

The first approach involved making assumptions about future adjustments to legal permanent resident status. We carried these assumptions forward for 30 years, on the stipulation that anyone arriving by the end of FY 2002 and not adjusted by September 30, 2032 should not be considered a legal immigrant. We assumed further that future adjustments would have the same distribution by duration of residence in the US (within categories of country of birth and broad admission class) as those who adjusted in the last three years of the empirical series, FY 2000 through FY 2002. “Filling out” the year-of-arrival series can thus be accomplished by replicating cases in the three-year period over 30 years, calculating the resulting arrival dates, and discarding all cases (the large majority) that would not have arrived by September 30, 2002. We stress that such assumptions do not put a lot of weight on the actual behavior of DHS in carrying out adjustments; we are only hypothesizing a constant duration-of-residence distribution on adjustments that would occur if adjustments rise to demand as they did in 2000 to 2002. Were DHS to fail to meet the demand, an already-substantial backlog of adjustment applications would become larger. Were DHS adjustments to outpace applications, the backlog could shrink. The assumption is impaired by DHS activity only to the extent that the size of the backlog affects the decisions of pending adjustees to remain or not to remain in the US. We observe further that while the existing adjustment application backlog of about 1 million grew in the late 1990s, it did not change much between 2000 and 2002. Had we chosen to replicate the experience of FY 1995 to FY 1999, there would undoubtedly have been an artificial lag in the duration distribution caused by an increasing inability of then-INS to keep adjustments apace with applications. This approach, so defined, does not specify what future trend in adjustments is realistic, but that can be a subject of hypothesis, as will be discussed below.

The second approach involves projecting one sector, employment immigration, by an external data series, while discarding the actual arrival data from recent years. The series comes from DHS and INS, and consists of the annual number of temporary workers admitted to the U.S. with non-immigrant visas, including temporary workers in the H-1, H-2, H-3, as well as O, P, and Q visa series, NAFTA workers, U.S.-Canada Free Trade Agreement Workers, and spouses and children of these groups (US Department of Homeland Security, 2003, Immigration and Naturalization Service, 2003, and earlier issues.) The basis for this assumption is that for each year, non-immigrants admitted in these categories represent a population at risk for those who will ultimately adjust to legal permanent resident status under employment provisions. We selected 1995 as a year for which almost all of the arrivals of future employment-based immigrants would have adjusted by 2002. This is born out by results of the process described in the previous paragraph for arrival year 1995. We disaggregated the data into ten country categories, nine of which embraced most of the employment-based immigrants, leaving the tenth category for a residual. For each country group, we calculated a ratio of immigrant arrivals for 1994 through 1996 (described in the previous paragraph) to non-immigrant arrivals of temporary workers from the statistical yearbooks. We assumed this ratio to remain constant until 2002, and applied it, for each year from 1995 to 2002, to the non-immigrant series to produce an estimate of future adjustees arriving in each year.

The results presented in tables 1 to 5 and figures 1 to 5 are the results of these calculations under various assumptions about the future level of adjustment demand through 2032 (first approach) in combination with the second approach. We stress that, while the assumptions of these projections relate to the period from 2003 to 2032, the intent is not indicate what we "think" future immigration will be; rather, we are trying to get a better sense of what arrivals were in the late 1990s through 2002. Hence, the tables and figures do not show data past fiscal year 2002.

Some Assumptions

To establish a range of plausible projection results, we made four assumptions about the future demand for adjustments through 2032, coupled with the special treatment of employment data. In each case we added the number of new arrivals that were not subject to estimation procedures.

Projection 1. We assumed that demand for adjustment to legal permanent residence in all four categories of immigration (employment, family provision, special immigrants, and refugees) remained constant for 30 years at the average level observed in 2000 through 2002.

Projection 2. We assumed that demand for adjustment to legal permanent residence under employment or family provisions, by country of birth, increased arithmetically, for each decade, as they had increased between the three-year averages of 1990-1992 and 2000-2002. For special immigrants and refugees, we assumed constant demand for adjustments, as in Projection 1. This amounted to assuming the increase of the decade 1991-2001 for the two large categories would continue for three additional decades.

Projection 3. We matched the constant assumptions in projection 1 for all categories except employment. For employment adjustees, we calculated a ratio of projection 1 arrivals to temporary worker non-immigrant arrivals, and held this ratio constant until 2002.

Projection 4. We matched the assumptions in projection 2 for all categories except employment. Family-based adjustees were assumed to increase; special immigrant and refugee adjustments were assumed to remain constant. For employment adjustees, we calculated a ratio of projection 2 arrivals to temporary worker non-immigrant arrivals, and held this ratio constant until 2002.

Projections 1 and 3 were identical for all categories except employment adjustees, as were projections 2 and 4. Projections 3 and 4 were very nearly identical for employment adjustees, since the choice of projection 1 and 2 in establishing the 1994-1996 base date for calculating the immigrant-to-temporary-worker ratio was of little consequence.

Findings

Following are some major broad findings from this activity, and they generally hold for any choice of the four models. They are based on information in tables 1 to 5, illustrated in figures 1 to 5.

1. Legal immigration, defined as the arrival of persons who would ultimately become immigrants rose more-or-less steadily through 1985, with some fluctuations that could be linked to overseas upheavals that generated refugees.

2. From 1985 to 1993, the increase became numerically more rapid, and the trend cut across the major types of immigration.

3. From 1993 to 1997, the trend turned downward, largely as a result of the trend in employment-based immigration.

4. A rapid rise occurred in the late 1990s through 2001, fueled initially by employment immigration, and ultimately by the much larger category of family-based immigration.

5. The leap in employment-based immigration in the late 1990s was truly without precedent, and wins support from two different estimation strategies, although neither is wholly unassailable.

6. The categories of immigration based on special provisions and refugee immigration show fluctuations determined primarily by changes in immigration law, in the case of special provisions and world events, in the case of refugees. The high-level of special-provision migration in the early 1970s is a function of a 1990s-based definition of "special." In the 1970s, many immigrants were admitted to the US under a general provision making immigration available to anyone born in the western hemisphere.

Following are some findings that are more of technical interest.

1. Even under the most conservative projection, there is a substantial difference in the early 1990s between admissions to legal permanent residence and arrivals in the same year. The 1990s strategy, applied to population estimates, of using the former as a proxy for the latter (except for refugees) can be seen to have served us poorly. Even the number of arrivals each year observed as adjustees by 2002 far exceeds the number of adjustees in the same year. This makes a case (if one was needed) for the critical interpretation of administrative data in estimating population. The fact of an increase in adjustees, indicating an underestimate of arrivals, was quite palpable by 1994.

2. The choice between projection 1 & 3 and projection 2 & 4 in the series for family-based immigration is quite decisive. The higher-level assumption in 2 & 4 produces results in the late 1990s and through 2002 more consistent with the historical trend. The lower assumption brings the trend somewhat closer to the reasoning embodied in the last pre-2000 US Census Bureau population projections.

3. Even the more conservative trend in family-based immigration suggests a level of net international migration (net of emigration, but including unauthorized and temporary migration) peaking well over 1.5 million in 2001. The apparent drop in 2002 is undoubtedly related, directly or indirectly, to the terrorist attacks in the late summer of 2001.

4. Maintenance of the higher-level assumptions on future employment-based adjustments is going to require a change in the numerical limitations on these preference categories. Should the changes not occur, the likely result may be a combination of something like the Projection 1 result, but possibly with a corresponding increase in the temporary population, the unauthorized population, and family-based immigration. This may be mitigated by some emigration of workers frustrated by their inability to obtain a permanent visa, but this is unlikely to be the sole result.

One issue of relevance to the retrospective view of intercensal population trends in the 1990s is the distribution of immigrants through the decade. Following are some findings specific to the period from April 1, 1990, to March 31, 2000. To this end, tables 6 and 7 show data for one-year intervals ending March 31 of each year in this decade. Table 6 shows results from the "Projection 1" assumption; table 7 shows the same results for the "Projection 4" assumption. These two provide a range for the balance of arrivals between the two halves of the decade, among the four possibilities considered.

The distribution of immigrants over the intercensal decade is quite sensitive to the choice of models, because a substantial portion of the immigrants, especially in the family provisions category, had not arrived by 2002. Under the first projection (table 6,) the result is near balance between the two halves of the decade. The boost that began in 1998 supports the second half of the decade, while the "hill" that peaked in 1993 and 1994 supports the first half. On the other hand, if we consider the second quinquennium alone, the trend is clearly a rise. Under the fourth projection (table 7,) immigration is weighted to the second half of the decade, as the rise through 1998, 1999, and early 2000 is more than needed to compensate for the peak in 1993 to 1994.

The selected categories in table 6 and 7 are motivated by underlying factors perceived to be causal. The late-weighted character of the trend for India and China exhibits the rise in employment-based immigration related to the influx of H-1B temporary work visas in the latter part of the decade, linked to the growth of the technical industry. By contrast, the early-weighted character of employment immigration from the Philippines is linked more to the service industries. Within the family immigration category, the large percentage (more than 40 percent, in both models) made up of people from Mexico and Central America, is undoubtedly linked to the rise in the legitimating of earlier unauthorized migration through marriage to a U.S. citizen or permanent resident, or through the marriage of parents or siblings.

Vulnerability of the findings

A number of factors need to be considered that could affect these results.

1. The trend in the late 1990s and early 2000s is quite sensitive to the assumptions about the number of immigrants already in the United States who have not adjusted to legal permanent residency, as can be seen by the difference between Projections 1 or 2 and Projections 3 or 4.
2. The imputation of year of arrival to cases where it was unknown played a major role in the findings for people who immigrated (or would immigrate, by extension) under family provisions. The proportion of missing values in this category beginning in fiscal 1998, especially among those born in Mexico, was enormous. There could be a bias among the reporting cases in favor of more recent arrival, if memory recall is a factor.
3. There could be a bias, mostly in the overall level, resulting from the displacement between year of "census arrival," or arrival to census-defined residence. If this bias exists, its direction is predictable, since census arrival would normally precede last arrival. Under the observed increasing trend in arrivals, this would tend to suggest that the overall level might have been somewhat higher, since arrivals are being judged to arrive during a more recent, higher period of the trend than actually occurred. For the most part, this bias would not affect the trajectory very much, since it would also indicate that some future adjustments judged to be future arrivals should have been placed before 2002. It could be argued that the apparent drop around 1998 and subsequent sharp rise in employment immigration may have been artificially enhanced by a forward displacement of year of "census arrival" by nonimmigrant year of arrival.
4. The data series include only one fiscal year that was past the terror attacks of September 11, 2001. Our "steady state" assumptions about future demand for legal permanent resident status do not allow for the possibility that more temporary arrivals may opt to emigrate rather than become legal permanent residents in the future.

Considerations regarding the less measurable components in the 1990s

These data provide no direct evidence about the trend in net unauthorized (illegal) migration, net migration of temporary workers, net migration from Puerto Rico and the outlying areas, or the emigration of legal residents. However, they may assist in some speculation regarding how the trend in these other components may appear over the decade.

As with some other components, the term "net unauthorized migration" is misleading with respect to the concept at hand. A significant portion of legal, family-based migration to the U.S. is composed of people who entered the country illegally; in some cases, this is even evidenced by a nonimmigrant last entry class of "EWI" (entered without inspection). In a larger number of cases, the class was "B2," a tourist visa, and it is fair to assume that a goodly number of these were overstayed. People who either enter without inspection or enter under tourist visas and neither depart nor adjust to LPR status, belong to the class of unauthorized immigrants. If they subsequently become legal (e.g., by marrying a U.S. citizen,) any time before 2032 they are included as legal immigrants, hence not unauthorized. Hence, the "missing" component is *unauthorized migration, including illicit entry, overstayed nonimmigrant visas, but net of eventual departures, and excluding future conversions to legal permanent residence.* While it is quite possible that illicit entries and arrival under visas to be overstayed were weighted to the latter part of the 1990s, it is also quite evident that conversion to legal permanent residence was weighted late. It is possible that departures with intent to legally reenter were weighted toward early in the decade, before the passage of Section 245i of the immigration code rendered this activity unnecessary. In short, there is no convincing evidence from these data as to how this component was weighted within the decade.

Another component subject to definitional misunderstanding is the net migration of temporary residents, actually, the *arrival of temporary residents net of departures, and excluding those ultimately adjusting to legal permanent residence.* In this case, there is considerable statistical evidence that this component was weighted to the last part of the decade, given the sharp increase in the granting of H-1B visas, NAFTA arrivals, etc. far in excess of adjustments to permanent resident status or emigration. The claim that departures were low compared to entries is supported by the simple fact that the greatest departure risk was for cohorts that entered several years earlier, when the annual number of entries was much lower. The growth in the number of temporary workers admitted under the provisions of NAFTA is also supportive of this hypothesis.

Emigration, or return migration, of legal permanent residents, is most likely to occur to foreign-born immigrants within a few years of their arrival, hence, it is reasonable to surmise that the number would increase with the trend in legal immigration. Other things being equal, this component would tend to decrease the arithmetic increase in immigration, without affecting the geometric increase, if we imagine that the ratio of emigrants to the population most at risk remains constant. Parenthetically, emigration may have been a greater factor after the terrorist attacks of September 11, 2001.

The trajectories of the smaller components of net migration from Puerto Rico and the outlying areas and native emigration (extended to emigration of long-term foreign-born residents per the logic of the previous paragraph) are impossible to predict from these data.

Conclusion

While the trajectory of legal foreign-born immigration (here treated as the immigration of people who will ultimately become legal permanent residents) in the years since the mid-1990s is subject to legitimate dispute, we have presented some evidence that international migration continues to increase. A conservative assumption regarding future demand of legally temporary and unauthorized residents for legal immigration in the coming decades indicates a fluctuating trend during the 1990s, followed by a sharp increase through 2001. If we assume a continuation of the trend of the 1990s in the demand for legal adjustments, coupled with the assumption that administrative factors do not force temporary residents to depart who would otherwise adjust, there should be a sustained increase in permanent arrivals during the 1990s, peaking in 2001. Without precise evidence about the "difficult" components of migration to and from the United States, a post-2000 assumption that net migration to the US was between 1.0 and 1.5 million appears quite plausible.

References

U.S. Department of Homeland Security, *Yearbook of Immigration Statistics, 2002,* U.S. Government Printing Office, Washington, D.C. 2003.

U.S. Immigration and Naturalization Service, *Statistical Yearbook of the Immigration and Naturalization Service, 2001,* U.S. Government Printing Office, Washington, D.C., 2003, as well as earlier issues of this publication.

Table 1. Legal Immigrants, FY 1972 to FY 2002 (numbers in thousands)

Year ending Sept. 30	Total by FY of arrival				Actual arrivals by FY			FY of immigration
	Projection 1	Projection 2	Projection 3	Projection 4	All arrivals	Adjustees	New arrivals	
1972	398	398	398	398	397	100	297	383
1973	406	406	406	406	405	91	314	400
1974	417	418	417	418	416	99	316	390
1975	499	499	499	499	497	187	310	386
1976	451	452	451	452	450	132	318	399
1977	442	443	442	443	440	107	332	462
1978	526	527	526	527	524	153	371	601
1979	549	550	549	550	546	214	332	460
1980	684	686	684	686	681	341	339	530
1981	657	659	657	659	653	274	379	596
1982	561	564	561	564	556	241	315	594
1983	551	555	551	555	544	207	337	560
1984	574	578	574	578	564	219	345	544
1985	605	611	605	611	592	236	356	570
1986	641	648	641	648	623	246	376	602
1987	658	667	658	667	634	247	387	594
1988	718	730	718	730	686	308	378	606
1989	822	838	822	838	778	375	403	602
1990	898	918	898	918	837	401	437	656
1991	921	947	921	947	843	401	442	702
1992	1,005	1,038	1,005	1,038	910	398	512	809
1993	1,029	1,072	1,029	1,072	915	380	536	888
1994	1,003	1,056	1,003	1,056	867	377	491	787
1995	938	1,002	934	997	776	396	380	721
1996	940	1,019	950	1,028	747	326	421	913
1997	900	996	914	1,008	669	289	380	781
1998	892	1,009	951	1,065	607	250	357	653
1999	1,004	1,146	1,059	1,194	647	246	402	646
2000	1,070	1,238	1,099	1,256	637	229	407	849
2001	1,085	1,288	1,086	1,269	576	165	411	1,064
2002	1,027	1,263	1,019	1,223	462	78	384	1,063

Table 2. Employment Provision Immigrants, FY 1972 to FY 2002
(numbers in the thousands)

Year ending Sept. 30	Total by FY of arrival				Actual arrivals by FY			FY of immigration
	Projection 1	Projection 2	Projection 3	Projection 4	All arrivals	Adjustees	New arrivals	
1972	33	33	33	33	33	12	21	34
1973	34	34	34	34	34	13	20	28
1974	34	34	34	34	34	13	21	33
1975	32	32	32	32	32	12	20	30
1976	31	31	31	31	31	14	17	27
1977	32	32	32	32	32	17	14	26
1978	37	37	37	37	37	23	14	33
1979	47	47	47	47	47	27	20	40
1980	49	49	49	49	49	27	22	47
1981	53	53	53	53	53	28	25	48
1982	55	55	55	55	55	29	27	57
1983	56	56	56	56	56	28	28	62
1984	55	56	55	56	55	28	28	52
1985	58	59	58	59	58	29	30	53
1986	64	64	64	64	63	32	32	57
1987	68	69	68	69	68	32	36	58
1988	73	73	73	73	72	35	37	59
1989	77	77	77	77	76	38	37	58
1990	88	89	88	89	86	48	38	58
1991	105	106	105	106	102	61	41	60
1992	123	124	123	124	119	67	52	117
1993	122	124	122	124	117	64	53	119
1994	115	118	115	118	107	61	46	102
1995	96	100	91	95	85	64	22	81
1996	94	100	104	109	80	56	24	117
1997	107	115	121	127	88	68	19	87
1998	79	89	138	145	53	38	15	77
1999	93	108	148	156	58	43	15	56
2000	140	159	168	176	93	71	22	106
2001	198	226	199	207	128	87	41	179
2002	198	238	190	198	99	59	40	174

Table 3. Family Provision Immigrants, FY 1972 to FY 2002
(numbers in thousands)

.

Year ending Sept. 30	Total by FY of arrival		Actual arrivals by FY			FY of immigration
	Projection 1/3	Projection 2/4	All arrivals	Adjustees	New arrivals	
1972	198	198	197	49	148	185
1973	218	218	217	49	168	208
1974	225	225	224	51	172	209
1975	220	220	218	52	166	197
1976	237	238	236	54	181	217
1977	269	270	267	62	205	246
1978	332	333	330	82	248	326
1979	384	386	382	96	286	360
1980	390	392	388	95	293	378
1981	399	401	395	97	298	382
1982	389	392	384	98	286	378
1983	415	418	409	100	309	395
1984	437	441	428	112	317	399
1985	462	467	451	125	326	421
1986	492	499	477	133	344	440
1987	494	502	475	127	348	434
1988	520	531	495	160	335	424
1989	568	583	535	185	350	438
1990	597	616	553	196	357	449
1991	608	633	551	188	363	455
1992	655	687	584	176	408	505
1993	687	728	603	165	438	540
1994	670	720	569	168	401	501
1995	620	680	501	179	322	465
1996	669	743	527	178	348	604
1997	635	724	466	146	320	535
1998	624	731	426	124	301	476
1999	688	816	455	113	342	485
2000	700	849	430	95	336	588
2001	686	862	386	57	329	679
2002	638	835	318	16	303	676

Table 4. Special Provision Immigrants, FY 1972 to FY 2002
(numbers in the thousands)

Year ending Sept. 30	Total by FY of arrival	Actual arrivals by FY			FY of immigration
		All arrivals	Adjustees	New arrivals	
1972	134	134	14	120	137
1973	127	127	11	116	133
1974	125	125	10	115	123
1975	123	123	9	115	122
1976	114	114	5	109	117
1977	108	108	3	104	112
1978	101	101	2	99	111
1979	15	15	1	14	16
1980	15	15	1	13	18
1981	57	57	1	55	60
1982	4	4	1	2	3
1983	3	3	2	0	1
1984	5	4	4	0	0
1985	9	7	7	0	1
1986	15	12	12	0	1
1987	21	17	14	3	6
1988	35	29	23	6	10
1989	54	44	37	7	11
1990	79	66	37	29	34
1991	74	57	35	22	32
1992	83	64	30	35	55
1993	79	58	24	34	93
1994	84	60	19	40	60
1995	78	51	15	36	55
1996	89	60	12	48	62
1997	81	50	10	40	52
1998	81	49	8	40	47
1999	92	57	12	45	63
2000	102	63	14	48	90
2001	94	51	10	40	98
2002	91	44	3	41	87

Table 5. Refugee Immigrants, FY 1972 to FY 2002
(numbers in thousands)

Year ending Sept. 30	Total by FY of arrival	Actual arrivals by FY			FY of immigration
		All arrivals	Adjustees	New arrivals	
1972	33	33	24	9	28
1973	27	27	18	9	31
1974	33	33	25	8	26
1975	123	123	114	10	36
1976	69	69	59	10	38
1977	33	33	24	9	78
1978	56	56	46	10	132
1979	102	102	90	12	44
1980	230	230	218	11	87
1981	148	148	148	0	107
1982	113	113	113	0	157
1983	77	77	77	0	103
1984	77	76	76	0	92
1985	76	76	76	0	95
1986	71	70	70	0	105
1987	75	74	74	0	97
1988	91	91	90	0	113
1989	124	123	114	9	96
1990	134	133	120	13	114
1991	134	133	117	16	156
1992	144	143	125	17	133
1993	141	138	127	11	136
1994	135	131	128	3	124
1995	144	139	138	1	119
1996	88	80	79	1	130
1997	77	65	64	1	107
1998	108	80	79	0	53
1999	130	77	77	0	42
2000	129	50	49	1	65
2001	106	11	11	0	108
2002	100	1	0	0	126

Figure 1
Legal Immigrants, FY 1972 to FY 2002

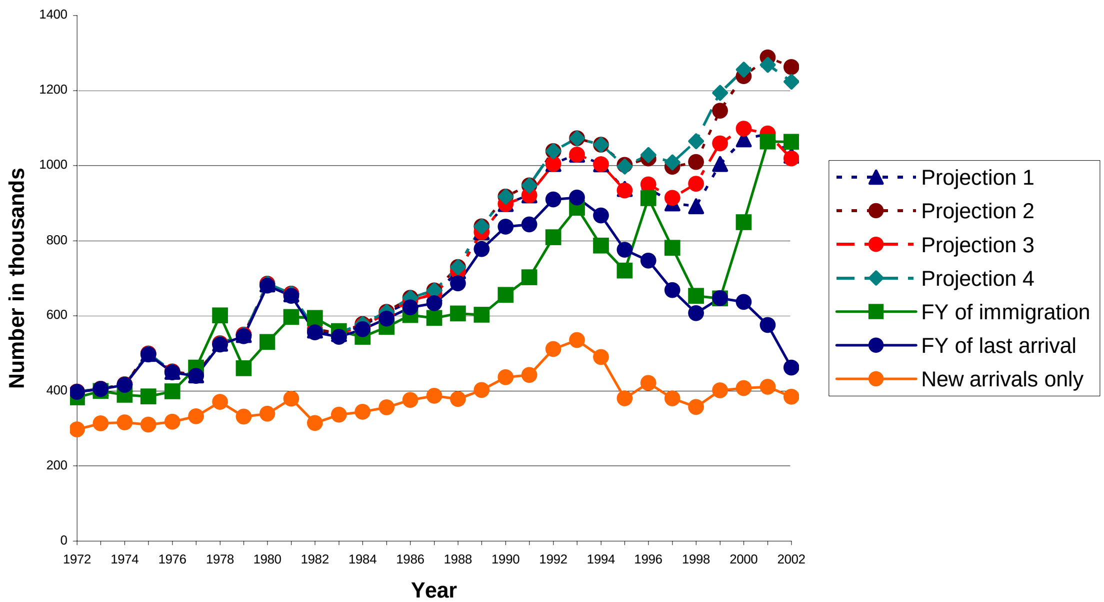


Figure 2
Employment Immigrants, FY 1972 to FY 2002

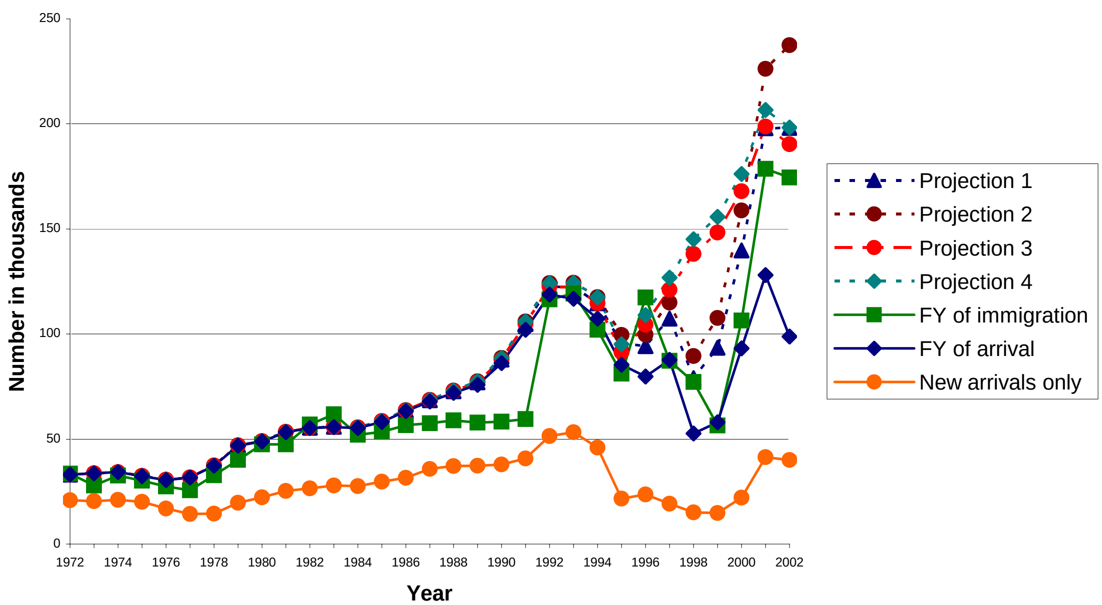


Figure 3
Family Provision Immigrants, FY 1972 to FY 2002

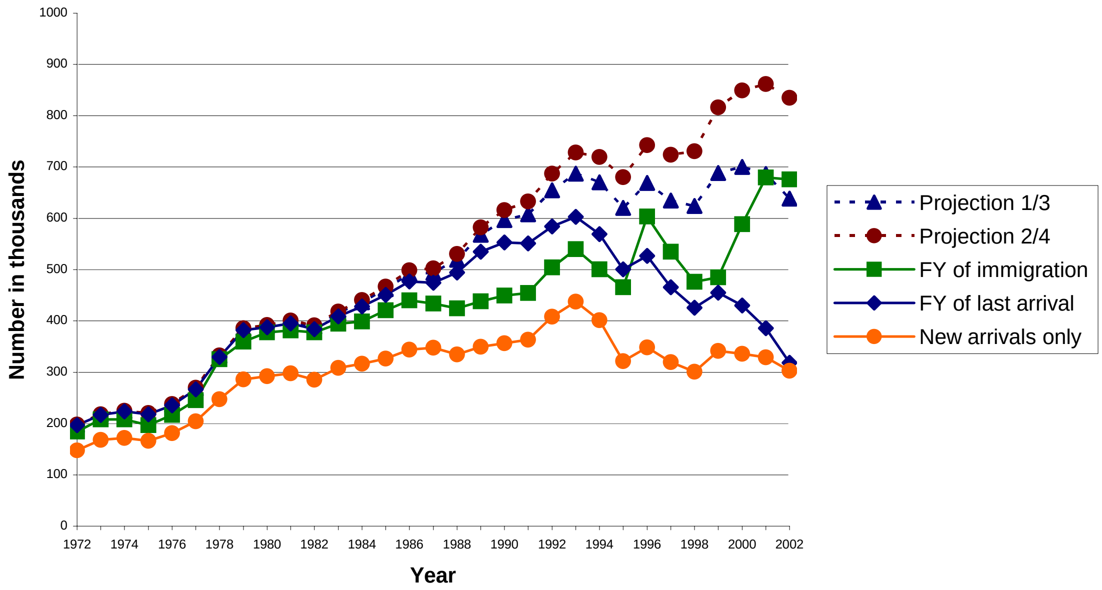


Figure 4
Diversity and Special Provision Immigrants, FY 1972 to FY 2002

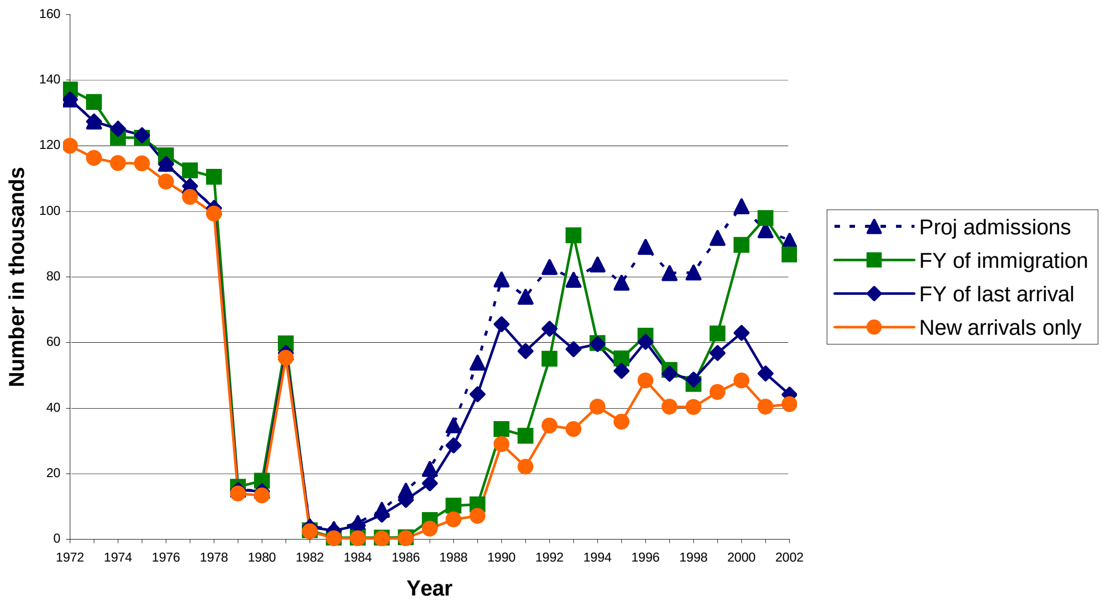
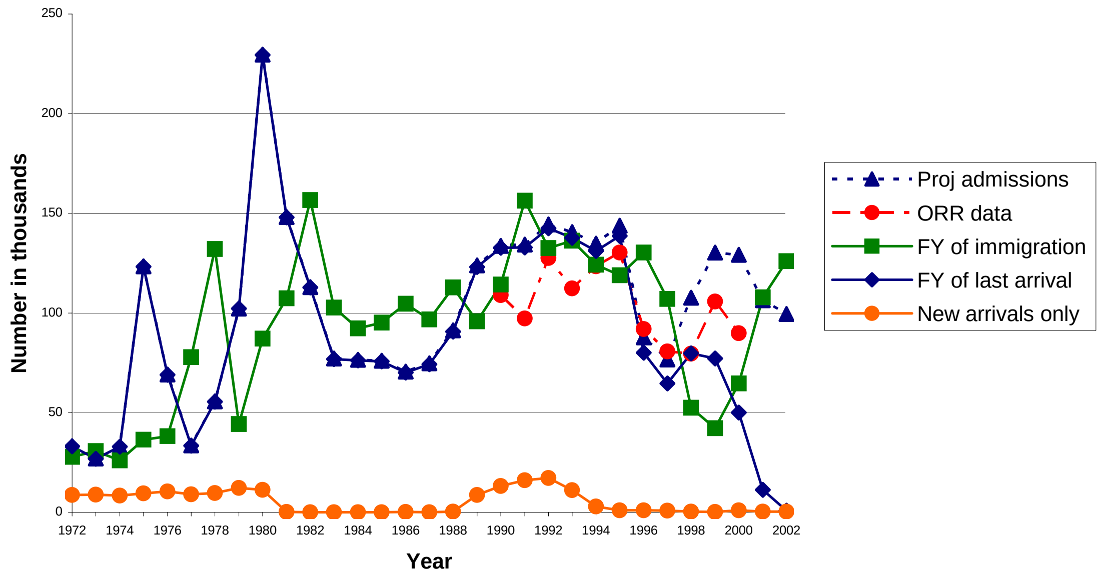


Figure 5
Refugee Immigrants, FY 1972 to FY 2002



This page intentionally left blank.

Forecasting Age Distribution Curves

Tucker McElroy and William Bell
U.S. Census Bureau

Disclaimer This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress. The views expressed on statistical, methodological, technical, or operational issues are those of the authors and not necessarily those of the U.S. Census Bureau.

1 Introduction

The construction of population projections typically involves forecasting of several demographic components of change—births, deaths, immigration, and emigration—each with considerable demographic detail (e.g., age-race-sex for single years of age and several race/origin groups). This presents a forecasting problem of high dimension. Along with producing forecasts with this level of detail it is generally desired that forecasts reflect an age pattern consistent with age patterns observed in historical data, which typically are quite pronounced. Also, these properties are sought in long-term forecasts (say, 50 to 75 years ahead) though based on historical time series that may be of relatively moderate length, often less than the desired forecast horizon. This setting presents some challenging methodological problems for forecasting.

In this paper we investigate the application of time series methods to the forecasting of data on legal immigration to the U.S. The available data consist of historical estimates of several categories of legal immigration from 1972 to 2002. We focus, for illustration, on the category of "Hispanic employment immigrants," which denotes legal immigrants and their dependents arriving from countries of predominantly Hispanic race/origin, who have immigrated for reasons of employment. The historical data provide estimates for each category of the number of immigrants (aggregated as males plus females) for single years of age from 0 to 99 and then 100+. The same level of detail is desired in the forecasts. This data is easily transformed into a suite (or time series) of 31 age distributions simply by dividing the number of persons of a given age in a given year by the total number of persons in that particular year. If these age distributions can be forecasted ahead, they can then be multiplied by forecasts of total Hispanic employment immigrants to obtain forecasts of immigrants by age. Forecasts are desired up to 50 years ahead.

As noted above, the creation of such long-term forecasts from a small data set presents many challenges. Clearly, any forecasts should themselves be distributions, i.e., for each forecast year, the forecasted values at each age should be non-negative and should sum to one over ages. In addition, one wishes to model and forecast the main features of the distribution (such as the modes) while ignoring small perturbations that appear to be just "noise" in the historical data. This amounts to selecting a parsimonious approximation of the age curves that can be readily forecasted, while remaining true to the data's structure. Finally, since the forecast horizon is so long, it is desirable to impose constraints on the forecast methodology to outlaw implausible distortions in the ultimate forecasts of the age-distributions. These considerations have led us to consider a combination of logistic transformation, principal components analysis, time series modelling, cubic spline smoothing, and Bayesian forecast attenuation. The combination of these statistical methods represents an innovative approach to achieving our forecasting objectives.

This paper focuses on the above methodologies and their application to the Hispanic employment immigrants data. Section 2 discusses the employment immigrant data that provides the basis for our forecasts, and Section 3 discusses the details of our methods. Section 4 discusses the application of these methods to the forecasting of Hispanic employment immigrants. Section 5 gives a summary of the approach, and an appendix contains a proof of a mathematical result used in Section 3.

Readers are cautioned that the focus in this paper is on investigation of methods for dealing with the forecasting problem, with the Hispanic employment immigrants data used for illustration. The actual forecast results are, at this stage, experimental, and should not be regarded as any sort of official Census Bureau projections.

2 Historical Data on Legal Immigration

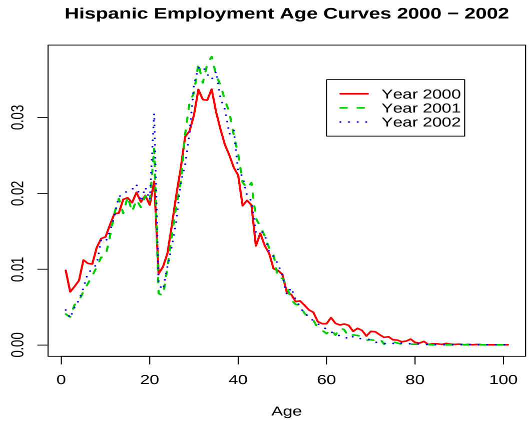
The first step in forecasting legal immigration is to develop corresponding historical data. This task was undertaken by Hollmann (2004), who developed historical estimates of legal immigrants by age for four general types of immigrants (family, employment, refugee, and special) for each of four race/ethnic groups, nominally labelled Hispanic, Black, White, and Asian (the last three groups all referring to the non-Hispanic population). The four race/ethnic groups do not actually correspond to reported races of individual immigrants (as such detailed information is not available), but rather to a classification of countries of origin into four groups according to the predominant race/ethnicity that immigrants from these countries reported in the 2000 census. We focus here on forecasting the time series of "Hispanic employment immigrants," which means immigrants from countries of predominantly Hispanic race/ethnicity *and their dependents* by virtue of employment provisions in immigration law. (The inclusion of dependents of the actual employment immigrants means that the age range of the employment immigrants as estimated by Hollmann range from 0 to 100 and above.) Essentially the same considerations apply to modelling and forecasting time series of employment immigrants for the other three race/ethnic groups, and similar considerations apply to forecasting the "family immigrants" series as well. (Family immigrants are family members of previous immigrants, admitted via the family provisions of immigration law.) The refugee and special immigrant data are somewhat different, showing more erratic behavior.

The data sources available for constructing estimates of legal immigration have some limitations that translate into errors in the historical estimates. Particularly worth noting here is the fact that a person does not become an "immigrant" until he or she has achieved legal permanent residence status in the U.S., and when this occurs the person is deemed to have immigrated as of the date of their last entry to the U.S. This "date of last entry" is often several years before the achievement of permanent residence status, implying that at any time there is a substantial pool of people already in the U.S. who will achieve immigrant status in the coming years and then be assigned as immigrants to the current year or to some past year (whenever they last entered the U.S.). To account for this pool of potential immigrants Hollmann (2004) was thus forced to impute substantial numbers of immigrants into the last years of his historical estimates. Such a process unavoidably incurs errors, and since these errors increase towards the end of the series, these errors have important implications for forecasting. In particular, although we shall forecast the series taking the data "as is," we must keep in mind that capturing every detail of the age-pattern of the historical data, and reflecting this in the forecasts, is not necessarily justified. Some irregularities in the data, particularly in the last years, may be due to data errors.

3 Methodology

The forecasting problem at hand poses four main challenges: preserving the distribution property, dealing with the high dimensionality of the data, projecting "signal" while dispensing with noise, and attenuating the excesses of forecasts over an extended forecast horizon. One approach to handling the first three issues at once is to fit a parametric family of distributions to the data, such as a gamma curve, and then forecast the curve parameters. Rogers and Castro (1981) describe use of this approach with migration data. Bell (1997) discusses relative advantages and disadvantages of this approach, in general. The primary disadvantage is the difficulty in finding a curve that depends on a small number of parameters yet provides an adequate approximation to the data. For the Hispanic employment immigrants series, and for the employment immigrant series of the other three race/ethnic groups, the data are bimodal (see Figure 1 below) with some irregularities that can be regarded as "noise," but with others that represent real features of the data that should be preserved in forecasting. This means that finding a simple parametric curve to fit to this data would be difficult, and an adequate approximation would probably require piecing together two or more curves resulting in a moderate set of parameter series to forecast. We use other techniques to address the four problems noted. To preserve the distribution property, Section 3.1 discusses how we use a generalization of the logistic transformation proposed by Aitchison (1987). To deal with the high dimensionality of the data we use the principal components approach (PCA) proposed by Bozik and Bell (1987) – see also Bell 1997, which provides an excellent approximation to the data in terms of very few "parameters." In fact, as discussed

Figure 1: Age distribution curves for Hispanic Employment, 2000 – 2002



in Section 3.2, we use an approximation based on one principal component for the mean corrected data, a version of the approach used by Ronald Lee and his colleagues (e.g., Lee and Carter 1992). The one principal component approximation succeeds in capturing the features of the data used here that appear to be real, at least to the extent that appears important for forecasting. As noted in Section 3.3, however, the approximations to the age distributions include some irregularities that are exacerbated by long-term forecasting. To remove these irregularities we smooth the principal component vector and the mean age distribution curve using cubic splines (Section 3.4). Another problem with long-term forecasts noted in Section 3.3 is that historical trends in these data, extrapolated indefinitely into the future, yield implausible results. Section 3.5 discusses how, with a nominally Bayesian approach, we bring in "prior information" about the plausible range for future values of the series to "attenuate" the point forecasts to prevent implausible results.

3.1 Logistic Transformations

We require a method that ensures that our forecasted distributions will also be distributions, i.e., curves that are non-negative with unit integral. A generalization of the basic logistic transformation to a multivariate context achieves this objective. Let the data be given as x_{it} for ages $i = 0, \ldots, 99, 100$ (the last age actually being 100+), and let $t = 1, \ldots, 31$ represent the years 1972 to 2002. For each year, we then obtain ratios

$$r_{it} = \frac{x_{it}}{\sum_{j=0}^{100} x_{jt}} \qquad i = 0, \ldots, 99, 100,$$

so that for each t, r_{it} defines an age distribution. Whatever forecasting methods we employ, we wish to ensure that the result is also an age distribution. The generalized logistic transformation (Aitchison 1987), when inverted, will constrain the forecasts to be positive and sum to one across ages. For each t, let

$$\gamma_{it} = \log\left(\frac{r_{it}}{r_{100,t}}\right) \qquad i = 0, \ldots, 99,$$

which defines a 100 by 31 data matrix. This transformation is reversed via

$$\begin{aligned} r_{it} &= \frac{e^{\gamma_{it}}}{(1 + \sum_{j=0}^{99} e^{\gamma_{it}})} \qquad i = 0, \ldots, 99 \\ r_{100t} &= \frac{1}{(1 + \sum_{j=0}^{99} e^{\gamma_{it}})}. \end{aligned}$$

The transformed data γ_{it} can take on any real number value, so there is no constraint on the forecasts of the γ_{it}, but the inverse transformation guarantees that we obtain an age distribution for each time t. One proviso is that $r_{it} > 0$ is necessary for the transformation. There are actually many zeros in the historical Hispanic employment immigrant data x_{it}. Since the counts x_{it} tend to be fairly large when nonzero, however, we modified the data by adding one to each x_{it}.

3.2 Principal Components Analysis

The multivariate series γ_{it} has high dimension (100) but is relatively short (31 years). To simplify matters we reduce the dimension of the forecasting problem by using the principal components analysis (PCA) approach. The general approach was proposed in Bozik and Bell (1987) and is discussed further by Bell (1997). Lee and Carter (1992) used a version of PCA based on a one principle component approximation to mean corrected data of log U.S. mortality rates. Here we also use a one principle component approximation to the Hispanic employment immigrant data with mean correction, where the data are the logistically transformed ratios γ_{it}. Irregularities in the migration data (note Figure 1), coupled with our knowledge that errors in the historical data grow towards the end of the series, suggest use of a low-dimensional PCA approximation. Also, the need to forecast age distributions for each of 16 groups (four immigrant categories for each of four race/ethnic groups), along with the total number of immigrants for each of these groups (which we don't consider here), mandates that we attempt to minimize the number of principle components used. Bell (1997) notes that mean correcting the data, or alternatively subtracting from the data each year the values from the last year of data, is helpful when using a low-dimensional PCA approximation.

Let $\gamma_t = (\gamma_{0t}, \ldots, \gamma_{99t})'$ be the column vector of the transformed r_{it} for year t. We subtract a baseline curve defined as a single summary measure of all the curves; this could be the mean curve (the average over time) or the last curve γ_{31}, for example. We will model the deviations from the chosen baseline, forecast the deviations ahead in time, and add back the baseline curve. Here we use the mean curve $\bar{\gamma} = \sum_{t=1}^{31} \gamma_t/31$ as the baseline. To apply the PCA approach to the centered curves $\gamma_t - \bar{\gamma}$, we compute the sum-of-squares and cross products matrix

$$S = \sum_{t=1}^{31} (\gamma_t - \bar{\gamma})(\gamma_t - \bar{\gamma})'$$

and determine its eigenvalue-eigenvector decomposition

$$S = \Lambda D \Lambda',$$

where D is diagonal consisting of the eigenvalues of S (by convention, ordered from greatest to smallest) and Λ has columns consisting of the corresponding orthonormal eigenvectors of S. Such a decomposition always exists, because S is non-negative. For any $J \leq 30$, the submatrix Λ_J consisting of the first J eigenvectors of Λ is of dimension 100 by J. The J-dimensional principal component approximation is obtained by regressing, for each year, the data vector $\gamma_t - \bar{\gamma}$ on Λ_J. The resulting regression coefficients are

$$\beta_t^J = (\Lambda_J' \Lambda_J)^{-1} \Lambda_J' \gamma_t. \tag{1}$$

Ordinarily in principle components $\Lambda_J' \Lambda_J$ equals the identity matrix, so that we get the principle components transformation $\beta_t^J = \Lambda_J' \gamma_t$, (Mardia, Kent, and Bibby 1979) but we will later smooth the columns of Λ_J destroying their orthonormality, so we retain the general expression as given above. The corresponding approximation of γ_t using J principal components is then

$$\hat{\gamma}_t^J = \bar{\gamma} + \Lambda_J \beta_t^J. \tag{2}$$

When $J = 0$, we set $\hat{\gamma}_t^0 = \bar{\gamma}$ by convention. Higher values of J improve the approximation of γ_t, but at the cost of a higher dimensional problem in forecasting β_t^J. Here we wish to keep J very low, but we still wish to accurately capture the main features of the curves.

In order to forecast the curves using PCA, we generate forecasts of the principal component series β_t^J. Given forecasts $\hat{\beta}_{31+h}^J$ for $h = 1, 2, \ldots, 50$ (see below for time series forecasting methods), we substitute into (2) to obtain forecasted curves

$$\hat{\gamma}_{31+h}^J = \bar{\gamma} + \Lambda_J \hat{\beta}_{31+h}^J.$$

Here we focus on the case that $J = 1$. A victory of this method was the excellent level of approximation obtained even for $J = 1$. As described in Mardia, Kent, and Bibby (1979), the ratio of the first J eigenvalues of S to the trace of S measures the proportion of total variation in the data that is explained by the first J principal components. Here this proportion was .92 for $J = 1$, and adding a second principal component only increases this proportion to .95. We thus set $J = 1$ so β_t^1 is a univariate time series that can be easily modelled and forecast.

3.3 Forecasting Principal Components

The one principal component series β_t^1 can be modelled and forecasted using univariate time series techniques. For our application, a simple ARIMA (Box and Jenkins 1976) model was deemed sufficient considering that the data limitations (length of series and errors in the historical data) did not warrant a very refined treatment. For the Hispanic employment immigrants data, we found that a decent fit was given by the ARIMA$(1, 1, 0)$ model with a trend (slope) constant:

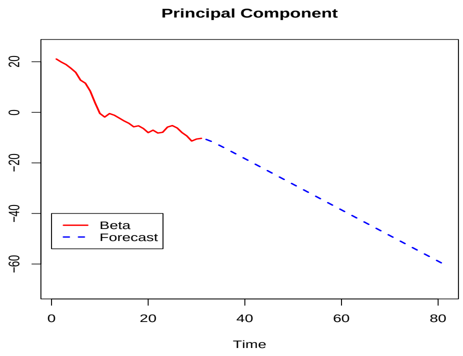
$$(1 - \alpha B)\left[(1 - B)\beta_t^1 - \mu\right] = \epsilon_t \tag{3}$$

where ϵ_t is a white noise sequence of variance σ^2, μ is a slope constant estimated via generalized least squares regression given α, the autoregressive parameter. Using the modelling capabilities of the program X-12-ARIMA (U.S. Census Bureau 2002), we obtained $\hat{\mu} = -1.01$, $\hat{\alpha} = .46$, and $\sigma^2 = 1.71$. The interpretation of the downward trend of β_t^1 is not obvious, since its product with Λ_J forms the approximation to γ_t. Since the model for β_t^1 is nonstationary, the forecasts will decrease in an unbounded fashion, possibly resulting in large deviations from the original age curves. In practice, we found this to be the case, especially when viewed 50 years out. Below in Figure 2 is a graph of β_t^1 together with 50 forecasts. The linear forecast pattern of β_t^1 creates implausible age distribution curves in the latter years of the forecast period. Empirically, we observed the following phenomena:

- Modes of the curve become increasingly high and narrow, to an unfeasible degree.
- Modes migrate to the older age groups, creating untenable results.
- Small noise perturbations in the original curves become large spikes in the long-term future.

These are all caused by the long-term behavior of the forecasts of β_t^1, which accentuate small pertur-

Figure 2: Plot and forecasts of principal component series for Hispanic Employment



bations in the principal components approximation to the age distributions into much larger perturbations in the forecasted age distributions 50 years out. On the other hand, the historical pattern of the series β_t^1 shows a steady downward trend that clearly should be reflected in the short-term forecasts. The problem is thus that the model (3) provides a reasonable description of the historical behavior of the series, and plausible short-term forecasts, but implausible long-term forecasts. Furthermore, *any* model or forecasting procedure that extrapolates the historical downward trend of β_t^1 indefinitely into the future will produce implausible long-term forecasts.

The first two problems noted above can be resolved by attenuating the forecasts of β_t^1. While simply truncating the forecast at a pre-specified limit would address the problems, this would yield somewhat strange results around the time point of truncation. Instead, we achieve a gradual attenuation of the forecasts by imposing a prior probability density function in the forecast period, as described Section 3.5 below. The amplification of the irregularities in the age distributions that appear to be "noise" can be addressed also through smoothing of the mean curve $\bar{\gamma}$ and Λ_1. When $J = 1$, the principal component approximation for the age distribution curve γ_t is $\bar{\gamma} + \Lambda_1\beta_t^1$. Since β_t^1 is a scalar (univariate) time series, if both $\bar{\gamma}$ and Λ_1 are smooth over age then all the forecasted age distributions will be smooth. Actually, we can selectively smooth $\bar{\gamma}$ and Λ_1 to maintain any non-smooth features of the age distributions that appear to be real while smoothing away those that appear to be "noise." In Section 3.4 below, we investigate the use of cubic spline smoothers for this purpose.

The other product of the ARIMA forecasts is a standard error at each future time point. These standard errors increase as a function of the forecast horizon h, and are used in the Bayesian methods discussed below. As discussed in Bozik and Bell (1987), it is possible to translate such standard errors on forecasts of β_t^1 (actually, we need to use the full variance-covariance matrix of the forecast errors) into standard errors for the forecasted age distribution curves, but we do not pursue this calculation in this paper.

3.4 Cubic Spline Smoothers

Smoothing via cubic splines is discussed in Hastie and Tibshirani (1990). The basic concept is to fit a cubic polynomial to every pair of consecutive data points on the desired curve. This only provides two constraints for 4 unknowns; the remaining constraints are obtained from smoothness conditions between adjacent cubics. For a smoother curve, one can leave out certain data points, or adjust the smoothing parameters which govern the goodness of fit in the polynomial fitting.

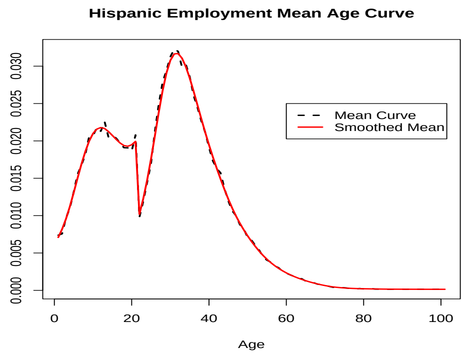
In our implementation, we used the *smooth.spline* function in the *R* programming language (R Development Core Team 2004), with a smoothing parameter chosen such that small noise perturbations were eliminated, while preserving the major features of the curve. For our applications, we spline smoothed both the mean curve $\bar{\gamma}$ and the Λ_1 curve in the PCA decomposition. It is necessary for both of these curves to be smooth in order to ensure that forecasts are also smooth. However, it is important to avoid oversmoothing, and thus some care in the selection of smoothing parameters is required. For $\bar{\gamma}$, we used an automatic choice of the smoothing parameter, whereas for Λ_1 we used the value of 0.5. We made our choice for Λ_1 based on aesthetic considerations.

One complication associated with using the spline smoother is that certain features believed to be intrinsic to the curve could get identified as noise, due to their structure. For example, the sharp drop in the Hispanic employment immigrants age distribution at age 22 is not an anomaly due to noise, but is a real feature of the data. Specifically, this drop reflects the fact that the "employment immigrants" below age 21 are primarily dependents of the actually employed immigrant, while few young adults over age 21 qualify as dependents. Hence, a sharp drop-off in dependents of employment immigrants between ages 20 and 22 results in the observed large drop in the aggregate (actually employed immigrants plus dependents) age distribution. On the other hand, the value at age 21 reflects a combination of dependents and actually employed immigrants, with a boost relative to the slightly lower ages due to age misstatement (an incentive to report oneself as 21 to qualify). Thus, we may wish to smooth out somewhat the peak at 21 as noise, but we wish to retain the drop at age 22 in the forecasts. Cubic splines applied to the full age distribution will automatically smooth out the drop at age 22. In order to force the preservation of this feature, we spline smooth the curve in two separate applications – up to age 21 and then 22 and up. Figure 3 below demonstrates the spline smoother on the transform of the mean age curve $\bar{\gamma}$. Note that the spline smoothing takes place on the logistic transformed data γ_t; when the transform is reversed, as depicted in Figure 3, the resulting curve is still smooth, but satisfies the properties of an age distribution, as discussed in Section 3.1 above.

3.5 Bayesian Attenuation

Bayesian methods have been used in time series modelling and forecasting – see, e.g., Thompson and

Figure 3: Mean age curve and smoothed mean age curve for Hispanic Employment



Miller (1986). Below we formulate a fairly general approach to forecasting with *a priori* beliefs about the future. Let x denote an n-dimensional vector of observed data, and let y denote an unknown future value (this can be generalized to a multivariate scenario, but here y is scalar for simplicity). Let Model 1 portray the scenario that x and y are jointly normal:

$$\begin{bmatrix} y \\ x \end{bmatrix} \sim \mathcal{N}(\mu, \Sigma), \tag{4}$$

where $\mu = [\mu_y, \mu_x]'$ and Σ is the covariance matrix. Then we can easily write down the joint multivariate normal density $p_1(y, x)$, as well as the conditional density $p_1(y|x)$. Model 2 specifies bounds for y to be the numbers L and U (with $L < U$); this could be extended to form bounds L_i and U_i for each x_i, but we won't pursue this here. So the joint density $p_2(y, x)$ is essentially just a truncated version of $p_1(y, x)$, and is given by:

$$\begin{aligned} p_2(y, x) &= p_1(y, x) 1_{\{L<y<U\}} / c \\ c &= \Pr{}_1[L < y < U] \end{aligned} \tag{5}$$

where $\Pr_1(\bullet)$ denotes the probability of the event computed as if Model 1 were true, and $1_{\{L<y<U\}}$ is the indicator of the event that y is between L and U. Roughly speaking, we wish to combine time series forecasts with prior assumptions about what values are allowed. This is done by determining $p_1(y|x)$

purely from our time series model, and combining this with predetermined limits L and U for the future value y . The forecast conditional on (i) the data, and (ii) our *a priori* assumptions about the future values, has density $p_2(y|x)$, the conditional density in Model 2. Proposition 1, whose proof is in the appendix, gives a formula for this density.

Proposition 1 *The conditional density in Model* 2 *is given by*

$$p_2(y|x) = \frac{p_1(y|x)1_{\{L<y<U\}}}{\Pr_1[L<y<U|x]}.$$

The minimum mean square error prediction of y from x in Model 2 *is given by*

$$E_2(y|x) = \int y\, p_2(y|x)\, dy = \frac{\int_L^U y\, p_1(y|x)\, dy}{\Pr{}_1[L<y<U|x]}.$$

Remark 1 *The quantity* $\Pr{}_1[L < y < U|x] = \int_L^U p_1(y|x)\, dy$ *forms the appropriate normalization for* $p_2(y|x)$ *and is easily computed, since* $p_1(y|x)$ *is known.*

We apply Proposition 1 to the forecasts of the principal component series β_t^1. Here $(\beta_1^1, \beta_2^1, \ldots, \beta_{31}^1)$ plays the role of the data vector x, and any particular future value β_{31+h}^1 is our y. Our ARIMA model will provide $\hat{y} = E_1(y|x)$ as well as the variance V of $\hat{y}$, which does not depend on x. Thus

$$p_1(y|x) = \frac{1}{\sqrt{V}}\ phi\left(\frac{y-\hat{y}}{\sqrt{V}}\right)$$

where ϕ denotes the standard normal density, and hence

$$p_2(y|x) = \frac{\phi\left(\frac{y-\hat{y}}{\sqrt{V}}\right) 1_{\{L<y<U\}}}{\sqrt{V}\left(\Phi(\frac{U-\hat{y}}{\sqrt{V}}) - \Phi(\frac{L-\hat{y}}{\sqrt{V}})\right)}$$

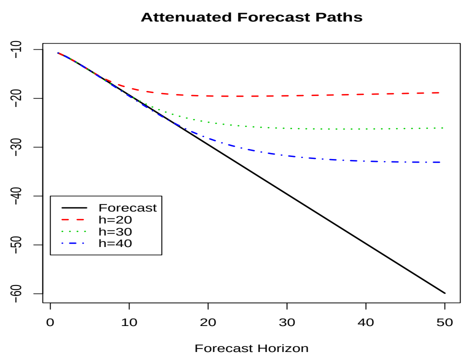
where Φ is the cumulative distribution function of the standard normal random variable. Finally, by integrating against y we obtain the optimal estimate under Model 2:

$$E_2(y|x) = \hat{y} - \sqrt{V}\,\frac{\phi(\frac{U-\hat{y}}{\sqrt{V}}) - \phi(\frac{L-\hat{y}}{\sqrt{V}})}{\Phi(\frac{U-\hat{y}}{\sqrt{V}}) - \Phi(\frac{L-\hat{y}}{\sqrt{V}})}. \qquad (6)$$

When a particular forecast $\hat{y}$ from the unconstrained Model 1 is extremely high or low, the Bayesian attenuation modifies this prediction towards the midpoint $(L+U)/2$.

The next practical question is "how should one choose L and U so as to best attenuate the forecasts?" The values of L and U have no obvious interpretation, and therefore some amount of trial and error is required. Figure 4 below plots the principal component forecasts $\hat{\beta}_{31+h}^1$ for various choices of L. U was set to the threshold of zero (greater than the maximum of all the forecasts), since no attenuation in this direction is necessary. Some "data-driven" choices for L are the various forecasts themselves, e.g., $L = \hat{\beta}_{31+h}^1$ for forecast horizons $h = 20, 30, 40$.

Figure 4: Forecast paths of Hispanic Employment with Bayesian attenuation. The original ARIMA forecast is plotted, together with attenuated forecasts with lower limit L is equal to the ARIMA forecast at horizon $h = 20, 30, 40$.



4 Application to Hispanic Employment Immigration

Here we summarize the procedures used on the Hispanic employment immigrants age distribution data.

1. Transform the data: add 1 to all data x_{it}, and normalize to form ratios. Apply the generalized logistic transformation, obtaining γ_t.

2. Compute the mean curve $\bar{\gamma}$. This is then spline-smoothed, taking any jags or special data features into account. We used age 21 as a break-point.

3. Compute the sum of squares and cross products matrix of $\gamma_t - \bar{\gamma}$, using the smoothed mean curve instead of the usual $\bar{\gamma}$. Find the eigenvectors to obtain the Λ matrix.

4. Select J and spline-smooth each column of Λ_J. We used $J = 1$, and customized the smoothing parameter to the value 0.5 in the R program.

5. Compute β_t^J via the formula (1).

6. Formulate a time series model for the principal component time series β_t^J. In our application $J = 1$, so we need only a univariate time series model for β_t^1. An ARIMA (1,1,0) model with trend constant was selected.

7. Forecast the principle component series, obtaining point forecasts and standard errors at each forecast horizon.

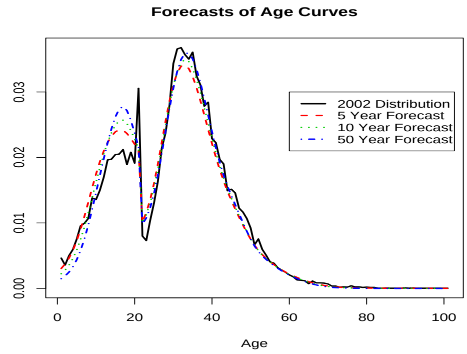
8. Modify the forecasts using Bayesian attenuation. One must decide upon an appropriate prior distribution for the forecast. We used the 30-year forecast value of the principal component as the lower limit in our Bayesian prior, and 0 as the upper limit.

9. Take the modified forecasted principal components, and apply (2). This gives the spline-smoothed, Bayesian-attenuated forecasts of the logistic-transformed data.

10. Undo the logistic transform to obtain the forecasted age distribution curves.

This is essentially the implementation used in our R program. Steps 1 through 5 are done in R, while steps 6 and 7 use the ARIMA modelling and forecasting capabilities of X-12-ARIMA. Then the output is read back into a second R program, which completes steps 8 through 10. The resulting forecasted age distribution curves for Hispanic employment immigrants are displayed in Figure 5. The forecasts have

Figure 5: Forecasted age distributions for Hispanic Employment, at 5, 10, and 50 year horizons, compared to the final year of data.



the desired properties discussed in the introduction: they are actual distributions, they are not overly simplistic (i.e., they don't smooth over relevant features of the data), and they are locally smooth. The use of PCA ensures that the curves give a fairly accurate representation of the data, with the accuracy controllable through the number of principal components used. The spline smoothing takes care of noise in a local fashion. Finally, the Bayesian attenuation ensures that forecasts at horizon $h = 50$ are not implausible. The difference between the age distribution in the last year of data and that from the first year of forecasts is somewhat large due to our use of $\bar{\gamma}$ as the baseline curve. If we instead used γ_{31} as the baseline curve, the initial forecasts would conform more closely to the particular behavior of the last year of data. Since the error in the data is highest in the last year, tying the forecasts more closely to the pattern in the last year is not necessarily desirable.

5 Summary

The PCA approach has previously been applied to age-specific fertility and mortality curves. The application here to the age distribution of Hispanic employment immigrants posed some different challenges. For U.S. fertility and mortality rates for major race groups (e.g., white or nonwhite) the data could be regarded as quite accurate, providing some rationale for using sufficient principal components to provide a very accurate approximation. The immigrant data are of lesser accuracy, particularly in the last years of the data, so there was less reason to use more than one principal component to obtain a very accurate approximation.

In addition, the nature of the appropriate forecast functions differ across the applications. For fertility relatively flat forecasts were appropriate, as fertility rates have not shown extended downward or upward trends (since the post-war "baby boom" and subsequent "baby bust.") (Log) mortality rates, on the other hand, have consistently moved downward over time, so that forecast functions with downward linear trends continuing into the distant future appear reasonable. For the immigrant age distributions the historical data produce a first principle component series showing a steady downward linear trend, but forecasting this trend to continue indefinitely eventually produces implausible age distributions. Hence, the forecasts of the principle component series were attenuated using a Bayesian approach. Also, the mean age distribution curve and principle component vector were spline smoothed because use of the unsmoothed data led to accentuation of irregularities in these vectors, producing implausible forecast age distributions in the long term.

Finally, since the data we wished to forecast here are age distributions it was necessary to produce forecasts that are themselves age distributions (nonnegative values that sum to one). We achieved this objective by applying the generalized logistic transformation to the data, forecasting in the transformed scale, and inverse transforming the results to yield forecasted age distributions.

Acknowledgements The authors thank Fred Hollmann and Ward Kingkade of the Census Bureau for providing the immigration data, and communicating the age forecasting problem.

6 Appendix: Proof of Proposition 1

The desired conditional density is $p_2(y|x) = p_2(y,x)/p_2(x)$, and

$$\begin{aligned} p_2(x) &= \int p_2(y,x)\,dy \\ &= \int p_1(y,x)\,1_{\{L<y<U\}}\,dy/c \\ &= \int_L^U p_1(x)\,p_1(y|x)\,dy/c \\ &= p_1(x)\,\mathrm{Pr}_1(L<y<U|x)/c, \end{aligned}$$

from which it follows that

$$\begin{aligned} p_2(y|x) &= \frac{p_1(y,x)1_{\{L<y<U\}}/c}{p_1(x)\,\mathrm{Pr}_1(L<y<U|x)/c} \\ &= \frac{p_1(y|x)1_{\{L<y<U\}}}{\mathrm{Pr}_1(L<y<U|x)}. \end{aligned}$$

The formula for $E_2(y|x)$ follows at once. □

The methods of Bayesian attenuation can easily be generalized. Let $f(x)$ be a probability density function, and let $p_2(y,x) = p_1(y,x)f(y)/c$, where $c = \int p_1(y)f(y)\,dy = E_1[f(y)]$. Then (6) becomes

$$E_2(y|x) = \hat{y} + \sqrt{V}\frac{\int z\phi(z)f(\hat{y}+\sqrt{V}z)\,dz}{\int \phi(z)f(\hat{y}+\sqrt{V}z)\,dz}.$$

If f is symmetric about its mean $\mu = \int yf(y)\,dy$, then $E_2(y|x) = \hat{y}$ when $\hat{y} = \mu$.

References

[1] Aitchison, J. (1987), *The Statistical Analysis of Compositional Data*, London: Chapman and Hall.

[2] Bell, William R. (1997) "Comparing and Assessing Time Series Methods for Forecasting Age-Specific Fertility and Mortality Rates," *Journal of Official Statistics*, **13**, 279-303.

[3] Bozik, J. and Bell, W. (1987), "Forecasting Age Specific Fertility," SRD Research Report No. RR-87/19, U.S. Census Bureau, available at http://www.census.gov/srd/papers/pdf/rr87-19.pdf.

[4] Box, G.E.P. and Jenkins, G. M. (1976), *Time Series Analysis: Forecasting and Control*, San Francisco: Holden Day.

[5] Hastie, T. J. and Tibshirani, R. J. (1990), *Generalized Additive Models*, Chapman and Hall.

[6] Hollmann, Frederick W. (2004), "Toward a Legal Immigrant Data Series for Population Projections," paper presented at the 2004 meeting of the Southern Demographic Association, Hilton Head Island, South Carolina.

[7] Lee, Ronald D. and Carter, Lawrence R. (1992), "Modeling and Forecasting the Time Series of U.S. Mortality," *Journal of the American Statistical Association*, **87**, 659-671.

[8] Mardia, K., Kent, J., and Bibby, J. (1979), *Multivariate Analysis*, San Diego: Academic Press.

[9] R Development Core Team (2004), *R: A language and environment for statistical computing*, Vienna, Austria. http://www.R-project.org

[10] Rogers, A. and Castro, L. (1981), *Model Migration Schedules*, RR-81-30, Laxenburg, Austria: International Institute for Applied Systems Analysis.

[11] Thompson, P. A. and Miller, R. B. (1986), "Sampling the Future: A Bayesian Approach to Forecasting From Univariate Time Series Models," *Journal of Business and Economic Statistics*, **4**, 427–436.

[12] U.S. Census Bureau (2002), *X-12-ARIMA Reference Manual*, Version 0.2.10, U.S. Census Bureau, available at http://www.census.gov/srd/www/x12a/x12down_pc.html.

This page intentionally left blank.

Volatility and Interpretation of Forecast Errors

Session Chair: Jeffrey Busse, U.S. Geological Survey (jbusse@usgs.gov)

Aberration Control in Time Trend Forecasting

Elliot Levy, U.S. Department of Commerce (elliot_levy@ita.doc.gov)

This presentation is an overview on the subject of setting upper and lower bounds for forecast error to assessing results of a prediction technique. A parsimony concept is applied against useless forecast input, e.g. linear dependent logjam. Use of quality control limits is the main emphasis that applies a loss function between actual and forecast. Also a method of 2-way forecast error adjustment is presented in an attempt to control forecasts that go awry. Costs of tracking various trend methods of forecast are presented, too.

Note: The author of this paper also has a related manuscript. Contact the author directly for more information.

The Importance of Macroeconomic Conditions on FDI and Agricultural Trade

Christine Bolling, Economic Research Service, U.S. Department of Agriculture (hbolling@ers.udsa.gov)

The dollar has been weakened since 2000, creating a changed environment for U.S. agricultural trade and foreign direct investment (FDI) compared to the macroeconomic environment of a strong dollar of much of the 1990s. Exchange rate fluctuations are only one of the major macroeconomic factors affecting trade and FDI, with incomes of other countries perhaps being the most important. The purpose of this study is to evaluate the effects of the weakened dollar and projected income growth on FDI and U.S. agricultural trade in the near future, drawing on recent studies carried out in ERS.

This page intentionally left blank.

CONCURRENT SESSIONS II

Research Projects Affecting Bureau of Labor Statistics Projections (BLS) Methods and Results

Session Chair: Michael Pilot, Chief, Division of Occupational Outlook, Bureau of Labor Statistics (pilot.michael@bls.gov)

The Educational Attainment of Occupations

Ian Wyatt, Bureau of Labor Statistics (wyatt.ian@bls.gov)

In the past, occupations projected by the Bureau of Labor Statistics have been assigned to training and education categories selected to represent the most likely path to career success in an occupation. Aggregating occupations within these education/training categories may lead to serious over-estimates of the number of jobs available to those with only a high school education or less. This study uses Current Population Survey data to examine the actual educational attainment of individuals in specific occupations and provides a natural hierarchical method of sorting occupations that reflects increasing levels of skill, education, and training. The resulting skills categorizations far better represents the likelihood of success in particular occupations for those with less than college education.

Modeling Disaggregated Producer's Investment in Equipment and Software (PIES) as a Submodel for the BLS Projections System

Kathryn Laurence, Bureau of Labor Statistics

Economic literature, based primarily on equations estimated for investment aggregates, generally supports the Neoclassical Model as the best performing approach to describing the behavior of producers' investment in equipment and software (PIES). The projections program in the BLS uses a more detailed set of PIES subcategories, leading to a breakdown of the Neoclassical Model for some of the subcategories. This discussion presents results of modeling the more detailed PIES subcategories and examines those with poor performance, suggesting reasons why the model breaks down in these cases. Since the BLS projections of detailed PIES are controlled back to a macro aggregate, the research is used to determine optimal distribution of the residual between the two models to the detailed subcategories.

Coding Occupations for Risk of Offshore Outsourcing

Norman C. Saunders, Bureau of Labor Statistics

The BLS prepares medium-term projections of industry and occupational employment demand projections, published ten years out and on a biennial basis. Clearly, some of the occupations which we forecast are at considerable risk of displacement due to offshore outsourcing. Lacking good data on the offshoring phenomenon, it was deemed necessary to better inform our occupational demand analysts about which of the 700 detailed occupations we study are the most at risk to this phenomenon. This paper presents a scoring method for assigning a risk-of-offshoring index to each of the occupations projected in BLS. Preliminary results appear to confirm that the concept works well at identifying various occupations which need closer watching over the coming decades as the offshoring phenomenon plays itself out.

This page intentionally left blank.

Modeling Disaggregated Producer's Investment in Equipment and Software (PIES) as a Submodel for the BLS Projections System

Kathryn Laurence
Bureau of Labor Statistics

Introduction

The Office of Occupational Statistics and Employment Projections (OOSEP) within the Bureau of Labor Statistics (BLS) is responsible for developing ten year projections of employment by industry and occupation. In the development of these projections, OOSEP uses several component models including models of each of the four elements of final demand: consumption, government spending, net exports, and investment. This paper will focus on the investment model to be implemented in the 2004-2014 projection cycle.

Most research to date supports the Modified Neoclassical model as the best performing model to describe past investment behavior. While other models are stronger in theory, the Modified Neoclassical performs better with the available data. Therefore, OOSEP will be using this model for the 2004-2014 projections. However, literature regarding investment models focuses on investment as an aggregate. At best, investment in structures is separated from the remaining investment. OOSEP will model investment except structures, also known as private investment in equipment and software (PIES). In order to provide more detailed categories for employment projections, PIES is disaggregated by OOSEP into twelve categories each of which is modeled and forecasted.

The model was run for each of the twelve subcategories as well as aggregate categories. This served as a test of whether research performed on PIES as an aggregate is relevant for all or some of the more detailed categories. Since subgroups do not necessarily behave similarly to their aggregate, it is not surprising to find that the model explains movements in some disaggregated PIES categories better than others. In general, the Modified Neoclassical model fails to capture the movement in high technology sectors. Since computers and peripheral equipment and software account for a substantial percentage of aggregate PIES especially in later years, these results reveal the need for further research in modeling investment in the high technology subcategories of PIES.

Literature Review

Building a concise macroeconomic model to describe investment spending has challenged researchers and policymakers for quite some time. Models which are stronger in theory often fail empirically. Personal consumption expenditures have been a much more straightforward portion of final demand to model than investment. Unlike customers who spend in very similar ways, businesses invest very differently from one another. Therefore, finding aggregate variables to explain investment and subsequently obtaining data to support these variables has proven to be an arduous task. Some of the more frequently studied models of investment are the Accelerator, Neoclassical, Modified Neoclassical, Euler equation, and Brainard-Tobin Q models or variations thereof.

The Accelerator model is the most basic in its description of investment spending. According to this model, investment is determined by recent output levels and previous capital stock. In fact, it assumes that output and demand for capital stock are proportional to one another. The structural form is as follows:

$$I_t = \sum_{i=0}^{n} a_i Q_{t-i} + cK_{t-1}$$

Where: I = investment
Q= output
K= capital stock

The model recognizes that there is a lag between when a business decides to implement a new project and when this venture is put into practice. For some projects, this time period is much greater than others. Consequently, greater demand for current output causes a gradual increase in capital spending in future time periods. Unfortunately, the Accelerator model implicitly assumes that businesses are very limited in their ability to substitute factors of production. The model performs best when the ratio of capital to output moves smoothly over time. The underlying assumptions prove to be too restrictive and the empirical results for the Accelerator model have been disappointing.

The Neoclassical model is a bit less prohibitive in its assumptions. It is based on the production function and profit maximization. The optimal stock of capital in the Neoclassical model is defined as proportional to the output divided by the cost of capital:

$$K^* = \alpha(Q^* / UCC^*)$$

Where: UCC = user cost of capital

Unlike the Accelerator model, this model includes the cost of capital as a determinant in businesses investment decisions. Using this definition along with the production function, Neoclassical researchers arrived at the following equation:

$$I_t = [\sum_{i=0}^{n} a_i (Q_{t-i} / UCC_{t-i}) +$$

$$\sum_{i=0}^{n} b_i (Q_{t-i} / UCC_{t-i-1}) + cK_{t-1}$$

When academics tested this model with the available data, they were not pleased with the results. After further empirical studies, they realized that output variables exert a much stronger influence on future investment decisions than do price variables. Combining output and pricing effects into one variable prohibited the model from exhibiting this characteristic.

Once researchers understood the importance of separating the output and price variables, they made a few minor adjustments to the Neoclassical model in order to arrive at the Modified Neoclassical model. The stock of capital, according to this model, is a function of the ratio of the optimal to the existing stock of capital.

$$K_t / K_{t-1} = (K_t^* / K_{t-1})^{\lambda}$$

Defining investment as the change in capital stock resulted in the following:

$$I_t / K_{t-1} = \Delta K_t / K_{t-1} \approx \lambda(\log(K_t^*) - \log(K_{t-1}))$$

Inserting the optimal stock of capital from the Neoclassical model into this definition yields a model which splits output and price effects into two separate variables:

$$I_t / K_{t-1} = \lambda[\log\{\alpha(Q^* / UCC^*)\}] - \log(K_{t-1})$$

Which can be further simplified to:

$$I_t / K_{t-1} = \lambda[\log(\alpha) + \log(Q^*) - \log(UCC^*)] - \log(K_{t-1}) \quad (1)$$

Results from the Modified Neoclassical model were much better than the previous models, however, there was still room for significant improvement.

Next came the Brainard-Tobin q model and the Euler models which better account for expectations and technology and often include variables to capture adjustment costs. The Q model holds that investment is positively related to q, the ratio of the financial value of the firm to the replacement costs of its existing capital stocks. While the model appears strong from a theoretical standpoint, it is not popular in practice. Many explanations for the poor performance have been offered including problems arising from the assumption of homogenous capital, too large of adjustment costs, mismeasured capital stock, and the likely inadequacy of available data to estimate q (Chirinko, 1888-1893).

In 1976 Robert Lucas published a critique in which he argued that agents were likely to change their investment decisions in response to alterations in the behavior of policymakers. Models of investment up to this time were not capturing this behavior. This argument shifted macroeconomics towards forward looking models with expectations and variables of taste and technology. The Euler equations were a direct response to this critique and have dominated academic research since that time. Because Euler equations are more involved than earlier models, we will not discuss them in detail here except to note that empirical results show they are more successful in explaining investment than the q models. Since Euler models are quite complicated and adequate data has not been found, they have not yet performed as well as some of the earlier models. The hope is that with continued research, modifications to the Euler equations and further data availability will at some point yield results superior to earlier models of investment.

Model Specifications

Given the results from previous literature, OOSEP will be using the Modified Neoclassical model of investment as the basis for projecting the investment portion of final demand for the 2014 employment projections. Combining the constants λ and α from equation 1 results in the reduced form of the model:

$$I_t / K_{t-1} = a_1 + a_2 * \log(Q^*) + a_3 * \log(UCC^*) + a_4 * \log(K_{t-1}) \quad (2)$$

As mentioned previously, OOSEP models PIES at a much more detailed level. For our purposes, PIES is disaggregated into the following subcategories derived from the Bureau of Economic Analysis' (BEA) National Income and Product Accounts (NIPA) and Input-Output (I-O) tables:

1: Computers and Peripheral Equipment
2: Software
3: Communication Equipment
4: Other Information Processing Equipment
5: Autos
6: Trucks, Buses, and Truck Trailers
7: Aircraft
8: Ships, Boats, and Railroad Equipment
9: Industrial Equipment
10: Other Equipment
11: Scrap
12: Residential Equipment

Equation 2 is modified to incorporate these categories by including the variable i=1 to 12. This brings us to the final form of the regression equation:

$$I_t^i / K_{t-1}^i = a_1 + a_2 * \log(Q^*) + a_3 * \log(UCC^*)] + a_4 * \log(K_{t-1}^i) \quad (3)$$

Aggregate models will also be run for categories one through three, four to twelve, and all twelve categories.

The investment and capital stock variables are determined endogenously while output and the user cost of capital are exogenous to the model. Output and investment, according to theory, are positively correlated. Therefore, in each of the models Q should have a positive coefficient. An increase in the user cost of capital discourages investment and should be associated with a negative coefficient. The coefficient of the lagged capital stock in this equation represents the rate at which investors move toward their observed need of capital and their existing level as well as the rate at which aged capital is replaced. According to economic theory, previous capital stock is expected to hold a negative relationship with investment.

The assumptions imbedded in the Modified Neoclassical model of investment are numerous. For example, the Lucas critique reminds us that the model cannot capture changes to investment as a response to policy makers adjustments to regulations. Expectations of firms are not included anywhere in the equation. Moreover, the model does not contain any variables to capture technological change. Changes in technology certainly have important effects on investment decisions especially in the detailed categories of computers and software. The model is very restrictive in its assumptions; however, empirically it has performed better than models which are stronger from a theoretical standpoint.

Procedural Aspects

Another component model used in the OOSEP employment projections is the MA model, a proprietary macro-economic model supplied by Macro-Economic Advisors. This model compiles historical data from numerous sources, uses this data to arrive at some of their own variables, and then forecasts all of the data over the given projection period. The macro model is a work in progress. Updates are made continually throughout the projections process. The comparisons of results published in this paper are not necessarily to the final version of the macro-model for OOSEP's projections to 2014. OOSEP does not need to incorporate a separate model for investment in structures because forecasts of both residential and nonresidential are provided by the MA model.

While the Modified Neoclassical model will provide the detailed subcategory preliminary estimates of PIES, the total projection for PIES is constrained to the forecast derived in the MA model. In other words, the subcategories must chain weight to the MA model forecast of PIES. BEA relies upon chain-type annual-weighted indexes, also referred to as Fisher indexes to measure real output and prices. Since the majority of our data is provided by BEA, we also use the chain weighting procedure in our projections process. Any aggregate data referred to in this paper was arrived at using the chain weighting technique[1].

In order to chain weight the sub-aggregate PIES columns, the deflators for each category were projected using the following equation:

$$\log(\mathrm{P_i}) = \mathrm{a}_0 + \mathrm{a}_1 * \log(\mathrm{PQB}) + \mathrm{a}_2 * \log(\mathrm{P_i}(\text{-}1)) + \mathrm{a}_3 * \log(\mathrm{P_i}(\text{-}2))$$

Where: P = the deflator for the given category, i
PQB = the GDP deflator
P(-1) = the deflator in the last period

The residual between the MA model projection and the chain weighted forecast for PIES from the Neoclassical Model of investment was then proportionally distributed back to each of the detailed pieces of

[1] For more information on chain weighting, please see the Landefeld and Parker paper listed in the references.

investment. It should be noted that poor projection results could be a result of not only the Modified Neoclassical model but also the price forecasts.

Data Collection

In the majority of available literature, models are not used to forecast investment. Instead, they focus on explaining past investment behavior. Using the model to project investment complicates the research because all of the included variables must be available for the forecasted time period. Investment values were the only variable for which no future data was needed. Therefore, we obtained historical values for each of the twelve subcategories from section five of BEA's NIPA tables.

For the remaining variables, we relied heavily upon the MA model for data. The estimate for output was provided by the macro-model's Gross Domestic Product, Nonfarm business less housing (QB). Historical data for QB was compiled by MacroEconomic Advisors from the BEA's NIPA accounts. Also from the macro-model, the user cost of capital was measured by the Marginal Rental Price of producer's durable equipment, computers and software (RPDC) and the Marginal Rental Price of producer's durable equipment, except computers and software (RPDO). The only exception was autos for which the macro-model provides Rental Price of Consumer Durables, motor vehicles (RCDMV).

Capital Stocks historical and forecasted data were also provided by the macro-model for total stock of producer's durables (KPD). The breakout was also available for KPD into the stock of producer's durables computers (KPDC), and stock of producer's durables except computers (KPDO). Once again the MA model used historical data from BEA NIPA accounts in order to arrive at their historical series for each of the capital stock variables. Additional Capital stock data is published in BEA's Fixed Asset tables for each detailed subcategory of investment. To calculate historical real values for capital stocks, we applied the quantity index from table 2.2, "Chain-Type Quantity Indexes for Net Stock of Private Fixed Assets, Equipment and Software, and Structures by Type", to the value in 2000 from table 2.1, "Current-Cost Net Stock of Private Fixed Assets, Equipment and Software, and Structures by Type".

In order to run the model, these individual capital stocks needed to be forecasted. We assumed that the depreciation rate over the projection period, 2003-2014, would equal the average depreciation rate for 1990-2002. The investment model was then run iteratively along with the following capital stocks model:

$$K_t^i = \delta^i * K_{t-1}^i + I_t^i$$

Where: δ = the depreciation rate

The only exception to this was scrap (i=11) because there is no individual capital stock published for scrap. Therefore, for this sub-category of investment capital stock was estimated by KPDO from the macro-model and the following investment model was used:

$$I_t^i / KPDO_{t-1} = a_1 + a_2 * \log(Q^*) + a_3 * \log(UCC^*) + a_4 * \log(KPDO_{t-1})$$

Although BEA publishes earlier data, the MA model at the time of our estimation included historical data only for 1967 through 2002 and projected data through 2014. Therefore, our model is based on data from 1967 through 2002 and forecasts investment from 2003 through 2014.

Regression Results

Two lags of output and user cost of capital were incorporated into the Modified Neoclassical model of investment in order to account for time lags in business decisions resulting in the following equation:

$$\begin{aligned} \mathrm{I_i/KS_i(\text{-}1)} &= \mathrm{a_0 + a_1 * \log(QB) + a_2 * \log(QB(\text{-}1))} \\ &+ \mathrm{a_3 * \log(QB(\text{-}2)) + a_4 * \log(RPD)} \\ &+ \mathrm{a_5 * \log(RPD(\text{-}1)) + a_6 * \log(RPD(\text{-}2))} \\ &+ \mathrm{a_7 * \log(KS_i(\text{-}1))} \end{aligned}$$

For the first two subcategories, computers and peripheral equipment and software, RPDC was used in place of RPD. RCDMV was utilized for subcategory five, autos. For the remaining groups, RPDO was employed. In the equation for scrap, KPDO was used in place of KS_i. RPDC was also used in the aggregate equation for subcategories one to three and capital stock was estimated by KPDC. For the aggregate of four to twelve, RPDO and KPDO were included. In the overall PIES regression, both RPDC and RPDO were included in the equation as well as KPD. A dummy variable was included for ships, boats and railroad equipment (i=8) in order to account for a severe level drop in the early 1980's[2].

[2] The dummy variable is equal to zero for years prior to 1981, 0.3 for 1981, 0.7 for 1982, and 1 for years 1983 and forward.

The results of the detailed regressions for each of the subcategories as well as the three aggregate equations are presented in the attached Table 1. Coefficients for each variable and equation are provided in this table as well as the corresponding probabilities of their t-statistics. In the equations for categories 1, 4, 9, and 10, at least one of the lags on each of the GDP, rental cost of capital, and capital stock variables are significant at the five percent level. The capital stock for equations for 3, 5, 6, 11 and 12 are not significant at the five percent level. Recall that the capital stocks for the detailed categories are forecasted under the assumption that the depreciation rate stays at its average annual rate from 1990-2002. Perhaps this assumption is invalid for some of these categories. Categories 2, 6, 8, and 12 do not hold significant coefficients on the user cost of capital or its lags. This is not surprising given that most previous literature recognizes output variables are much more important to investment decisions than pricing variables. Most worrisome is the fact that groups 2, 3, 7, 8, and 11 do not have significant coefficients on any of the output variables. The aggregate equations for categories one to three and overall PIES do not contain significant coefficients for the lagged capital stock or on the user cost of capital.

Output is expected to be associated with positive coefficients while capital stock and user cost of capital variables should be paired with negative coefficients. While each of the lags may not be of the correct sign, they should be correct when the coefficients on the lags are summed. For the most part, the signs on the coefficients in the equations are as expected. The coefficients for output, however, sum to a negative value for equations 4, 5, 11, and 12. While all other PIES values are positive, scrap is a negative value and behaves much differently. Therefore, the negative coefficient may be reasonable. For the remaining equations, this is especially worrisome. Capital stocks variables hold positive signs on their coefficients for equations 4 and 11, but their impacts are not as significant as those from the output variables[3]. The rental price of capital carried a positive sign in the equations for 11, 12, 1-3, and PIES[4].

The magnitudes of the coefficients are as expected. Analysis of the coefficients along with the relative size of the data shows that output effects contribute more to investment in all of the equations than either lagged capital stock or the rental price of capital. This is especially true for the aggregate equations in which it should be noted that all of the current output variables were significant at the one percent level. The next table includes the R-squared value, Durban Watson Statistic, and the probability of the F-statistic for each of the equations:

Model	***R^2***	***DW***	***Prob(F-stat)***
1	0.6677	0.6233	0.0001
2	0.3055	0.4123	0.1520
3	0.5322	0.6429	0.0023
4	0.8747	1.6668	0.0000
5	0.8997	1.7638	0.0000
6	0.7390	0.5153	0.0000
7	0.6311	1.3443	0.0001
8	0.8985	1.0843	0.0000
9	0.8169	0.8299	0.0000
10	0.8296	0.9780	0.0000
11	0.6258	0.8173	0.0002
12	0.8983	0.7285	0.0000
1-3	0.9680	0.3838	0.0000
4-12	0.8863	0.5693	0.0000
PIES	0.9412	0.7143	0.0000

The R-squared values for the most part are high or very high with exception to software and communication equipment. Notice that the R-squared values for the aggregates are for generally higher than the sub aggregate equations. This is especially true for groups one through three. The Durbin Watson statistics, however, indicate that there is a problem with serial correlation in most of the equations. The presence of autocorrelation could be magnifying the R–squared values. High p values presented in table 1 indicate that there are most likely problems with multicollinearity as well in the equations. These suspicions are confirmed in the next table which presents the pairs of variables in which variables are highly correlated.

Eq.	**Correlated Pairs (>.8)[5]**
1	QB & RPDC, QB & KS, RPDC & KS
2	I & QB, I & RPDC, I & KS, QB & RPDC,QB & KS, RPDC & KS
3	I & QB, I &KS, QB & KS
4	I & QB, I & KS, QB & KS
5	I/KS & QB, I/KS & KS, I &QB, I & KS, QB & KS
6	I & KS, QB & KS
7	QB & KS
8	None
9	I & QB, I & KS, QB & KS

[3] This is taking into account not just the size of the coefficient but also the relative magnitude of the variables.

[4] Only RPDC in the PIES equation has the incorrect sign, not RPDO. However, the coefficient on RPDC is larger than that on RPDO so their sum is also positive.

[5] All of the correlations are calculated for the variables as they appear in the equation. For example, QB is actually log(QB) and KS is KS_i. I chose not to include lags to minimize table space, but in general if the variable is correlated, all of its lags are correlated as well.

10	I & QB, I & KS, QB & KS
11	QB & KPDO
12	I/KS & QB, I/KS & KS, I & QB, I & KS, QB & KS
1-3	I/KPDC & QB, I/KPDC & RPDC, I/KPDC & KPDC, I & QB, I & RPDC, I & KPDC, QB & RPDC, QB & KPDC, RPDC & KPDC
1-4	I & QB, I & KPDO, QB & KPDO
PIES	I/KPD & QB, I/KPD & RPDC, I/KPD & KPD, I & QB, I & RPDC, I & KPD, QB & RPDC, QB & KPD

The theory behind the modified neoclassical model, however, supports that all of these variables are important determinants of investment. Therefore, although multicollinearity is present, we leave all of the variables in the equation.

Forecasting Investment

The annual totals from the chain weighted sub aggregate models are as low as about ninety percent of the macro-model results but in later years approach the MA model projection. In fact, by 2014, the results chain weight to greater than ninety eight percent of the macro-model result. Since OOSEP publishes point projections and not time series forecasts, this left only a small residual to be redistributed back to the detailed models at least for the projection year.

The average annual growth rates presented in the following table appear reasonable for most equations. Included are annual average growth rates for the entire historical period, the last thirteen years, the last five years, and the entire projected period.

Group	Historical	1990-2003	1998-2003	Pro-jected
1	1.3762	1.2771	1.2467	1.1250
2	1.4115	1.1311	1.0940	1.1367
3	1.0879	1.0921	1.0909	1.0998
4	1.0686	1.0462	1.0611	0.9271
5	1.0434	1.0136	0.9837	1.0072
6	1.0529	1.0550	0.9790	1.0166
7	1.0638	1.0471	1.0253	1.0551
8	1.0213	1.0294	0.9898	1.0274
9	1.0282	1.0108	0.9875	1.0195
10	1.0316	1.0199	1.0156	1.0125
11	1.0632	1.0165	0.9183	0.9931
12	1.0496	1.0160	1.0441	0.9941
1 to 3	1.1799	1.1527	1.1237	1.3610
4 to 12	1.0364	1.0246	1.0079	1.0118
PIES	1.0669	1.0686	1.0523	1.0855

The projected growth rates for the majority of the equations are very close to the actual rates in the past five and thirteen year periods. The forecasted growth rate for equation twelve (residential equipment) is a bit low. Computers and peripheral equipment, software, other information processing equipment, and aggregate results for groups one to three appear highly unreasonable when compared to historical data.

Low forecasted growth rates for computers and peripheral equipment and other information processing equipment are exacerbated by low forecasts in the early years of the projection period. The forecasted 2003 levels for categories one and four respectively are 14.4% and 17.0% below the actual data for these years. Scrap is 60.9% high and residential equipment 24.3% low. Software was also projected 19% higher in 2002 than the level published in the NIPA accounts. The projected aggregate for groups one to three is 51.6% higher than the published level. The aggregate for groups four to twelve fares much better at only 9.4% over actual levels and total PIES was only 5.6% above where it should be.

Implications

The Modified Neoclassical model of investment performed much better for some sub-aggregate groups of investment than others. OOSEP must keep in mind the strengths and weaknesses of the results of these equations. As the projection results are analyzed and hand changes made, categories with the weakest results should receive the majority of changes needed to reach the MA model and to correct for any anomalies. Analysts should resist making changes to those categories in which regression results were stronger. At both the aggregate level and as individual categories, results from groups four through twelve were much stronger than those from groups one through three; computers and peripheral equipment, software, and communication equipment. Findings were also disappointing in the scrap and residential equipment categories, but these categories only represent a small component of PIES. Therefore, we are not nearly as concerned at this time with their performances as the high technology categories. Sub-categories one through three grew much more quickly in the past than the remaining groups. Because investment was forecasted in real dollar values, rapidly changing technology affected these categories in two important ways: high depreciation rates and the need for businesses to replace this type of equipment often. These two effects magnified one another. Therefore, the sectors showed significantly faster growth than other portions of PIES. Because these categories do not behave in the same manner as the remaining PIES subcategories, they most likely require a separate model to explain their behavior. Research, however, continues to focus on PIES as a whole. Because computers and peripheral equipment and software contribute a significant and growing percentage of the total PIES value, the need for a model which captures their growth is ever more important.

Table 1. Regression Results

	a_0 c	a_1 QB	a_2 QB(-1)	a_3 QB(-2)	a_4 RPD	a_5 RPD(-1)	a_6 RPD(-2)	a_7 KS_i(-1)[6]
1	1441.888	1013.788	1048.801	-1755.043	-1755.283	310.7329	828.5233	-296.3971
	(.8439)	(.2589)	(.3609)	(.0439)	(.0031)	(.6656)	(.1334)	(.0041)
2	-4547.795	466.7656	211.0931	191.8316	-5.099343	38.13217	-203.9337	-336.8908
	(.1752)	(.3703)	(.7666)	(.7155)	(.9881)	(.9326)	(.5494)	(.0417)
3	-526.0287	325.6465	-18.29115	-103.6397	98.14476	48.32438	-328.5981	-87.48233
	(.6004)	(.1033)	(.9465)	(.5924)	(.3466)	(.7466)	(.0073)	(.1500)
4	3973.476	-226.0669	186.7752	-276.4682	-205.1414	-207.2658	-47.37898	90.18611
	(.0000)	(.1445)	(.3818)	(.0502)	(.0211)	(.0785)	(.5859)	(.0010)
5	3499.519	1198.920	-1111.776	-220.0271	-425.8638	298.7177	-326.1996	-98.43961
	(.0000)	(.0006)	(.0139)	(.3772)	(.0275)	(.3408)	(.1407)	(.0783)
6	1283.858	791.7926	250.3912	-989.4782	-206.2681	91.75956	-177.5352	-126.8506
	(.0036)	(.0042)	(.5064)	(.0004)	(.1771)	(.6550)	(.2887)	(.0771)
7	-700.7270	73.25000	36.77337	129.2709	287.6917	-205.0902	-89.76605	-236.3583
	(.0578)	(.6630)	(.8814)	(.4367)	(.0044)	(.1327)	(.3788)	(.0000)
8[7]	122.4719	48.47091	25.33829	-34.50405	20.46308	-14.17236	-49.46535	-44.55565
	(.7634)	(.5041)	(.7972)	(.5832)	(.5987)	(.7992)	(.2956)	(.0328)
9	703.0831	58.54974	148.2089	-135.9032	15.61220	-8.450189	-80.77108	-140.4948
	(.0000)	(.2810)	(.0670)	(.0146)	(.5972)	(.8420)	(.0175)	(.0000)
10	1276.938	89.58004	179.1948	-215.0956	-79.55265	-10.64705	-123.5902	-143.8612
	(.0000)	(.3018)	(.1668)	(.0158)	(.1081)	(.8775)	(.0250)	(.0000)
11	-34.84447	-3.636132	5.128868	-2.871751	-3.492911	2.173969	7.866520	3.215858
	(.0000)	(.4309)	(.4406)	(.5615)	(.1672)	(.5446)	(.0071)	(.3078)
12	507.6077	94.76942	139.6977	-275.0191	8.396725	-15.78835	57.60706	-39.44185
	(.1093)	(.3898)	(.3815)	(.0131)	(.8916)	(.8551)	(.3828)	(.0880)
1-3	-92542.93	6054.547	2269.789	1080.830	-172.9859	1542.255	1470.281	-323.2599
	(.0000)	(.0100)	(.4679)	(.6351)	(.9073)	(.4416)	(.3693)	(.4652)
4-12[8]	1022.083	174.4752	68.56096	-126.2348	-38.64344	-47.69774	-111.1368	-183.9503
	(.0012)	(.0078)	(.4103)	(.0803)	(.3165)	(.3256)	(.0069)	(.0014)

	a_0 c	a_1 QB	a_2 QB(-1)	a_3 QB(-2)	a_4 RPDO	a_5 RPDO(-1)	a_6 RPDO(-2)	a_7 RPDC
PIES	-3131.186	435.8491	138.5576	-173.7455	21.28252	31.90587	-73.07879	-116.633
	0.0004	0.0003	0.323	0.11	0.7328	0.7514	0.3033	0.102

	a_8 RPDC(-1)	a_9 RPDC(-2)	a_{10} KPD(-1)
PIES cont.	35.98323	138.3685	-51.45009
	0.7407	0.0588	0.3865

[6] KPDO was used in place of KS_i for equations 11 and 1-3 and KPDC in equation 4-12
[7] The dummy variable for equation 8 had a coefficient of -61.65337 (0.0001)
[8] The log(RCDMV) variable for equation 4-12 had a coefficient of 41.17257 (0.2040) and Dummy had 13.98515 (0.2890)

Bibliography

Abel, Andrew B. and Janice C. Eberly. "A Unified Model of Investment Under Uncertainty". *The American Economic Review*. 84:5 (December, 1994). 1369-1384.

Anderson, G. J.. "A New Approach to the Empirical Investigation of Investment Expenditures". *The Economic Journal*. 91:361 (March, 1981). 88-103.

Bean, C. R.. "A New Approach to the Empirical Investigation of Investment Expenditures: A Comment". *The Economic Journal*. 91:361 (March, 1981). 104-105.

Chirinko, Robert S.. "Business Fixed Investment Spending: Modeling Strategies, Empirical Results, and Policy Implications". *Journal of Economic Literature*. 31:4 (December, 1993). 1875-1911.

Galeotti, Marzio. "Specification of the Technology for Neoclassical Investment Theory: Testing the Adjustment Costs Approach". *The Review of Economics and Statistics*. 72:3 (August, 1990). 471-480.

Kopcke, Richard W. and Richard S. Brauman. "The Performance of Traditional Macroeconomic Models of Businesses' Investment Spending". *New England Economic Review*. 2 (2001). http://www.bos.frb.org/economic/neer/neer2001/neer201a.pdf

Landefeld, J. Steven and Robert P. Parker. "BEA's Chain Indexes, Time Series, and Measures of Long-Term Economic Growth". *Survey of Current Business*. (May, 1997). 58-68.

Li, Hong. "Estimation and Testing of Euler Equation Models with Time-Varying Reduced-Form Coefficients". *Princeton University; Job Market Paper*. (November, 2004). http://www.princeton.edu/~hli/jobpaper.pdf

Oliner, Stephen, Glenn Rudebusch, and Daniel Sichel. "New and Old Models of Business Investment: A Comparison of Forecasting Performance". *Journal of Money, Credit and Banking*. 27:3 (August, 1995). 806-826.

Pacheco-de-Almeida, Goncalo and Peter Zemsky. "The Effect of Time-to-Build on Strategic Investment Under Uncertainty". *RAND Journal of Economics*. 34:1 (Spring, 2003). 166-182.

Roberts, John M.. "Modeling Aggregate Investment: A Fundamentalist Approach". *The Federal Reserve Board: Finance and Economics Discussion Series*. August, 2003. http://www.federalreserve.gov/pubs/feds/2003/200348/200348pap.pdf.

This page intentionally left blank.

Coding Occupations for Risk of Offshore Outsourcing

Norman C. Saunders, Research Coordinator
Office of Occupational Statistics and Employment Projections

The Office of Occupational Statistics and Employment Projections (OOSEP) has a dual mission--first, to produce historical statistics of detailed occupational employment and wages, and second, to prepare annual, national-level projections of aggregate economic activity; labor force by age, sex, race, and ethnicity; output, productivity, and employment for about 200 detailed industries; and, for each of these industries, the occupational make-up of their employment.

The primary audience for the BLS projections are students, career counselors, and State Labor Market Information Offices. Secondarily, we serve academicians, other government agencies, the media, and policymakers. BLS's biennial projections appear in a series of articles in the ***Monthly Labor Review***, in the **Occupational Outlook Handbook**, the **Career Guide to Industries**, and **Occupational Projections and Training Data**.

The purpose of this paper is to discuss an index of the "risk of offshoring" developed over the past year in OOSEP. Our 2000-2010 projections, published in November of 2001, were developed in a time of almost irrational exuberance. The U.S. economy was on its longest recovery period since the end of World War II, the Federal government was in surplus for the first time in decades, and the few nagging warning signs of trouble ahead hadn't adequately penetrated our consciousness. Known for pessimistic projections, OOSEP was pleased to publish in 2001 perhaps our most optimistic projections ever. We were, unfortunately, overtaken by the times.

By the time we were deep into the preparation phase of our 2002-2012 projections, published in February 2004[1], we were dealing with a collection of harsh realities:

- The tech bubble had burst, leading to calamitous impacts on equity valuations.
- Massive accounting frauds at Enron and other corporations undermined confidence in the upper management of many U.S. companies.
- The entire nation was dealing with the shock and agony of the 9/11 terrorist attacks.
- We were dealing with a war against terrorism (mainly being fought in Afghanistan) and a war against "weapons of mass destruction," being fought in Iraq, both of which were exacting a price in terms of U.S. lives and a return to Federal deficits on a grand scale.
- Finally, we began to have a sense that a new economic phenomenon was emerging, one which would impact white-collar service-producing occupations--offshore outsourcing.

All of these factors had negative affects on our 2012 projections but we were haunted by the nagging sense that we really didn't understand the offshoring phenomenon and we had very little idea what the potential might be for serious job displacement in the U.S.

After a few months of internal discussion, the decision was taken to form a team to take on this problem. The team was comprised of two occupational analysts, an industry specialist, and a macroeconomist. The focus of the team's analysis was twofold. First, the team was asked to collect all the information possible on offshoring--what data was there, what research was underway, and what anecdotal material was available.

Determinations on this first goal came rapidly. Hard data on offshoring was in exceedingly short supply. Several data sources were being mined, notably the Mass Layoff Statistics program in the Bureau of Labor Statistics[2], the survey of U.S. Direct Investment Abroad, carried out annually by the Bureau of Economic Analysis, and Commerce Department attempts to piggy-back offshoring questions onto the periodic Institute for Supply Management Survey. [3]

A few serious research programs are underway, but for the most part these are aimed at putting the phenomenon of offshore outsourcing in an economic context and suggesting ways in which reliable statistical data could be collected to better gauge the actual impact of offshoring.

Much of what we know of the offshoring of white-collar service-producing jobs is anecdotal in nature and stems from an infamous series of proclamations from Forrester Associates and Gartner Associates, both firms who provide consulting services to other U.S. firms wishing to avail themselves of offshoring as a cost-cutting mechanism.

The second focus of the OOSEP team was to determine the characteristics of occupations which made them more or less at risk of being moved offshore. The ultimate purpose was to inform our staff of occupational analysts involved in researching the outlook for the detailed occupations included every other year in the OOH. Hopefully the result of this information would be more reasonable staffing pattern projections for the coming decade.

The team developed rather lengthy lists of characteristics on both sides of the question, beginning with characteristics of work favorable for performance in the U.S.:

- Work in which there is uncertainty about what the customer wants or what the specifications should be
- Projects that require highly iterative development processes
- Work that requires a high degree of personal interaction with end-users/clients
- Work that crosses many disciplines
- Applications with complex procedures, including substantial manual intervention and data fixes
- Applications that involve a high degree of integration with other systems developed and maintained on-shore
- Work involving nuances or deep cultural understanding
- Work in which much of the knowledge exists only in the minds of the on-shore staff
- Analytical tasks, leading-edge research and non-rule-based decision-making
- High levels of creativity, innovation, insight, “thinking outside the box”
- High management requirements
- Process design and business analysis
- Technology and systems integration (applications, hardware and networks)
- Fusion of industry knowledge, high-level skills, and business process expertise

Another detailed list was formulated for work favorable for performance outside of the U.S.:

- High wage differential with similar occupation/level in destination country
- High labor intensity
- Clearly defined requirements, little nuance
- Repetitive tasks
- Rule-based decision-making and problem solving
- Documented or easily transferred content and process knowledge
- Discreet, separable; low degree of interaction across different services, applications
- Low degree of personal interaction with end-users, clients
- Stable applications with minimum of “firefighting”
- Long projected useful life to amortize offshore set-up costs
- Low-to-medium business criticality
- Less time-sensitive, longer transition periods
- Projects involving simple and standard hardware and software
- Digital, Internet-enabled
- Low setup barriers
- Low-to-medium technical complexity
- Not multi-disciplinary
- Projects in business areas in which offshoring is a broadly accepted concept
- Tightly defined work processes
- Stable process

These can be readily summarized as: work most likely to be offshored is that which includes repetitive tasks, has clear requirements with little nuance, or little personal interaction with end-users or clients, is less time-sensitive with longer transition periods, digital, Internet-enabled and not multi-disciplinary.

Work least likely to be offshored is that which crosses many disciplines, requires a lot of interaction, includes a lot of uncertainty about the specifications, involves nuances or deep cultural understanding, and depends on creativity and innovation.

These are only our first cut at defining the characteristics which mitigate for and against offshoring. The team anticipates that the lists will evolve over time as analyses are carried out based on the lists.

Once the team had developed the characteristics which best represented offshoring tendencies, the next step was to design a method for determining whether those characteristics applied to specific occupations.

The team decided that a useable approach was that utilized by Forrester Associates in their initial estimates of particular occupations at risk of offshoring, namely a series of questions were derived to reflect the foregoing characteristics.

The questions were to be answered with "Yes" or "No", where a Yes answer indicated a greater risk of offshoring. In all, eight questions were developed, with "Yes" answers being scored as a 1 and "No" answers being scored as a 0[4]. The individual scores were then summed, resulting in an occupational score ranging between 0 and 8. The individual occupational scores were then normalized to a 0-100 scale.

Following are the eight questions, with commentary.

Question 1. Can the occupation be successfully carried out without being onsite or requiring a security clearance?

This was our so-called "deal breaker," i.e. if this question was answered with a "No", then a score of 0 was assigned to the occupation and none of the remaining 7 questions were evaluated.

Some examples: Most Physicians need to be onsite to examine patients and Janitors need to be onsite to clean floors but do tax preparers, graphic designers, or architects need to be onsite?

Question 2. Are computers, telephones, or other telecommunications equipment used by the occupation extensively?

If all of the occupation's communications with coworkers, clients, or customers could be done by telephone, or e-mail, or fax, then communication needs would not prevent the jobs from being offshored.

Question 3. Can the work of the occupation be routinized or handled by following a script?

If someone is not completely fluent in English or familiar with cultural issues, this potential handicap could be compensated for by structuring their communication with clients or customers so that they are trained to ask a routine series of scripted questions to which responses are limited, questions that that would be answered by only “yes” or “no” or very simple responses, and then, depending on the response, scripted follow-up questions could be asked.

Question 4. Can the tasks of the occupation be carried out with little knowledge of social issues, industrial organization, or other local knowledge?

Do culturally diverse customers or clients interact with, or respond to, members of the occupation in subtly different ways? Is an understanding these differences important for the occupation successfully completing tasks? Some examples of “No” answers could be Counselor, Social worker, Psychologist, Funeral director, Advertising sales agent, or Teacher.

Question 5. Are the products or services produced by the occupation and the inputs required to do the tasks telecommunicable, Internet-capable, or easily and cheaply transportable?

How easily can the product or service produced by the occupation be conveyed to the customer or client? A photographer, for example, can e-mail a digital picture from anywhere.

How easily are the inputs to the work of the occupation transported to a distant location? A school photographer cannot readily transport an elementary school class to Asia to take their pictures.

Question 6. Is the product or service produced by the occupation modular in nature?

By “modular”, we are asking "Can a good or service produced by the occupation be separated from the work of other occupations that contribute to a larger product?"

A garment can be made of fabric woven in the U.S., designed and cut in the U.S., and sold to customers in the U.S., but assembled by sewing machine operators abroad. Motor vehicles are modular, assembled from parts and sub-assemblies made in many locations. Some software is comprised of code modules that perform specific functions linked together with a custom front end.

Question 7. Can the tasks of the occupation be carried out without any special license or other regulatory requirements?

Certain types of work are regulated by State and local governments in the interest of public safety and consumer protection.

Some occupations that generally must be licensed include architects, barbers, cosmetologists, lawyers, and physicians.

Licensure varies in other occupations, such as accountants and auditors, electricians, and engineers.

Question 8. Does the median wage of the occupation fall within the middle 2 quartiles of wages for all occupations in the U.S.?

Workers in occupations in the lowest quartile of earnings are often low paid, so offshoring the jobs results in little cost savings.

Workers in occupations in the highest quartile of earnings have the potential for large cost savings, but often their skills are so valuable that offshoring the jobs would not be considered.

Workers in the middle two quartiles are the best candidates for offshoring: relatively good potential cost savings and comparable skills may be readily available offshore.

As a test of the process, the team applied the questions to the 329 detailed occupations which are contained in the following nine major occupational groups of the Standard Occupational Classification/2000:

11	Management
13	Business and Financial Operations
15	Computer and Mathematical
17	Architecture and Engineering
19	Life, Physical, and Social Science
27	Arts, Design, Entertainment, Sports, & Media
29	Healthcare Practitioner and Technical
41	Sales and Related
43	Office and Administrative Support

The SOC identifies 821 detailed occupations in 23 major groups but these nine major groups were felt to best represent the target group of white collar service producing jobs.

The team applied the eight questions, compiled the results, and computed index estimates. The results scoring in the 75-100 % range encompassed:

- 86 detailed occupations, or
- 11.5 million employed persons, accounting for
- 8% of total employment and
- 26% of the occupations considered in the analysis

Widening the base to those scoring 50% or higher included:

- 112 detailed occupations, or
- 17.7 million employed persons, accounting for
- 12% of total employment and
- 34% of the occupations considered in the analysis

One has to be cautious about how one interprets these figures. They do not mean that 12 million people will lose their jobs as a result of offshoring or that 8% of total employment will disappear to other countries. Rather, they should be used as an indicator of the boundaries of white-collar service-producing occupations which are more likely to be affected in the coming years by offshoring.

Other preliminary results are shown in Table 1 and appear to validate the scoring technique as a method of identifying groups of occupations that appear to be losing jobs. Consider that employment fell between 2000 and 2010 by 0.5% per year. Then consider that the occupations selected in the 50% and over category declined at a rate of 1.5% annually and that those in the more stringent grouping of 75% and over declined during the same period by 1.8 percent each year.

Table 1. Preliminary Summary Results of the Offshore Coding Exercise

	2000	2002	2010	2012	2000-02	2000-10	2002-12
Total employment	145.6	144.0	167.8	165.3	-0.5	1.4	1.4
75%+	11.5	11.1	12.4	11.7	-1.8	0.8	0.8
Share of total	7.9	7.7	7.4	7.1			
50%+	17.7	17.2	20.5	19.0	-1.5	1.5	1.0
Share of total	12.2	11.9	12.2	11.5			

Clearly the technique has captured groupings of occupations with poorer historical job performance than the overall level of employment in the economy.

Further, consider that the total employment and the 75% employment are both predicted to grow at the same rate in 2002-2012 as they were in the 2000-2010 projections. At first glance this might lead you to think that knowing about offshoring in the second round of projections led to no differences. Since we're starting at a lower takeoff point in 2002 than in 2000, clearly we moderated our overall employment growth in the later projections. Further, examining the much slower growth in the 50% category of occupations between the two projections leads us to conclude that even our imperfect understanding of offshoring in 2002 led to more reasonable projections in that set than in the earlier.

The actual sorted list of occupations isn't included with this paper since the project reported on here was meant primarily as a test of the process rather than a presentation of the results.

The process has been put in place as an accompaniment to the staffing pattern projections developed in the 2004-2014 OOSEP projections and will be used to evaluate the same set of occupations as in this analysis. The evaluation, however, will be carried out by the occupational outlook experts on the staff, resulting in a more informed process that will, hopefully result in a cleaner list of affected occupations. We hope to be able to report more fully on the results of that analysis later in 2005.

FOOTNOTES

[1] Normally, BLS projections are released in November of odd-numbered years. The 2012 projections were delayed three months due to the shift from the SIC to the NAICS industry classification system, a shift which left us without many of our traditional historical time series upon which our various models are based. BLS anticipates a return to publication in the November issue of the Monthly Labor Review with the 2004-2014 projections, to be released in late 2005.

[2] "Mass Layoff Statistics Data in the United States and Domestic and Overseas Relocation," Sharon P. Brown, Chief, Division of Local Area Unemployment Statistics, Bureau of Labor Statistics, U.S. Department of Labor, presented at EU-US Seminar on "Offshoring of Services in ICT and Related Services," Brussels, Belgium, December 13-14, 2004.

[3] John Tschetter, Office of the Secretary, U.S. Department of Commerce, personal communication.

[4] We experimented with allowing the response over a continuous range between 0 and 1, a response to the "it depends" school of thought. While this remains an open question, our initial findings were that while the scores changed slightly as a result, the rankings of the scores were virtually unaffected. Subscribing to the KISS school of thought, we elected to keep the 0 or 1 approach for now, subject to a later evaluation of detailed results.

This page intentionally left blank.

Forecasting Strategic Issues Facing the Veterans Health Administration II

Session Chair: Donald Stockford, Veterans Health Administration, U.S. Department of Veterans Affairs (donald.stockford@va.gov)

Forecasting the Veteran Population I: Imputation of Age and Race

Peter Ahn (peter.ahn@va.gov), Jin Kim, Stephen Meskin (stephen.meskin@va.gov), Eddie Thomas Jr. (eddie.thomas@va.gov), and Cathy Tomczak (cathy.tomczak@va.gov)
Department of Veterans Affairs

In forecasts of the veteran population an imputation method, derived using Lagrange multipliers, was used to solve three problems: 1) Imputing the joint distribution by age group (under 65 and over 64), gender, period of service, and state from the marginal distributions in Census 2000 SF3; 2) Imputing the joint distribution by single year of age, gender, period of service, race, and state, that match published SF4 marginal distributions; and 3) Imputing the joint distribution by age group (5-year), gender, race, state, and projected year from Census 2000 SF4 and a projection by age, gender, and state.

Forecasting the Veteran Population II: County Estimates

Peter Ahn (peter.ahn@va.gov) and Stephen Meskin (stephen.meskin@va.gov), Department of Veterans Affairs

Not only is all politics local but most VA services are local. Managing VA requires estimates of the veteran population in counties many years into the future. The traditional ratio method starting with the Census and then relying on county projections of the total population by age and gender is modified to account for the foreign born population and the number of armed forces in each county.

Estimates of Veterans Classified by VA Health Care Enrollment Priority, 4/1/2000 -9/30/2030: The Method of Successive Subtraction

George Sheldon, Office of the VA Actuary, Department of Veterans Affairs (george.sheldon@va.gov) or (george.g.sheldon@census.gov)

Approximately 7 million out of 25 million veterans are enrolled for VA health care where enrollment priority is largely based on disability status and income. Disability status is measured with respect to both service-connected disability entitlement status and clinically-defined “catastrophic” disabled status. We ask, “How would the entire veteran population be classified as of April 2000 if all were enrolled for VA health care?” Starting with Census 2000 sample data, we use a method of “successive subtraction” to take into account the hierarchical nature of the prioritization system and the fact that veterans often have several characteristics relevant to prioritization.

This page intentionally left blank.

Practical Forecasting Issues

Session Chair: Malcolm Harris, Manager, Market Intelligence and Support, U. S. Postal Service (malcolm.c.harris@usps.gov)

Forecasting When the Data Breaks from History: Postal Volumes After Anthrax and 9/11

Malcolm C. Harris, Sr., Manager, Market Intelligence and Support, U. S. Postal Service (malcolm.c.harris@usps.gov)

How does a forecaster deal with a decisive break in the historical data? The U.S. Postal Service has a long history of accurate revenue and volume forecasts based on econometric models. These forecasts drive budgeting, strategic planning, capital and facilities planning, pricing, revenue requirements for ratemaking, and workload allocations for the $69 billion government owned enterprise. In Fall, 2001, the Postal Service experienced its largest volume decline since the Great Depression after the terrorist attacks on 9/11 and the delivery of lethal Anthrax spores through the mail. This paper describes the approach developed in response to this challenge.

Forecasting Customer Data to Prioritize Marketing Efforts at the Postal Service

Stacey D. Harrison, Mathematical Statistician, Marketing Strategy and Support, U.S. Postal Service

The U.S. Postal Service's Marketing group has developed analytical models to improve its efforts to acquire, grow and retain the Postal Service's customer base. These models incorporate forecasting procedures utilizing historical customer transactional data and inputs from outside sources. The Revenue Forecasting, Customer Valuation and Growth-Defection Predictive models provide the various internal customers with strategic and actionable insight through the delivery of revenue expectations, strategic value targets and customer migration trends. This paper presents the forecasting methods used in these models. It also describes how the models contribute to the Postal Service's revenue generation efforts.

A Comparison of USDA's Agricultural Export Forecasts with ARIMA-based Forecasts

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

If USDA's forecasts of U.S. agricultural exports are no more accurate than those of an easily updatable model based on trends in monthly data, then using trends is preferred to USDA's more extensive efforts. This study compares the accuracy of USDA's FY2001-04 forecasts with forecasts based on trends for each commodity. ARIMA models utilizing the monthly data available at the time each USDA forecast was published were estimated. Out of 24 separate commodity forecasts examined, USDA forecasts were superior to ARIMA forecast in only 9 cases. ARIMA forecasts were superior in 11 cases, and there was no difference in 4 cases.

This page intentionally left blank.

Forecasting Customer Data to Prioritize Marketing Efforts at the Postal Service

Stacey Harrison
Marketing Strategy and Support, U.S. Postal Service

Introduction

Since the establishment of the position of Postmaster General by the Second Continental Congress in July of 1775 and the subsequent creation of the Post Office Department[1], the U.S. Postal Service has been tasked with the obligation of delivering mail. Its original duty was to ensure the delivery of communications between the Continental Congress and its army during the Revolutionary War. The Postal Reorganization Act of 1970 eliminated the Post Office Department and created the self-supporting organization we know today as the U.S. Postal Service. This "reorganization" signed into law by President Nixon allowed the Postal Service greater freedom to manage its operations in a more business like manner to achieve its mission of binding the nation together. The resulting impacts of this law can be seen today in a few of the Fiscal Year 2004 highlights:

• The number of delivery points has grown to over 142 million addresses[2]
• Total Factor Productivity growth of 2.4 percent for FY 2004 and 16.8 percent cumulative since 1972 respectively[3]
• $69 billion dollars of revenue generated from the delivery of over 206 billion mail pieces[4]
• Commercial revenue accounted for 72% or $50.4 billion dollars of the FY 2004 total revenue. Retail and Other revenues comprised the remaining 28% or $18.6 billion dollars[5]

A few things are clear from the first two Fiscal Year 2004 highlights. The increase in Total Factor Productivity demonstrates the success of the Postal Service's operations management. The Postal Service sells its services primarily to businesses. Commercial revenue is the largest contributor to revenue by a factor of just under three when compared to retail and other combined. The dominance of the commercial share of revenue is also illustrated in the proportion of mail originating from nonhouseholds versus households. In FY 2003 the ratio of mail pieces originating from nonhouseholds with respect to households was nearly eight to one or 178.5 and 22.8 billion respectively. Conversely, households received over 142 billion pieces of the 201 billion pieces of FY 2003 mail volume[6]. Furthermore, the $18.6 billion dollars in non-commercial revenue is generated by 20.5 million small businesses[7], which currently do not have an official commercial relationship with the Postal Service, and 111.2 million U.S. households[8]. Therefore, commercial revenue is the predominant source of revenue.

Given the importance of the revenue base to a self supporting business entity, the Marketing division of the Postal Service has developed strategies and products to enhance its revenues from the commercial and retail segments. Products and solutions such as, but not limited to, money orders, passport services, gift cards and self serve Automated Postal Centers (APCs) have assisted in generating retail revenue. Workshare discounts and Negotiated Service Agreements (NSAs) are a few of the initiatives that contribute to the generation of commercial revenue. Recognizing the importance of the commercial contribution to revenue, Marketing has developed an infrastructure and associated business rules to facilitate the implementation of these strategies and product solutions. Sound strategies and effective programs are based on data and analysis. This paper describes the models used to inform and focus the Postal Service's marketing. In the following sections, I describe the data and market segmentation that are used in these models. Next I present the Customer Valuation Model (CVM) and the forecasting techniques used in the CVM.

[1] U.S. Postal Service, *An American History 1775-2002*
[2] U.S. Postal Service, *2004 Comprehensive Statement on Postal Operations, p. 44*
[3] U.S. Postal Service, *2004 Comprehensive Statement on Postal Operations, p. 70*
[4] U.S. Postal Service, *FY 2004 Revenue, Pieces and Weight Report*
[5] U.S. Postal Service, *FY 2004 Financial Performance Report*
[6] U.S. Postal Service, *2003 Household Diary Study*
[7] Based on U.S. Small Business Administration's estimation of 22.9 million small businesses for 2002 minus 2.4 million active accounts in the U.S. Postal Service's Preferred commercial account segment
[8] U.S. Census Bureau, Current Population Reports, *America's Families and Living Arrangements: 2003*

Then I describe the segment forecasting model. The last section is a conclusion.

Background

In order to efficiently allocate internal resources and effectively manage its customer base, a structure has been developed for the commercial sector which divides this base into segments largely based on historical revenue. Accounts with the highest levels of postage revenue, greater than ten million dollars, are classified in the National segment and comprise 33% or 16.6 billion dollars of FY 2004 revenue. National accounts are typically large corporations and leaders in their respective industries. Second, the Premier segment, which accounts for 44% or $22.1 billion dollars of FY 2004 revenue, consists of those accounts with revenues between $250,000 and $10 million dollars. Typically, but not necessarily, a Premier account is smaller in postal revenue and size when compared to a National account. The Preferred Plus segment accounts have been modeled as high growth potential accounts. Essentially the Preferred Plus segment is an elevated subset of the Preferred segment. This elevation is based on criteria which identifies accounts as having the potential to be promoted to Premier segment status. Lastly, and the largest group by customer count, is the Preferred segment. This is a catch-all segment for those accounts not classified in the first three segments. The overwhelming majority of these accounts can be classified as small businesses. Given the inherent nature of these accounts the churn, the rate at which new accounts are added and existing accounts are declassified as active**,** associated with this segment is immense. For example, in Marketing's Commercial Business Customer Information System database, henceforth referred to as CBCIS, the number of Preferred segment customers is over 6 million. However, when accounting for the churn associated with this segment, the number of active accounts, defined as those with positive revenue within the prior fiscal year and year-to-date period, is reduced to just over 2.4 million.

Account segments are further defined by the business rules which dictate the manner in which Marketing treats these segments. The first three segments, National, Premier and Preferred Plus are managed accounts. That is to say an account manager is assigned to each of these accounts to promote and service their interactions with the Postal Service. However, based on limited resources, brief customer life cycle and their inherent low rate of return, the Preferred segment customers are unmanaged. Therefore, the relationship or contact with these businesses is more limited. In addition to allocating accounts into, largely, revenue segments, the Postal Service further breaks down customers by industry classifications. Using these categories the Postal Service can tailor its programs and sales efforts within each industry and segment.

ANALYTICAL MODELS

There exists a rich source of historical commercial customer transactional detail stored in the CBCIS database. Given the large size and high level of granularity of the data, the Postal Service is able to conduct a multitude of analyses to increase its understanding of the market and customer base. As a result of this understanding, Marketing is able to make informed business decisions, form strategies and achieve its organizational goals. One such input to the higher level of market/customer knowledge is the output of analytical models. These analytical constructs are based on transactional data such as customer information, salient customer identification characteristics, time periods, and associated revenue and volume data according to product. The Customer Valuation, Revenue Forecasting, Growth-Defection, and Growth-Defection Predictive models are a few examples of the analytical tools which utilize the transactional data stored in CBCIS. Of these four models, two incorporate forecasting: the Customer Valuation Model (CVM) and Revenue Forecasting models.

CUSTOMER VALUATION MODEL (CVM)

Knowing a commercial customer's product usage and associated revenues and volumes allows relatively simple comparisons and trend analyses. These analyses provide high level insights but are limited in scope. In contrast, the Customer Valuation Model provides additional insight beyond the one dimensional view of percentage change comparisons.

Purpose

The purpose of the CVM is to identify the value of each customer and each customer's potential for growth. Prior to explaining the methodology of the model a discussion with regard to the purpose of the model is in order. Not surprisingly, the objectives of the model are aligned with the business rules which were set forth to manage the commercial customer segment. First, the model calculates to which commercial segment a customer should be assigned. Next, a second purpose of the model is to align internal resources so that the Marketing organization can better serve the needs of the most valuable current and future customers. In order to fulfill the second purpose, the model determines the growth potential for each commercial customer. Lastly and closely related to the first purpose, the model prioritizes customer interaction in order to maximize the return on investment.

Fundamentally, the purpose of any model or analysis will be aligned with an internal unit's functional objectives and/or goals. The CVM model serves multiple Marketing groups by assisting them to achieve their objectives. Internal customers such as Sales, Advertising, Market Research, Segment Managers, and the Business Service Network (BSN) use the outputs of the model to better manage their business processes of focusing on the voice of the customer.

Inputs

The model not only uses transactional revenue data but also incorporates Postal Service cost-revenue analysis data, package market estimates supplied by an outside vendor, and the top 100 leading national advertisers according to Advertising Age. Utilizing these various inputs creates a model which provides a multidimensional view of the customer and its performance with respect to Postal Service revenue, contribution, and market share.

Outputs

Outputs of the CVM, as the name suggests, are based on several value categories. Value categories provide information which enables the organization to gain a multidimensional view in which to base their strategies and decisions regarding the customer. The various value outputs are:

Actual Value – Defines the current worth of a customer to the Postal Service. The operational premise is that even if the organization does not alter the interaction and/or treatment of the customer, the customers they retain represent a certain value to the Postal Service.

Total Strategic Value – Defines the unrealized growth potential of a customer. The assumption is that if the Postal Service were to modify its interaction and/or treatment of the client by focusing on specific customer needs, the customer's value to the Postal Service would grow. The amount by which a customer's value grows is the Strategic Value. Total Strategic Value is the summation of three components: Strategic Value Advertising share, Strategic Value Expedited Package Services (EPS) share, and Strategic Value Other share.

Strategic Value Advertising share – Defined as the unrealized opportunity for Advertising mail.

Strategic Value EPS share – Defined as the unrealized opportunity of package services such as the Postal Service's Express, Priority, and Parcel shipment products.

Strategic Value Other share – Defined as the unrealized opportunity for those products not in Advertising share or EPS share in which the Postal Service has a reasonable ability to influence. Residual Meter, International, and 15% of First-Class Mail are the products included in the Other share.

Value Tier – Defined as a designation resulting from the combination of the Actual and Strategic Value deciles for a customer. The four designations are Most Valuable Customers-MVC, Most Growable Customers-MGC, Migrators-MIG, and Below Threshold-BZ.

Tier Level – Defined as a designation, between one and five with one being the highest, based on Actual Value. The number of customers per tier level is not fixed. Limited fluctuations occur with each update of the model.

Business Applications

Workload Points

One business application is its Workload Point assignment. The model produces scores for account managers and BSN representatives to differentiate accounts based on their Actual and Strategic Values. This provides information to the sales and service organizations in terms of where to align the focus of their efforts. Customers with higher values are given added attention. Each account is assigned Actual Value Workload Points (8, 4, 3, 2, 1, 0) and Strategic Value Workload Points (8, 4, 3, 2, 0). Regardless of segment status or account designation, customers within the managed account base are assigned

Workload Points based on their Actual and Strategic Value.

Campaign Customer Selection

The CVM is an effective tool for advertising campaign targeting; it identifies customers by both growth potential and current value. For example, Inside Sales uses the Tier Level and Value Tier values to develop calling lists which target customers for sales efforts. Developing leads for Package Services involves using Strategic Value Expedited Package Services (EPS) share to select Preferred Plus accounts with the most potential for growth in package services spend.

Identifying New Preferred Plus Accounts

The CVM was used to select the Preferred Plus accounts. These accounts have the potential to achieve Premier status (i.e. higher Postal Service revenues) with the assistance of added attention and treatment. These accounts were classified using five modeling attributes:

1. Growth/Defection model output, a revenue performance measurement which indicates whether a customer is growing relative to its industry peers
2. Actual revenue growth
3. Revenue
4. Actual Value
5. Strategic Value

Preferred Sub-Segmentation

Given the much larger number of smaller customers that make up the Preferred Plus segment, sub-segmentation is needed. The marginal benefit from employing focused strategic treatment to this group is large. Therefore, in order to foster the growth of accounts belonging to the Preferred segment, customer subsets are created which are conducive to targeting and receptive of treatment strategies. Three criteria, derived from the CVM model, are used to form these groups, namely, Actual Value, real growth, and revenue.

Methodology

The CVM model is updated semi-annually on a fiscal year basis (after months 6 and 12). Each value category output, Actual Value, Strategic Value Advertising Share, Strategic Value EPS share, and Strategic Value Other share has a relatively unique methodology.

Actual Value Methodology

A high level overview with regard to computing the Actual Value component of the CVM model illustrates four major process rules:

1. Actual Value is based on three historical years and two years of forecasted contribution.
2. Managed accounts, National, Premier, and Preferred Plus are scored at the account level.
3. Unmanaged Accounts are scored at the site level.
4. Mail Service Provider (MSP) accounts are not scored.

There are 3 main steps in computing Actual Value:

Step 1: Assigning Cost Information

Based on the Postal Service's Cost Revenue Analysis (CRA) report, cost per piece data is matched to product classes applied to the product-category level. Prior to processing the matching component, certain assumptions have to be made concerning the calculation of cost per piece with regard to the several product classes:

Postage Due and Pre-Cancelled Stamps
-Customer meters and postage due assumed product cost equal to First-Class stamp revenue.
-Pre-cancelled stamps revenue is assumed to be associated with Standard A product costs

Catalogs, Bound Printed Matter
Assumed cost and revenue per piece for bound printed matter

Residual Meter
Calculated a contribution margin based on a weighted average across the following five product classes: Single-Piece Letters, Priority Mail, Express Mail, Parcel Post, and International Mail

Step 2: Determining Contribution

Operating under the principle that defines customer contribution as a function of revenue and cost of services sold, contribution is computed by customer, product, and month. Contribution can be calculated by this equation, Contribution = Revenue * Contribution_Margin; where Contribution_Margin meets the following condition:

Contribution_Margin = (Revenue per Piece – Cost per Piece) / Revenue per Piece

Step 3: Calculate Actual Value

Actual value is computed as the Net Present Value (NPV) based on two years of forecasted future contribution using a 12% discount rate.

Example: $FC_{t=1} / (1 + d) + FC_{t=2} / (1 + d)$ ^^{2} where, FC = Forecasted Contribution, t = year, and d = discount rate.

Forecasting Element

Two years of future contribution is forecasted by using the Holt-Winters (exponential smoothing with a trend component) method. Depending on the smoothing constant (w) chosen, a figure from zero to one, more weight is given to recent or more historical periods. A value closer to one gives more weight to recent periods and a value closer to zero lends more weight to prior periods. The projections are computed by the following equation:

$\pi = w * x + (1 - w) * \pi_{t-1} + \varepsilon$ where,

π_t = smoothed value of contribution for the current time period
w = smoothing constant; a weight, ranging from zero to one, will be chosen so that the standard error of the estimate is minimized
x_t = contribution in the current time period
π_{t-1} = smoothed value in the previous time period
ε = error term

Strategic Value Advertising Share Methodology

How is a customer's strategic value calculated? There are two different methods for calculating Strategic Value Advertising share. Which is used depends on whether a company is listed in AdAge's Top 100 Advertisers list. Again, assumptions are required about the mix and proportion of product classes and the industry direct mail spend benchmark.

Method 1: For companies that are found in the Top 100 Advertisers according to AdAge

Step 1: Determine Direct Mail spend
- Assume postage is 37.6% of Direct Mail budget
- Direct Mail postage (most recent 12 months) = 100% Standard + 100% Catalogs (BPM) + 12% First-Class mail
- Direct Mail budget = Direct Mail postage/.376

Step 2: Determine Direct Mail spend percentage share of total advertising budget
- Determine Direct Mail budget share of total advertising for each of the accounts
- Calculate average and maximum shares by: segment, industry, and employee range group
- Grow maximum share by 10%. This is the ceiling for the particular peer group
- Determine gap between current Direct Mail budget and what it would be if that account was spending at the maximum + 10% level. The postage share of this (37.6%) gap is the Postal Service's opportunity for advertising mail.

Method 2: For companies not found in the list of the Top 100 Advertisers

In this method, growth is estimated by comparing current spend amongst customers that have the same industry classification, account segment, and employee count range.

Step 1: Calculate Direct Mail postage and budgets the same as in Method 1.
Step 2: Create groups of companies that have same industry classification, account segment, and employee count range.
Step 3: Cap the Strategic Value at the 75th percentile within the group.
Step 4: Determine Strategic Value based on difference between current spend and the third quartile.

For example:
- Group1: All Premier accounts in FINANCIAL SERVICES with 100 – 499 employees
- 75th percentile spend within group = \$500K
- Company A's spend = \$300K
- Company A's direct mail Strategic Value = \$500K - \$300K = \$200 K
- Company A's direct mail Strategic Value will be reported at the contribution level as well
- For those companies that are above the 75th percentile we redo the analysis within that group twice more for managed accounts and five times more for unmanaged accounts.

Strategic Value EPS Share Methodology

- EPS is the sum of three mail categories: Express, Parcels, and Priority.

- Managed and Unmanaged accounts are calculated separately.
- Market share data, supplied from an independent source, is incorporated at the industry segment code and employee size groupings.
- Cap the potential share the Postal Service can obtain through an analysis of high spend distribution within an industry segment/employee size group.
- Assume that accounts who are setting the ceiling today are maximized and do not present any additional opportunity for growth.

For example:

- Segment: Financial Services, Employees 1 – 19
- Average Postal Service share Express Mail = 5%
- Highest share the Postal Service has within this segment = 15% (Please note: this could also be at the 95th percentile depending on outlier analysis)
- Company B's current Express Mail spend (most recent 12 months) = $40 assume that to be 5% of total Express Mail budget
- Company B's Total Express Mail budget = $40 / 5% = $800
- Company B's Max spend Express Mail = $800 * 15% = $120
- Company B's Strategic Value Express Mail = $120 - $40 = $80
- Company B's Strategic Value Express Mail is also reported at the contribution level as well

Strategic Value Other Share Methodology

This calculation is very similar to Method 2 of the Strategic Value Ad Share calculation

Step 1: Calculate Other Mail spend for past 12 months (100% Residual Meter, 100% International, 15% First-Class) for each account.
Step 2: Create groups of companies that have the same industry, segment, and employee size range.
Step 3: Cap the Strategic Value at the 75th percentile within the group
Step 4: Determine Strategic Value based on difference between current spend and third quartile.
Step 5: For those accounts that are above the 75th percentile, redo this analysis within this group twice more for managed accounts and five times more for unmanaged accounts.

REVENUE FORECASTING MODEL

Forecasting revenue at the Postal Service assists the organization with its planning activities on many levels. Moreover, Marketing employs three different commercial revenue forecasts for its plan formulations. Two of three forecast models reside in Marketing: Customer Resource Management and Customer Analytics; the third is conducted by the Finance department. Finance's model generates the overall commercial and retail revenue forecast for the Postal Service and is the source referred to for official projections. The first two models tend to be customer centric and aligned with business applications specific to Marketing.

Overview of Forecasting Models Used by Marketing

Finance's model is an econometric forecast using historical product data and economic drivers. This model creates revenue projections at the product level. However, the model is limited in that it does not take other views into consideration. Since the model is based on the systematic economic drivers of the mail, the model may not fully reflect recent trends.

Customer Resource Management (CRM) utilizes, on an annual basis, a time series based forecasting model, using three years of historical CBCIS data, which produces projections by account, product, and fiscal year month for managed commercial accounts: National, Premier, and Preferred Plus. Specifically, the model uses the Holt- Winters method. Due to the lack of sufficient data points with regard to unmanaged commercial accounts, an indication of the high degree of churn, the forecast for the Preferred segment is conducted at an aggregated postal district level by product.

As with CRM, the Customer Analytics group also employs a time series based forecasting model compiled on a semi-annual basis. The model projects revenues by the combination of account segment and product. In addition, a time series forecast approach is generated for revenue per Postal Area for the nine areas which encompass 80 Postal districts.

Process, Inputs, and Outputs

A large internal consumer of the three forecasts is the Sales function of Marketing. Forecasts are crucial for the allocation of Sales resources and goal setting purposes. However, the forecasts alone do not dictate the final expectations of sales staff. There is a collaborative process by which the final expectations are established using the three forecasts and market intelligence gained by account representatives in the field.

The CRM forecasting model lends itself to the Sales function due to the nature of what is being forecasted. As mentioned previously, the CRM group conducts a forecast by customer account, product, and fiscal year month. Once the forecast is compiled, the headquarters unit of Sales divvies the results and disseminates them to the respective account managers and Area Sales analysts. The account managers, armed with the latest forecast results, attempt to discern whether the forecasts are reasonable by speaking with the clients for whom they have accountability in order to assess their current state of affairs and future performance expectations. Based on this intelligence gathering and current insight, the account managers will provide their feedback to the Area Sales analysts on the attainability of the projections. In addition, based on the staff interaction with their customer base, movement of forecast transfers between accounts will be noted in the feedback analysis. The Area Sales analysts will collect all feedback from their Area, review the suggested changes, and reconcile any forecast transfers for submission to the headquarters office of Sales. HQ Sales, upon reception of the feedback from Area Sales, will review the suggestions for credible reasons for modifications to an account's forecast. A reasonable qualitative explanation must be submitted to be eligible for consideration; not just a comment such as "forecast too high".

Upon incorporating the agreed to adjustments, the revenue projections are summed at the segment level: National, Premier, and Preferred Plus. A comparison is then made to the second Marketing forecast produced by Customer Analytics. The comparison is logical based on the fact that the Customer Analytics model produces forecasts by a segment-product combination. This model is another dimension by which to develop the ultimate Postal Service revenue goals by account. You may be asking, "Why not just sum the CRM results by account to obtain your segment and product goals?" Due to the more macro level of the segment level forecast the Customer Analytics model is more accurate. If the segment results vary too greatly the account level data produced by CRM will be reconciled to the Customer Analytics data at the segment level.

Since the projections produced by Finance are the official Postal Service forecasts, the total revenue goals for Sales must equal the commercial revenue total created by Finance. Therefore, given the requirement that the segment-product combination forecast must equal the official forecast, the account level projection will, by default, match the forecast by Finance.

Methodology: Customer Analytics Model

As a supplement to Finance's econometrics-based commercial revenue projections, Customer Analytics provides revenue forecasts using a different optimization – the time series model. The model considers cyclical fluctuations, seasonality, and past revenue performance to predict future Postal Service revenues. Using revenue projections from both the econometrics based and time series models, the Postal Service is provided a multifaceted view of future revenue trends. A segment-product combination and Area level forecasts are generated from the model. There are two phases to running the model: 1) forecasting revenue and 2) forecasting product type percentages to determine the product type composition of the revenue.

Phase 1: Forecasting Revenue

In this phase of forecasting revenue, nine models are estimated for each time series. Models with the best forecasting statistics are chosen for their forecast results. The statistic ultimately used to determine the selected time series is the Theil's U statistic-value. This statistic measures the accuracy, as opposed to the goodness-of-fit statistics such as F-statistic and R Square-statistic, of a given model. The equation for the Theil's U value is as follows:

$$U = \frac{\left(\frac{1}{n}\sum_{t=1}^{n}\left(Y_t^s - Y_t^a\right)^2\right)^{0.5}}{\left(\frac{1}{n}\sum_{t=1}^{n}\left(Y_t^s\right)^2\right)^{0.5} + \left(\frac{1}{n}\sum_{t=1}^{n}\left(Y_t^a\right)^2\right)^{0.5}}$$

Values of U will range from 0 to 1 where 0 indicates a perfect historical fit and one indicates the poorest fit. An example of a well performing model will have a Theil's U value in the range of 0.03 to 0.1.

In performing the estimation of the models, three forecasting methodologies are employed. Furthermore, three variants of these models are computed for a total of nine different models. The three standards are:

1) Exponential
This is an exponential smoothing model which performs well when the trend of the data is smooth or with few fluctuations.

2) Holt-Winters
This model performs well when the fluctuations in the data are considerable and the data exhibits high levels of seasonality.

3) Stepwise Autoregressive
This model captures inter-period inertia in its methodology. For example, the model would capture the inertia between the first and third periods, the inertia between the fifth and seventh periods, etc. The model performs better when high levels of autocorrelation are present.

Each variant to each of the three models is based on the trend designation of: 1) constant, 2) linear, and 3) quadratic for a total of nine models.

Phase 2: Forecasting Product Percentages

As with the segment-product combination forecasts, the forecasting of product percentages phase follows the same methodology of running the nine models and using the models with the optimal Theil's U statistic.

Area Revenue Forecast

This model also uses the Theil's U statistic when choosing the model with the best performance. As previously stated, the Postal Service Area revenue forecast is used as a benchmark for the Sales organization's revenue goal formulations. However, unlike the CRM Area forecast, this model incorporates all of the commercial segments not just the unmanaged Preferred segment.

Conclusion

A forecasting model can only be as good as the underlying data. The CBCIS data used in the models is very detailed. Customer, product, revenue, volume, product weight and industry are a few of the data fields. There is an abundance of rich and granular data; however, in many instances a complete picture of a customer is absent. This absence is due to an ultimate end user identification issue resulting from a relationship with mail intermediaries. These companies are defined for postal purposes as Mail Service Providers and Publishing and/or Printing establishments. The function of a MSP is to offer discounted mailing services to customers by commingling the mailings of many companies into one consolidated mailing to reap the benefits of discounted Postal Service workshare rates. An issue arises when an intermediary does not complete a mailing statement in its entirety. Intermediaries are required to identify the ultimate mailers for a given consolidated shipment; the policy is quite often not enforced. Hence, this is a concern to Marketing given that approximately 21% or $10.5 billion dollars of the commercial revenue in FY04 was a result of intermediary mailings. Of this total, only 24% or $2.5 billion dollars was attributed to the ultimate end user. Therefore, we cannot properly allocate the remaining $8 billion dollars. In addition, the data which cannot be allocated to an end user is recognized as revenue from the intermediary. Further exacerbating the issue is the inability to accurately analyze an intermediary's performance. Consolidation between inter-mediaries, inter-industry cannibalization of market shares and the aforementioned lack of manifest transparency are all factors which hamper the effectiveness of measuring the performance of mail intermediaries.

The potential resolution of the end user identification issue will lead to a higher quality of data; ergo, more accurate models. However, despite the limitations of the data, the models are able to provide results for the creation of insightful strategies and business decisions. The CVM and segment revenue forecasting models are tools which provide accurate, if not precise, direction for prioritizing marketing efforts at the Postal Service.

APPENDIX

Table 1
Historical Revenue by Source
(Data in Millions)

Revenue Source	FISCAL YEAR				
	2000	**2001**	**2002**	**2003**	**2004**
Commercial	47,592.2	48,957.3	48,801.4	50,661.0	51,088.7
Retail	16,106.2	16,459.1	16,752.2	17,334.4	17,221.0
Other	565.5	538.4	512.2	547.1	718.8
Total	**$64,264.1**	**$65,954.9**	**$66,066.0**	**$68,542.5**	**$69,028.5**

Source: USPS Revenue, Pieces and Weight report
Totals may not match due to rounding

Table 2
Customer Count by Commercial Segment

Commercial Segment	Number of Accounts
National	221
Premier	13,499
Preferred Plus	32,034
Preferred*	2,472,419
Total	**2,518,173**

Source: USPS-CBCIS database
*Preferred segment total only includes those accounts with positive revenue within the previous 17 months (Oct FY04 – Feb FY05)

Table 3
Historical Revenue & Volume Totals by Commercial Segment
(Data in Millions)

Commercial Segment	Revenue			Volume		
	FY 02	**FY03**	**FY04**	**FY02**	**FY03**	**FY04**
National	16,307	16,661	16,658	54,369	55,119	56,918
Premier	20,590	21,956	22,171	73,366	76,562	80,062
Preferred Plus	1,823	2,554	2,567	4,483	6,706	6,939
Preferred	9,228	8,757	8,993	15,135	12,643	14,209
TOTAL	**$47,948**	**$49,928**	**$50,388**	**147,353**	**151,031**	**158,128**

Source: USPS-CBCIS database
Totals may not match due to rounding
Total commercial revenue in CBCIS will not equal Revenue, Pieces, and Weight report due to different reporting systems.

Table 4
Revenue & Volume Totals by Industry Classification
(data in millions)

Industry Classification	Revenue			Volume		
	FY02	FY03	FY04	FY02	FY03	FY04
Financial Services	8,698	9,034	9,220	18,585	19,222	21,592
Government	2,243	2,326	2,285	4,018	5,216	5,534
Mail Order Catalogs E-tailer	2,834	2,835	2,742	9,714	9,559	9,412
Mail Service Providers	4.677	5,353	5,563	26,294	28,704	29,794
Manufacturing and Wholesalers	3,838	3,903	3,798	9,538	9,221	9,425
Publishing and/or Printing	4,853	5,042	5,069	17,880	17,755	18,498
Retail	3,351	3,586	3,657	13,437	14,003	14,350
Services	12,097	12,444	12,328	33,604	33,583	33,965
Telecommunications	1,213	1,323	1,385	4,088	4,327	4,667
Utilities	629	672	665	1,824	1,925	1,940
Not Listed or Classified	1,767	1,908	1,996	4,125	4,309	5,011
Unknown Segment	1,748	,1504	1,682	4,245	3,207	3,940
TOTAL	**$47,948**	**$49,928**	**$50,388**	**147,353**	**151,030**	**158,129**

Source: USPS – CBCIS database

Table 5
CVM: Top 10 Actual Value Accounts

Customer	Industry	Actual Value $
A	Financial Services	436,367,474
B	Financial Services	410,999,394
C	Financial Services	407,636,727
D	Telecommunications	365,236,137
E	Telecommunications	297,487,859
F	Financial Services	296,534,662
G	Financial Services	267,477,537
H	Financial Services	206,828,806
I	Government	196,752,823
J	Retail	193,244,592

Data source for CVM: USPS – CBCIS database

Table 6
CVM: Top 10 Strategic Value Accounts

Customer	Industry	Strategic Value $
F	Financial Services	107,847,142
B	Financial Services	95,822,541
K	Financial Services	74,472,044
L	Financial Services	67,931,978
M	Financial Services	66,678,028
N	Financial Services	65,501,829
O	Services	60,801,812
P	Financial Services	54,551,241
J	Retail	51,802,872
Q	Retail	49,809,250

Data source for CVM: USPS – CBCIS database

Table 7
Segment Owner View Forecast of Commercial Revenue by Segment per Product

FY05 Marketing Expectations
- Segment Owner View
- Prepared August 18, 2004

SEGMENT	PRODUCT	Actual FY02 YE	Actual FY03 YE	Actual FY04 YTD*	Forecast FY04 M11 - M12 FORECAST	Forecast FY04 YE	Growth from FY03	Forecast FY05 YE	Growth from FY04
National	CATALOGS (BPM)	$ 267,143,591	$ 290,253,617	$ 229,527,267	$ 55,439,524	$ 284,966,791	-1.8%	$ 286,534,201	0.6%
	EXPRESS	$ 21,849,081	$ 17,989,563	$ 12,162,994	$ 2,522,466	$ 14,685,460	-18.4%	$ 16,735,833	14.0%
	FIRST CLASS	$ 7,855,623,914	$ 7,696,467,765	$ 6,228,141,396	$ 1,207,778,186	$ 7,435,919,582	-3.4%	$ 7,358,987,136	-1.0%
	INTERNATIONAL	$ 80,931,587	$ 82,705,834	$ 77,937,798	$ 14,383,595	$ 92,321,393	11.6%	$ 99,531,125	7.8%
	MISC FEES	$ 4,773,534	$ 10,475,229	$ 3,449,031	$ 7,026,198	$ 10,475,229	0.0%	$ 11,606,554	10.8%
	PARCELS	$ 638,946,563	$ 647,218,981	$ 487,652,502	$ 63,629,574	$ 551,282,076	-14.8%	$ 416,369,274	-24.5%
	PERIODICALS	$ 819,998,767	$ 852,817,953	$ 706,367,357	$ 130,854,959	$ 837,222,316	-1.8%	$ 810,183,585	-3.2%
	POSTAGE DUE	$ 216,024,539	$ 226,824,728	$ 180,360,505	$ 31,839,152	$ 212,199,657	-6.4%	$ 204,095,608	-3.8%
	PRIORITY	$ 597,007,360	$ 495,549,896	$ 395,841,496	$ 70,302,120	$ 466,143,616	-5.9%	$ 378,540,658	-18.8%
	STANDARD	$ 5,629,272,893	$ 6,191,448,775	$ 5,438,098,874	$ 1,060,380,074	$ 6,498,478,948	5.0%	$ 6,540,298,609	0.6%
	National Total	**$ 16,131,571,829**	**$ 16,511,752,341**	**$ 13,759,539,220**	**$ 2,644,155,847**	**$ 16,403,695,067**	**-0.7%**	**$ 16,122,882,584**	**-1.7%**
Premier	CATALOGS (BPM)	$ 171,110,589	$ 184,166,527	$ 153,024,389	$ 35,619,377	$ 188,643,766	2.4%	$ 187,301,110	-0.7%
	EXPRESS	$ 36,019,046	$ 35,186,743	$ 24,601,980	$ 5,113,232	$ 29,715,212	-15.5%	$ 29,558,298	-0.5%
	FIRST CLASS	$ 11,135,700,181	$ 11,699,920,408	$ 9,668,729,000	$ 1,850,663,533	$ 11,519,392,533	-1.5%	$ 11,235,586,691	-2.5%
	INTERNATIONAL	$ 85,764,457	$ 80,356,393	$ 67,898,609	$ 13,307,742	$ 81,206,351	1.1%	$ 77,861,869	-4.1%
	MISC FEES	$ 2,088,493	$ 2,789,907	$ 2,465,502	$ 324,405	$ 2,789,907	0.0%	$ 3,091,217	10.8%
	PARCELS	$ 132,038,857	$ 127,011,592	$ 107,124,974	$ 18,520,546	$ 125,645,520	-1.1%	$ 115,365,592	-8.2%
	PERIODICALS	$ 1,031,094,792	$ 1,047,623,862	$ 838,134,651	$ 169,927,157	$ 1,008,061,808	-3.8%	$ 1,000,208,992	-0.8%
	POSTAGE DUE	$ 280,160,203	$ 327,401,958	$ 267,481,123	$ 53,841,038	$ 321,322,161	-1.9%	$ 317,496,044	-1.2%
	PRIORITY	$ 813,560,729	$ 731,938,114	$ 547,375,845	$ 106,880,704	$ 654,256,549	-10.6%	$ 612,271,510	-6.4%
	STANDARD	$ 7,225,736,901	$ 7,933,755,280	$ 6,701,514,448	$ 1,425,341,288	$ 8,126,855,736	2.4%	$ 8,076,556,261	-0.6%
	Premier Total	**$ 20,913,274,248**	**$ 22,170,150,784**	**$ 18,378,350,521**	**$ 3,679,539,022**	**$ 22,057,889,543**	**-0.5%**	**$ 21,655,297,583**	**-1.8%**
Preferred Plus	CATALOGS (BPM)	$ 13,450,507	$ 25,694,334	$ 19,792,683	$ 3,838,587	$ 23,631,270	-8.0%	$ 24,600,200	4.1%
	EXPRESS	$ 14,658,755	$ 17,946,857	$ 13,771,053	$ 2,772,343	$ 16,543,396	-7.8%	$ 18,510,177	11.9%
	FIRST CLASS	$ 1,113,678,745	$ 1,506,716,683	$ 1,210,509,047	$ 240,536,182	$ 1,451,045,229	-3.7%	$ 1,493,119,398	2.9%
	INTERNATIONAL	$ 1,387,130	$ 4,916,638	$ 6,608,859	$ 1,592,594	$ 8,201,453	66.8%	$ 9,407,143	14.7%
	MISC FEES	$ 10,094	$ 19,961	$ (46,407)	$ 66,368	$ 19,961	0.0%	$ 22,117	10.8%
	PARCELS	$ 2,042,455	$ 3,596,009	$ 3,884,934	$ 800,999	$ 4,685,933	30.3%	$ 4,693,237	0.2%
	PERIODICALS	$ 47,757,492	$ 60,416,184	$ 52,633,838	$ 10,563,910	$ 63,197,748	4.6%	$ 66,147,683	4.7%
	POSTAGE DUE	$ 21,996,729	$ 36,244,073	$ 33,446,304	$ 6,630,063	$ 40,076,367	10.6%	$ 42,230,742	5.4%
	PRIORITY	$ 105,587,934	$ 137,149,683	$ 112,260,617	$ 22,119,780	$ 134,380,397	-2.0%	$ 142,545,185	6.1%
	STANDARD	$ 599,065,151	$ 1,036,206,930	$ 816,356,932	$ 175,340,561	$ 991,697,493	-4.3%	$ 1,109,641,367	11.9%
	Preferred Plus Total	**$ 1,919,634,992**	**$ 2,828,907,351**	**$ 2,269,217,860**	**$ 464,261,387**	**$ 2,733,479,247**	**-3.4%**	**$ 2,910,917,249**	**6.5%**
Preferred (Unmanaged)	CATALOGS (BPM)	$ 29,122,481	$ 27,779,679	$ 27,800,157	$ 6,192,984	$ 33,993,141	22.4%	$ 34,574,402	1.7%
	EXPRESS	$ 45,864,911	$ 40,648,004	$ 32,888,790	$ 6,810,430	$ 39,699,220	-2.3%	$ 46,060,753	16.0%
	FIRST CLASS	$ 6,211,114,359	$ 5,952,420,487	$ 5,133,293,650	$ 1,034,399,097	$ 6,167,692,747	3.6%	$ 6,631,450,940	7.5%
	INTERNATIONAL	$ 9,946,633	$ 5,347,449	$ 8,566,633	$ 1,821,373	$ 10,388,006	94.3%	$ 11,665,238	12.3%
	MISC FEES	$ 57,972	$ 54,981	$ (94,141)	$ 149,122	$ 54,981	0.0%	$ 60,919	10.8%
	PARCELS	$ 8,884,999	$ 5,483,492	$ 7,623,791	$ 1,374,273	$ 8,998,064	64.1%	$ 12,397,289	37.8%
	PERIODICALS	$ 203,228,954	$ 198,310,853	$ 165,955,520	$ 34,260,286	$ 200,215,806	1.0%	$ 220,836,819	10.3%
	POSTAGE DUE	$ 102,970,062	$ 107,784,529	$ 112,785,088	$ 21,876,813	$ 134,661,901	24.9%	$ 139,220,921	3.4%
	PRIORITY	$ 633,600,490	$ 601,347,853	$ 518,657,808	$ 105,300,657	$ 623,958,464	3.8%	$ 678,945,718	8.8%
	STANDARD	$ 1,738,546,216	$ 1,478,465,961	$ 1,537,257,351	$ 306,136,471	$ 1,843,393,822	24.7%	$ 1,876,547,683	1.8%
	Preferred (Unmanaged) Total	**$ 8,983,337,077**	**$ 8,417,643,288**	**$ 7,544,734,647**	**$ 1,518,321,506**	**$ 9,063,056,153**	**7.7%**	**$ 9,651,760,682**	**6.5%**
Commercial Not in CBCIS			**$ 756,382,617**	**$ 529,201,752**	**$ 83,464,091**	**$ 612,665,843**	**-19.0%**	**$ 551,399,259**	**-10.0%**
Commercial Total			**$ 50,684,836,381**	**$ 42,481,044,000**	**$ 8,389,741,854**	**$ 50,870,785,854**	**0.4%**	**$ 50,892,257,357**	**0.0%**

Data source for Revenue Forecasting model: USPS-CBCIS database.

Table 8
Product Summary for Commercial Revenue Forecast

FY05 Marketing Expectations
- CBCIS Product Summary
- Prepared August 18, 2004

	Actual			Forecast				
PRODUCT	**FY02 YE**	**FY03 YE**	**FY04 YTD***	***FY04 M11 - M12 FORECAST***	***FY04 YE***	**Growth from FY03**	***FY05 YE***	**Growth from FY04**
CATALOGS (BPM)	$ 480,827,168	$ 527,894,157	$ 430,144,496	$ 101,090,472	$ 531,234,968	0.6%	$ 533,009,913	0.3%
EXPRESS	$ 118,391,793	$ 111,771,167	$ 83,424,817	$ 17,218,470	$ 100,643,287	-10.0%	$ 110,865,061	10.2%
FIRST CLASS	$ 26,316,117,199	$ 26,855,525,341	$ 22,240,673,094	$ 4,333,376,998	$ 26,574,050,091	-1.0%	$ 26,719,144,165	0.5%
INTERNATIONAL	$ 178,029,807	$ 173,326,314	$ 161,011,899	$ 31,105,305	$ 192,117,204	10.8%	$ 198,465,376	3.3%
MISC FEES	$ 6,930,093	$ 13,340,078	$ 5,773,985	$ 7,566,093	$ 13,340,078	0.0%	$ 14,780,806	10.8%
PARCELS	$ 781,912,874	$ 783,310,074	$ 606,286,201	$ 84,325,392	$ 690,611,593	-11.8%	$ 548,825,393	-20.5%
PERIODICALS	$ 2,102,080,005	$ 2,159,168,852	$ 1,763,091,366	$ 345,606,311	$ 2,108,697,677	-2.3%	$ 2,097,377,078	-0.5%
POSTAGE DUE	$ 621,151,533	$ 698,255,288	$ 594,073,020	$ 114,187,066	$ 708,260,086	1.4%	$ 703,043,314	-0.7%
PRIORITY	$ 2,149,756,513	$ 1,965,985,547	$ 1,574,135,765	$ 304,603,261	$ 1,878,739,026	-4.4%	$ 1,812,303,071	-3.5%
STANDARD	$ 15,192,621,161	$ 16,639,876,946	$ 14,493,227,605	$ 2,967,198,394	$ 17,460,425,999	4.9%	$ 17,603,043,920	0.8%
CBCIS TOTAL	**$ 47,947,818,146**	**$ 49,928,453,764**	**$ 41,951,842,248**	**$ 8,306,277,763**	**$ 50,258,120,011**	**0.7%**	**$ 50,340,858,099**	**0.2%**

**YTD is FY04 through Month 10*

Data source for Revenue Forecasting model: USPS – CBCIS database

Table 9
Area Forecast of Commercial Revenue

AREA LOC VIEW
- Prepared August 18, 2004

	Actual				Forecast					
AREA	**FY03 YE**	**Mkt Share**	**FY04 YTD**	**Mkt Share**	**FY04 YE**	**Mkt Share**	**Growth from FY03**	**FY05 YE**	**Mkt Share**	**Growth from FY04**
Capital Metro	$ 2,413,313,141	4.8%	$ 2,022,745,140	4.8%	$ 2,416,396,136	4.8%	0.1%	$ 2,451,524,918	4.8%	1.5%
Eastern	$ 9,000,908,326	17.8%	$ 7,620,696,066	17.9%	$ 9,114,546,201	17.9%	1.3%	$ 9,248,849,641	18.2%	1.5%
Great Lakes	$ 8,366,489,811	16.5%	$ 7,101,516,845	16.7%	$ 8,583,003,036	16.9%	2.6%	$ 8,778,262,733	17.2%	2.3%
Northeast	$ 3,817,039,312	7.5%	$ 3,192,011,622	7.5%	$ 3,826,768,314	7.5%	0.3%	$ 3,890,781,570	7.6%	1.7%
NY Metro	$ 4,552,860,843	9.0%	$ 3,800,607,676	8.9%	$ 4,494,589,939	8.8%	-1.3%	$ 4,538,790,300	8.9%	1.0%
Pacific	$ 6,004,216,718	11.8%	$ 4,974,615,844	11.7%	$ 5,944,771,020	11.7%	-1.0%	$ 5,624,943,104	11.1%	-5.4%
Southeast	$ 5,470,165,022	10.8%	$ 4,555,849,513	10.7%	$ 5,435,354,079	10.7%	-0.6%	$ 5,023,005,775	9.9%	-7.6%
Southwest	$ 3,931,742,565	7.8%	$ 3,231,777,460	7.6%	$ 3,878,481,054	7.6%	-1.4%	$ 4,025,356,716	7.9%	3.8%
Western	$ 7,128,100,644	14.1%	$ 5,981,223,834	14.1%	$ 7,176,876,073	14.1%	0.7%	$ 7,310,742,601	14.4%	1.9%
Grand Total	**$ 50,684,836,381**	**100.0%**	**$ 42,481,044,000**	**100.0%**	**$ 50,870,785,854**	**100.0%**	**0.4%**	**$ 50,892,257,357**	**100.0%**	**0.0%**

**YTD is FY04 through Month 10*
Note: Does not include Retail revenues

Data source for Revenue Forecasting model: USPS – CBCIS database

A Comparison of USDA's Agricultural Export Forecasts with ARIMA-based Forecasts

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

This study indicates that a number of USDA forecasts lack information that is readily available from monthly U.S. export data. This is determined by comparing the accuracy USDA's FY2001-04 forecasts with forecasts based on trends for each commodity. ARIMA models utilizing the monthly data available at the time each USDA forecast was published were estimated. Out of 24 separate commodity forecasts examined, USDA forecasts were superior to ARIMA forecast in only 9 cases. ARIMA forecasts were superior in 11 cases, and there was no difference in 4 cases.

Introduction

U.S. agricultural export forecasts are one subset of the voluminous information USDA provides on agriculture. From one perspective, these forecasts are but one small facet of a broad, integrated program of analysis. In addition to indicating developments in U.S. exports, published trade forecasts serve as useful indicators of USDA's perspective on current developments in global commodity markets. Similarly, the process of developing these forecasts may have positive externalities for other USDA priorities, both analytical and with respect to policy.

From another perspective, forecasting U.S. trade might be considered a diversion of resources that USDA could apply directly to other priorities. Published trade forecasts are only useful if they contain information not already published elsewhere. If USDA's published forecasts are no more accurate than forecasts anyone could develop from already published data, then USDA could increase public welfare by focusing on other priorities.

A balanced view of USDA's trade forecasting may lie between these two perspectives. Only a small percentage of the resources USDA devotes to U.S. export forecasting are devoted exclusively to this process, so the gains from eliminating the task may be small. Also, knowledge about trade clearly strengthens USDA's efforts on other commodity topics. However, if USDA's forecasts are no more accurate than those of an easily updatable model based on trends in monthly data, then rationality suggests there may be circumstances when using the model is preferred to USDA's more extensive efforts. At the very least, it suggests that the output of the model should be added to the information set available to USDA forecasters.

Methods and Data

This study compares the accuracy of USDA's fiscal year export value forecasts for FY 2001-04 with forecasts based on trends in each commodity's monthly exports. USDA's forecasts are published quarterly in the *Outlook for U.S. Agricultural Trade*, The trend forecasts were produced with ARIMA models utilizing the monthly data available at the time each USDA forecast was published. The models were specified and estimated with the Tramo/Seats software developed by the Bank of Spain. This software was incorporated by Eurostat into a software package, Demetra, which was this study's interface for Tramo/Seats.

For a given fiscal year (October-September), USDA forecasts U.S. agricultural export value by commodity 5 times. The first forecast is published in August, before the fiscal year begins, and updates are published in November, February, May and the following August. As an illustration, Table 1 compares the ARIMA forecasts of U.S. cotton export value with those USDA published each November during FY 2001-04.

International agricultural trade is in a constant state of flux. Economic development around the world has induced significant structural change for consumers, producers, and traders. For economists, structural change necessitates newly specified or estimated models. For forecasters, structural change means reorienting toward new countries, to different segments of the supply chain, or to different portions of the marketing year. This study's efforts were confined to the last 4 years to limit the impact of inevitable changes in world markets on the validity of its conclusions.

Historical data for the ARIMA modeling was downloaded from the Foreign Agricultural Service's website. In order to ensure a sufficient number of observations for ARIMA modeling, June 2000 was the end point of the oldest data set, and August 2000 was the date of the earliest USDA forecast analyzed.

Results

Table 1 shows that USDA's cotton forecast error each November during FY2001-04 was smaller than what an ARIMA-based forecast would have produced. Note that the software chooses 3 different ARIMA specifications over the 4 years studied, perhaps indicative of market volatility that hinders the accuracy of ARIMA forecasting. The ARIMA model's error in FY 2002 is extraordinarily large, and the ARIMA methodology's root mean squared error (RMSE) during these 4 years is substantially higher than USDA's as a result. The ratio between the RMSE of the ARIMA methodology and the RMSE of USDA's forecasts is 3.2 / 0.2 = 13.7. This was the highest ratio for any commodity for any of the forecast update months (Table 2). In Table 2, ARIMA modeling is less accurate than USDA if the ratio is greater than 1. This ratio provides a simple indication of relative performance.

ARIMA forecasts of aggregate commodities, like Grains and feeds, are the sum of forecasts by ARIMA models for each component of the aggregate. This includes a forecast of residuals for aggregate groupings. USDA's published forecast for Total U.S. agricultural exports is essentially a sum of its published forecasts of each component of agricultural trade. However, USDA's forecast of Grains and feeds exports (for example) is larger than its published forecasts of specific categories of grains and feeds. Therefore, there is an implied forecast of the remaining products. Table 2 indicates that USDA's RMSE has been at least twice as large as the error that ARIMA-based forecasting would have realized for this grains and feeds residual.

While cotton and the residual category for grains and feeds have relative RMSEs that clearly indicate the dominance of ARIMA or USDA forecasting, these are atypical. For most forecasts, the ratio is much closer to 1.0. To formally compare the accuracy of the methodologies, two statistics were calculated. A general measure of forecast accuracy has been developed by Diebold and Marino, the Morgan-Granger-Newbold test (Table 3). A sign test was also used to determine if the frequency of a given forecast's greater relative accuracy was significant during 2001-04 (Table 4).

While some patterns are apparent in Tables 3 and 4, one further step was taken to make these patterns clearer. A summary statistic was created for each commodity. To create this statistic, USDA's forecast of each commodity was assigned a score based on its performance in the MGN and sign tests. For each update month for which USDA's 2001-04 performance was significantly better (at least 10 percent significance) than the ARIMA-methodology's for a given test, a score of 1 was assigned. If the significance of USDA's dominance was 1 percent or better, a score of 2 was assigned. On the other hand, when ARIMA forecasting dominated, the scores were -1 and -2, respectively. A commodity's composite score is the sum of its MGN and sign test scores over the 5 update months. Conceivably, a commodity's score could be as high as 20 or as low as -20.

Table 5 ranks the scores in ascending order, and negative scores are more common than positive scores. The scores range from 8 to -9. Soybeans, soybean meal, and cotton have the best scores, while a number of high-value products and rice have the worst. Interestingly, the accuracy of USDA's estimates for Horticultural products in total is lower than for virtually all the forecasts of the components of the total. This isn't the case for any of the other aggregates: Grains and feeds, Oilseeds and products, and Livestock and products.

These results have an important implication for USDA's forecasting of U.S. total agricultural export value. A combination of USDA and ARIMA forecasts is more accurate than either alone. A forecast of total U.S. agricultural export value can be created by using ARIMA forecasts for all commodities with a composite score of -3 or below and USDA forecasts for all other commodities. This combined forecast would have been more accurate than USDA's forecast in 3 out of the last 4 years (except for the initial August release, for which there was a tie in one year). This frequency of dominance is not statistically significant. However, the MGN test indicates the combined forecast was significantly more accurate (at the 1 percent or better level) in November and February over 2001-04. This comparison does not take into account USDA's practice of revising commodity forecasts to create a total export value forecast that rounds to the nearest $500 million, but perhaps the rationale and implications of that practice bear examination.

Conclusions

Benchmarked against ARIMA-based forecasting, USDA's quarterly U.S. export value forecasts are more often dominated than dominant. As measured by a composite score, USDA forecasts were superior to ARIMA forecast in only 9 out of the 24 separate commodities examined. ARIMA forecasts were superior in 11 cases, and there was no difference in 4 cases. While it should be noted that ARIMA forecasts did not dominate a majority of the commodities, the onus is probably on USDA to dominate the ARIMA forecasts, which it has often failed to do.

USDA does not devote equivalent resources to each commodity's forecast. Some receive a great deal of attention, some very little. All USDA forecasts are approved by the World Agricultural Outlook Board (WAOB). Examination of the WAOB's publication, the *World Agricultural Supply and Demand Estimates* (*WASDE*), reveals different levels of detail for different commodities. Variations in the levels of detail correspond to variations in the intensity of USDA's forecasting efforts. Variations in USDA's accuracy also correspond to these variations in forecasting efforts.

Complete supply and demand estimates for the United States and other major producers, consumers, importers, and exporters comprise the greatest level of detail any commodity receives in the *WASDE*. The next level of detail is to provide supply and demand forecasts for only the United States. In each case, these forecasts are produced by interagency committees that meet monthly, reviewing developments in U.S. and world markets (see Vogel and Bange for discussion).

The commodities with the greatest level of detail in the *WASDE* can be grouped into a "high attention" category :

> Wheat, corn, rice, soybeans, soybean meal, soybean oil, and cotton.

The commodities with only U.S. supply and demand tables in the *WASDE* can be grouped into a "medium attention" category:

> Sugar, beef, pork, broilers, turkeys, eggs, and milk.

However, the majority of the 24 commodities examined in this study are not included in the *WASDE*. The interagency committees overseeing these forecasts meet less frequently, and the supply and demand estimates USDA provides for these commodities include only a small number of countries outside the United States (Table 6). These other commodities can be grouped into a "low attention" category.

Averaging the composite scores of commodities in the "high attention" category (8 forecasts) gives an average of 1.9, indicating the superiority of the USDA forecasts. The average of commodities in the "medium attention" category (3 forecasts) is -1.0, and the "low attention" category's average is -1.3 (13 forecasts).

The implication is that, for the majority of commodities included in USDA's quarterly *Outlook for U.S. Agricultural Trade*, publication of ARIMA-based forecasts of U.S. exports would be an improvement from previous efforts in terms of accuracy. These forecasts are primarily in the "low attention" category.

At the very least, the information embodied in ARIMA-based forecasts would make a useful contribution to USDA's analysis of these commodities. This also holds for some of the commodities already receiving a significant amount of forecasting resources. A forecast can only be considered rational if it embodies all information available when the forecast is developed. Software is now available that can readily provide this information, offering a viable opportunity to improve USDA's accuracy.

This study indicates that a number of USDA forecasts lack information that is readily available from monthly U.S. export data. The appropriate response to this challenge would vary by commodity and would be best implemented by specialists concentrating on these commodities. The "high attention" commodities, on average, have forecasts superior to ARIMA-based forecasts, appropriately enough. The advantages of adding such trend analysis to forecasters' information set are not immediately obvious. However, as circumstances change, it is appropriate to consider all options as any forecasting institution reviews its changing mix of priorities and resources.

References

Diebold, F.X, and R.S. Mariano (1995), Comparing predictive accuracy, *Journal of Business and Economic Statistics*, 13(3): 253-263.

DEMETRA 2.0, Service Pack 1 - Release (May 2002) http://forum.europa.eu.int/irc/dsis/eurosam/info/data/demetra.htm

U.S. Department of Agriculture, *Outlook for U.S. Agricultural Trade*, various issues 2000-2004.

U.S. Department of Agriculture, *World Agricultural Supply and Demand Estimates*, various issues.

Vogel, Frederic, and Gerald Bange (1999). *Understanding USDA Crop Forecasts*, Office of the Chief Economist, U.S. Department of Agriculture. Miscellaneous Publication No. 1554.

Table 1--November U.S. cotton export value forecasts, FY 2001-04

Fiscal year	ARIMA model historical data	ARIMA model specification	ARIMA forecast	USDA forecast	ARIMA error	USDA error
			--------------Billion dollars--------------			
2001	Oct 95 - Sep 2000	(0 1 0) (0 1 1)	2.8	2.4	0.7	0.3
2002	Oct 95 - Sep 2001	(0 2 0) (0 1 1)	8.3	2.1	6.3	0.1
2003	Oct 95 - Sep 2002	(1 0 0) (0 1 1)	2.0	2.6	-0.9	-0.3
2004	Oct 95 - Sep 2003	(0 1 0) (0 1 1)	5.3	4.3	0.8	-0.2

Table 2--Ratio of ARIMA RMSE to USDA RMSE for updates of fiscal year forecasts, 2001-4

	August	November	February	May	August
Grains and feeds	1.5	0.4	1.3	0.9	4.5
Wheat and flour	2.6	1.3	2.8	1.3	2.4
Rice	0.7	0.6	0.7	2.1	1.0
Coarse grains	0.6	0.3	0.7	0.3	1.1
Corn	0.6	0.3	1.0	0.1	1.6
(unpublished residual)	0.5	0.2	0.0	0.8	1.4
Feeds and fodders	1.0	0.9	0.2	0.4	0.6
(unpublished residual)	0.5	0.5	0.2	0.3	0.5
Oilseeds and products	1.3	1.9	1.7	5.4	4.6
Soybeans	1.4	2.0	2.1	5.6	6.4
Soybean meal	1.0	2.7	1.5	1.7	3.6
Soybean oil	1.9	2.0	2.8	1.0	0.9
(unpublished residual)	0.9	1.1	0.7	0.2	0.6
Livestock products	1.0	0.3	1.8	1.2	0.6
Beef, pork, and variety meats	1.1	0.3	2.9	1.8	0.6
Hides and skins	0.8	0.6	0.3	0.6	1.0
(unpublished residual)	0.5	1.3	0.7	0.4	0.7
Poultry and products	1.0	1.2	1.5	1.2	0.1
Broiler meat	--	--	--	--	--
Dairy products	0.9	0.6	0.9	1.1	0.7
Tobacco, unmanufactured	0.8	1.0	0.6	1.6	2.1
Cotton and linters	1.1	13.7	2.5	1.7	2.0
Seeds	1.2	0.4	0.6	0.6	0.8
Horticultural products	0.7	1.1	0.6	1.0	1.1
Fruits and preparations	1.0	0.8	1.0	1.1	1.0
Vegetables and preparations	1.9	1.7	1.3	0.7	1.2
Tree nuts and preparations	0.8	2.3	1.0	2.7	1.3
(unpublished residual)	0.6	0.6	0.6	0.7	3.3
Sugar and tropical products	1.2	1.1	0.9	0.3	1.8
Total	1.4	2.5	1.0	3.1	1.7

Note: "--" means forecast was not analyzed.

Table 3--Forecast with significantly greater accuracy (Granger-Newbold-Morgan test), 2001-04

	August	November	February	May	August
Grains and feeds		**ARIMA**[1,2]			
Rice	**ARIMA**	ARIMA	**ARIMA**	**USDA**	**ARIMA**
Coarse grains			ARIMA		USDA
Feeds and fodders	ARIMA	**ARIMA**	**ARIMA**	**ARIMA**	
Oilseeds and products				USDA	USDA
Soybeans			**USDA**	**USDA**	USDA
Soybean meal	**USDA**	**USDA**		**USDA**	**USDA**
Soybean oil			USDA	ARIMA	ARIMA
(unpublished residual)					**ARIMA**
Livestock products			USDA		
Beef, pork, and variety meats	**USDA**		**USDA**		ARIMA
Hides and skins	**ARIMA**	ARIMA			
Poultry and products	USDA	USDA	**ARIMA**		
Dairy products		**ARIMA**	**ARIMA**		
Tobacco, unmanufactured				USDA	**USDA**
Cotton and linters	**USDA**	**USDA**			USDA
Seeds	**ARIMA**	**ARIMA**	**ARIMA**	**ARIMA**	ARIMA
Horticultural products			**ARIMA**		
Vegetables and preparations			USDA		
Tree nuts and preparations			**USDA**	**USDA**	
(unpublished residual)			**ARIMA**	**ARIMA**	**USDA**
Sugar and tropical products	**USDA**	**USDA**			

Note: Commodities with neither methodology dominant in any month not included.
[1]Labels indicate a methodology's dominance with at least 10 percent significance
[2]Bold labels indicate dominance with at least 1 percent significance

Table 4--Significant number of years with superior accuracy, 2001-04

	August	November	February	May	August
Grains and feeds	ARIMA	ARIMA			USDA
Wheat and flour					USDA
Coarse grains	ARIMA	ARIMA			
Corn			ARIMA		
(unpublished residual)	ARIMA	ARIMA	ARIMA		
Feeds and fodders			ARIMA		
Oilseeds and products				USDA	USDA
Soybeans			USDA	USDA	USDA
Tobacco, unmanufactured					USDA
Cotton and linters	USDA	USDA			
Horticultural products			ARIMA		
Total				USDA	

Note: Commodities with neither methodology dominant in any month not included.
[1]Labels indicate a methodology's dominance with at least 10 percent significance

Table 5--Composite score of USDA relative forecast accuracy, 2001-04

Commodity	Score
Seeds	-9
Feeds and fodders	-8
Dairy products	-6
Rice	-5
Horticultural products	-4
(horticultural residual)	-4
Grains and feeds	-3
(coarse grains residual)	-3
Hides and skins	-3
(oilseeds residual)	-2
Tree nuts & prep.	-1
Coarse grains	-1
Corn	-1
Soybean oil	-1
(grains residual)	-1
(livestock residual)	0
Poultry and products	0
Fruits & prep.	0
Wheat and flour	1
Livestock products	1
Vegetables & prep.	1
Total	1
Beef, pork, etc	3
Oilseeds and products	4
Tobacco, unmanuf.	4
Sugar & tropical prod.	4
Cotton and linters	7
Soybeans	8
Soybean meal	8
Broiler meat	#N/A

Table 6--USDA's commodity supply and demand forecasts

	Updates each year[1]	Countries analyzed[1]
Cotton, grains, & oilseeds	12	101
Tobacco	11	123
Livestock	2	21
Dairy	2	15
Sugar & Tropical	2	105
Horticultural	2	8

Source: USDA PS&D Online Database

[1]Average of products in commodity group

ADDITIONAL DOCUMENT(S)

Exchange Rates, U.S. Agricultural Exports, and FDI

Christine Bolling
Economic Research Service, USDA

U.S. agricultural trade is an important source of earnings for U.S. agriculture and agribusiness. Additional sources of foreign earnings to U.S. agribusiness are from direct investments abroad, but they are seldom discussed.

The dollar has weakened since 2000, creating a changed environment for U.S. agricultural trade and foreign direct investment (FDI) compared to the macroeconomic environment of a strong dollar of much of the 1990s. Exchange rate fluctuations are only one of the major macroeconomic factors affecting trade and FDI, with income growth in other countries perhaps being the most important. The purpose of this study is to evaluate the effects of the weakened dollar and projected income growth on U.S. agricultural trade and FDI in the near future, drawing on recent studies carried out in ERS.

This paper highlights recent trends in U.S. agricultural trade and FDI in processed food, the near term outlook for income growth in other countries and outlook for exchange rate fluctuations. This paper draws on the ERS Baseline exercise and ERS trade outlook work (Westcott; Brooks, et. al.) to highlight the effect of the recent dollar devaluation on U.S. agricultural trade.

The principal addition of this research pertains to the effects of income growth and exchange rate fluctuations on outward FDI. This paper revisits the Somwaru-Bolling model (updated in 2004) to ascertain future prospects for growth in U.S. FDI abroad.

The Importance of U.S. Agricultural Trade

A positive agricultural trade balance has been an important part of the U.S. agricultural policy environment during the past 5 decades since export earnings have been an important factor in determining farm earnings (fig. 1). We export 20 percent of our agricultural production, including large portions of wheat, corn, and soybeans, bolstering market prices for these important U.S. farm products. Canada, Mexico, Japan, China, and South Korea are our most important export markets. Nine countries are export markets of more than a billion dollars, with Canada as a $9.6 billion market in 2004 (fig. 2). In all, the United States had a record $61 billion in agricultural export earnings in 2004.

The Importance of FDI in the Processed Food Industry

While U.S. export earnings fluctuated in recent years, food industry sales from the $30 billion investments of U.S. FDI abroad continued to grow to reach well over $150 billion in 2002, representing additional earnings to the U.S. agribusiness community (figs.3, 4). Food and beverage sales abroad from U.S. FDI included $17.2 billion to Canada, $17.1 billion to Mexico, and $16.3 billion to the UK. There were 18 countries where sales from FDI abroad were over a billion dollars. While U.S. agricultural exports to the EU-15 totaled $6.1 billion in 2002, U.S. FDI sales to those same countries totaled nearly $86 billion in the same year. While quoting these sales, it is important to note that the value added component from trade since the labor component from FDI and taxes paid to foreign governments actually stay abroad. Sales from outward FDI in the global processed food and beverage industries also exceed sales from inward FDI into U.S. processed food and beverage industries.

Trends in World Economic Growth and the Dollar Exchange Rate

According to Oxford Economic Forecasting, prospects for economic growth are uneven but strong for the year 2005, following the slower growth rates of the new millennium (fig. 5). Much of the strong economic growth is occurring in China, India, Argentina, Brazil, and Mexico, leaving the EU-15 with a less than 2 percent rate of economic growth in 2005. The United States is forecast to have nearly a 3.8 percent growth rate, despite the drag that high petroleum prices are putting on the economy.

Meanwhile, since 2000, the U.S. dollar has slid in relation to the major currencies, including the Japanese *yen* and the European Union e*uro*. In nominal terms, in an FDI-weighted exchange rate index, the dollar had declined nearly 15 percent, but nearly all the decline had been with the currencies of the developed countries, where it had declined nearly 30 percent from 2000 to January 2005. Many of the currencies of the developing world are pegged to the dollar, but in late 2004, some floating currencies such as the Brazilian *real* and Argentine *peso* had also appreciated with respect to the dollar. (An agricultural- trade weighted exchange rate index is available on the ERS web site www.ers.usda.gov/data/exchangerates/

The Weakened Dollar and Agricultural Exports

Econometric studies have already established that agricultural exports have a significant exchange rate elasticity (Batten and Belongia; Chambers and Just; and Cushman). The current ERS estimate is - 0.79, as incorporated in the USDA Baseline (Westcott). A weakened dollar with respect to currencies of developed countries is expected to cause a turnabout in U.S. agricultural exports, perhaps leading to new records in nominally valued export. The main obstacle to growth is high tariffs and non-tariff barriers such as food safety measures. Income growth in major importing countries is also cited in many studies as a determinant of agricultural exports. Much of the near term growth is expected to occur in middle income countries, where income elasticities for food are higher than in developed countries. China and India, starting from a lower income base and large global population centers, are also expected to be growing markets. Currencies of developing countries have not appreciated as much as developed countries, and they will be less of a driving force than income growth.

The Weakened Dollar and U.S. FDI Abroad

It has been hypothesized that the exchange rate effect on FDI was a counterpoint to trade, where a strong dollar led to an increase in FDI and sales while weakening prospects for U.S. agricultural exports. Conversely, the strong dollar led to increased agricultural imports and a decline in inward FDI. When the dollar appreciates, U.S. companies seek investments abroad because assets, labor costs, and raw material costs are less. Econometric studies in ERS also lead to that conclusion, although the relationship is stronger for developing countries than it is for developed countries. An example is the updated Somwaru-Bolling study that was based on the Gopinath-Pick-Vasavada methodology but covered additional countries. The original Somwaru-Bolling model is a four-panel equation system with foreign affiliate sales from FDI, U.S. processed food exports, foreign affiliate demand for labor and demand for foreign direct investment capital. Data for the studies was obtained from IMF *International Financial Statistics*, World Bank *World Development Indicators*, Bureau of Economic Analysis, and Economic Research Service FATUS. Price and income data was adjusted for inflation using a GDP deflator.

Figure 7 (empirical results from the updated FDI sales panel of the Bolling-Somwaru study) demonstrates the strong positive relationship between the exchange rate and FDI sales in the host country. It is interpreted as follows: as the dollar appreciates, it takes more of the host country currency for each dollar, so it is implied that as the dollar increases in value in relation to the host country (in real terms) sales from FDI sales increase. This is consistent with the Gopinath-Pick-Vasavada study, which covered the OECD countries only. Equally important is the generally positive relationship between U.S. FDI sales and income growth in the host countries. This relationship would be consistent with demand for processed food, whether it comes from domestic production, FDI, or imports. It would be expected that there would be a significant positive relationship between FDI sales and income growth.

From this empirical study, the signs of the coefficients would lead to the conclusion that a strong dollar leads to increased FDI sales, and conversely a weaker dollar would lead to a decline in U.S. outward FDI sales. The effect would be expected to be strong in developed countries as a group in comparison to developing countries since the dollar has not depreciated as much against the developing-country currencies as it has against the developed country currencies. Income growth in these countries *could* mitigate the effect of the currency fluctuations.

In conclusion, for the U.S. processed food industry, there may be a slowdown in real U.S. FDI sales abroad in the near future, but the effect will be different between developed and developing countries. The dollar devaluation is much deeper and occurred much earlier in developed countries. Moreover, income growth is expected to be slower in the European Union and Japan exacerbating the exchange rate effect. But, the effect of the slowdown in global FDI sales brought on by the depreciating dollar to developed country currencies may also be less severe because of mitigating circumstances. This is just because of the differences in the extent of the dollar devaluation between the two country groups. Because the dollar devaluation was also much less in developing countries, U.S. FDI faces more positive prospects in the developing world. Income growth is also going to be much sharper in the developing countries. In contrast, U.S. agricultural exports are expected to increase as a result of the weakened dollar.

References

Batten, D.S., and M.T. Belongia. “Monetary Policy, Real Exchange Rates, and U.S. Agricultural Exports,”American Journal of Agricultural Economics, Vol. 68, pp. 422-27, May 1986.

Brooks, Nora, Carol Whitton, and Ernest Carter, coordinators. Outlook for U.S. Agricultural Trade, AES-45. http://www.ers.usda.gov/publications/so/view.asp?f=trade/aes-bb/

Chambers, R.G., and R.E. Just. "Effects of Exchange Rate Changes on U.S. Agriculture: A Dynamic Analysis," American Journal of Agricultural Economics, Vol. 63, pp. 32-46, Feb. 1981.

Cushman, D.O. "The Effects of Exchange Rate Risk on International Trade," Journal of International Economics, Vol. 15, pp. 45-63, 1983.

Gopinath, Musimany, Daniel Pick, and Uptal Vasavada, "Exchange Rate Effects on the Relationship between FDI and Trade in the U.S. Food Processing Industry," American Journal of Agricultural Economics, Vol. 80, pp. 1073-79, Dec. 1998.

International Monetary Fund, International Financial Statistics, Washington, DC, selected issues.

Makki, Shiva S., Agapi Somwaru, and Christine Bolling. "Determinants of Foreign Direct Investment in the Food-Processing Industry: A Comparative Analysis of Developed and Developing Countries," Journal of Food Distribution, Vol. 35 No. 3, Nov. 2004.

Oxford Economic Forecasting. World Economic Prospects Monthly Review: April 2005.

Somwaru, Agapi, and Christine Bolling. "U.S. Foreign Direct Investment and Trade: Substitutes or Complements? The Case of the Food Processing Industry." Selected paper, AAEA Annual Meeting, Nashville, TN, Aug. 8, 1999.

U.S. Department of Agriculture, Economic Research Service (USDA-ERS). 1989-2000. Foreign Agricultural Trade of the United States, Washington, DC

U.S. Department of Commerce, Bureau of Economic Analysis. Survey of Current Business, selected issues.

U.S. Department of Commerce, Bureau of Economic Analysis. U.S. Direct Investment Abroad: Operations of Parent Companies and Their Foreign Affiliates. Washington, DC, selected issues.

Westcott, Paul, ERS Contact. USDA Agricultural Baseline Projections to 2014., OCE-2005-1, http://www.ers.usda.gov/Publications/oce051/

World Bank, 2004, *World Development Indicators.* The World Bank Group, Washington, DC

FEDERAL
FORECASTERS
CONFERENCE