The Sixth Annual

Federal Forecasters Conference - 1993

Papers and Proceedings

Crystal Gateway Marriott Hotel
Crystal City, Virginia
September 8, 1993

- Sponsoring Agencies -

Bureau of the Census ♦ Bureau of Economic Analysis ♦ Bureau of Mines
Bureau of Labor Statistics ♦ Central Intelligence Agency ♦ Department of Veterans Affairs
Economic Research Service ♦ Energy Information Administration ♦ Environmental Protection Agency
National Center for Education Statistics ♦ U.S. Geological Survey

THE SIXTH ANNUAL FEDERAL FORECASTERS CONFERENCE

PAPERS AND PROCEEDINGS

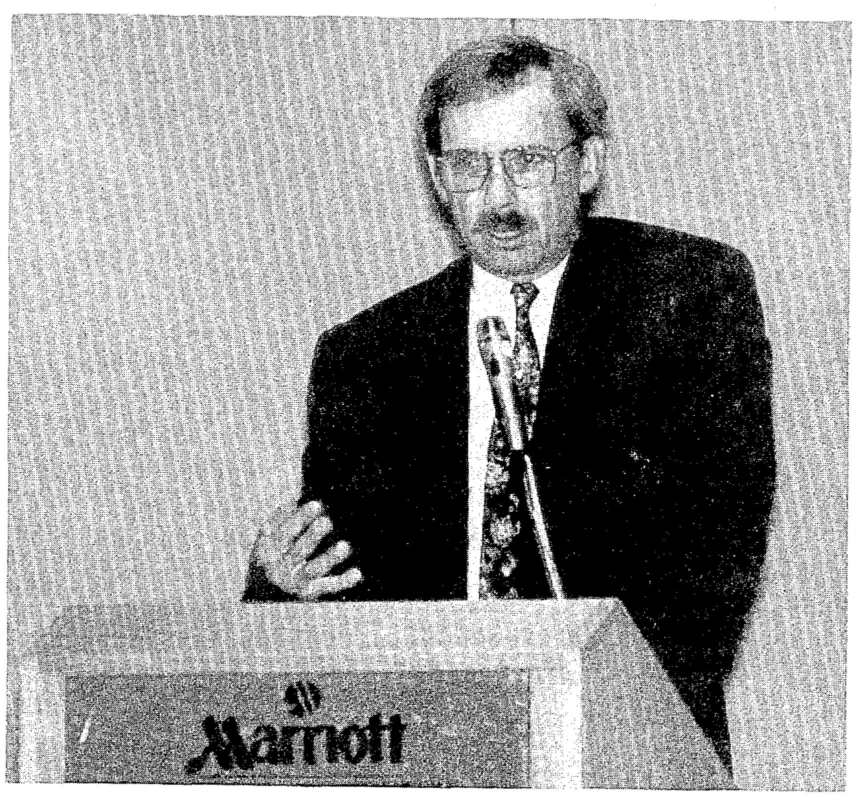
The Crystal Gateway Mariott Hotel
Crystal City, Virginia
September 8, 1993

Released for Printing: October 4, 1994

Printed with soy ink on recycled paper

SCENES FROM THE CONFERENCE

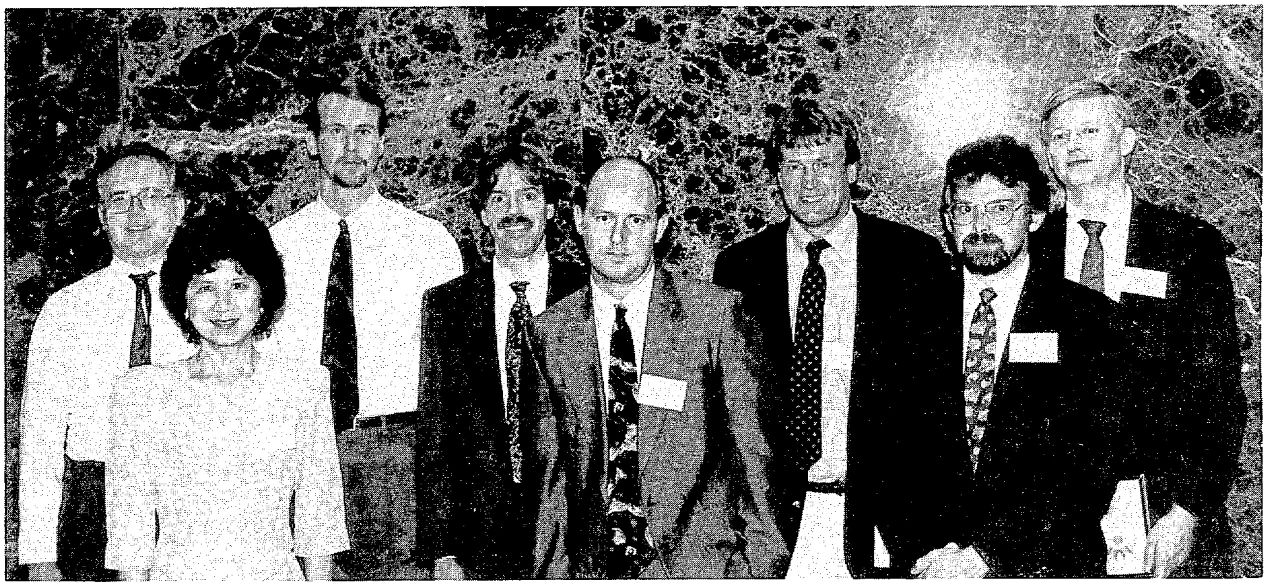
The 1993 Federal Forecasters Conference Organizing Committee



Larry Pettis, then Acting Administrator of the Energy Information Administration, delivers the opening remarks at the Conference.

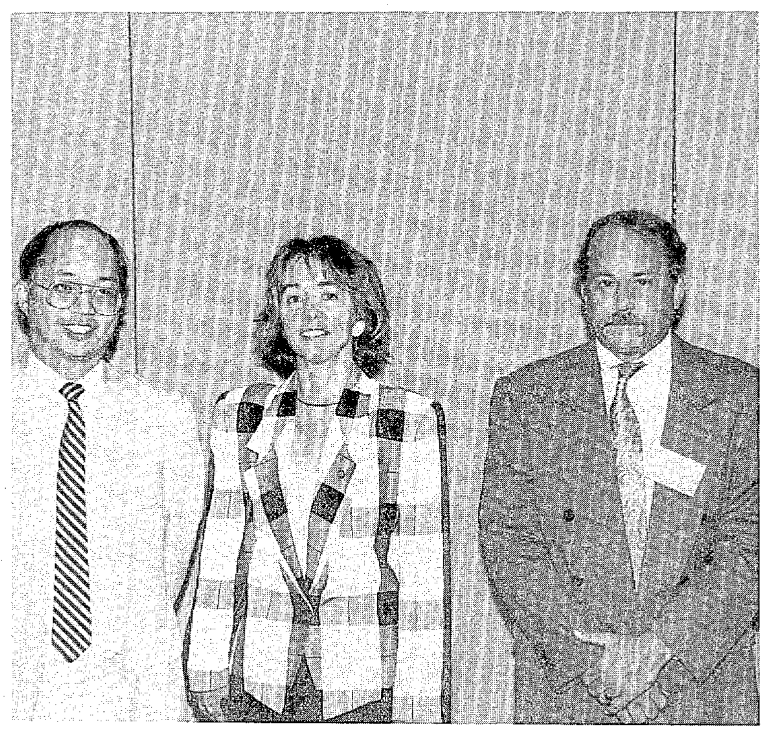
TABLE OF CONTENTS

SCENES FROM THE CONFERENCE



Award winners from FFC-93.

Captain Dennis Egan of the National Performance Review delivers the keynote address.



The panel discussants on the "Role of Forecasting in the Anticipatory Government": Michael Woo, Professional Staff Member for the Committee on Energy and Commerce; Heather Ross, Special Assistant to the President for Economic Policy at the National Economic Council; and Mark Rodekohr, Director of the Energy Demand and Integration Division of the Energy Information Administration.

MORNING SESSION--FFC/93

Welcome and Opening

Mr. Earley: I want to welcome you to the Sixth Annual Federal Forecaster's Conference. I am Ron Earley and I am the Chair of this year's conference. This Conference is a collaborative effort between 11 different sponsoring agencies and over 50 attending agencies. This year we have approximately 300 registrants. We're holding true to form in terms of the growth and expansion of the conference. We have moved the site to a different facility this year. I would like to get some comments from you on this move.

I do want to say that it's a real pleasure to welcome you here today, and I think we have an interesting program. We've expanded participation this year to broaden it beyond the typical Federal base that we have. Represented here today are a number of private organizations, State Governments, as well as academia.

I particularly want to welcome a number of foreign visitors that we have today. A gentleman from BLS called me and said he had a training program and wanted to bring some of his trainees to the conference. So we have a large contingent of international representatives today. It was a little bit tough squeezing some of their names on their name tags. We had a little difficulty with spelling too, but I think we worked that one out.

I was talking to Debbie Gerald and she mentioned a story about a coworker who attended one of the town hall meetings at the Department of Education. At that town hall meeting, Vice President Gore asked for examples of things in Washington where Government employees developed and implemented programs. One employee spoke up and gave as an example the Federal Forecaster's Conference. I truly think that it is a good example of a grassroots activity built around 11 sponsoring agencies and all of the attending agencies.

I would like to pay particular thanks to those people who have served with me on the organizing committee this year. From the Bureau of Mines, Sandra Absalom and Ching Yu. From the Bureau of Economic Analysis, Tom Lienesch and Zoe Ambargis. From the Bureau of the Census, Paul Campbell. From the Bureau of Labor Statistics, Howard Fullerton. From the National Center for Education Statistics, Debra Gerald. From the Economic Research Service, Steven MacDonald and Karen Hamrick. From the Environmental Protection Agency, Joe Abe and Dave Rejeski. U.S. Geological Service, Ethan Smith. Department of Veterans Affairs, Peg Young, and finally, the Methodology Center of the Central Intelligence Agency as always, no named individuals. And of course, my colleagues at the Energy Information Administration, Arthur Andersen, Louise Bonadies, and Doug Hale. I also want to thank Kay, Brian, and Jason for all of their efforts while I've been preoccupied. A special "Thanks" is reserved for Sharon Shears, Shirley James, Linda Walsh, John Weiner, and Angela Renfrow for their diligent work on the logistics for this conference.

I would also like to pay particular thanks to the past chairs of this conference, Tom Lienesch from the Bureau of Economic Analysis, Norm Saunders and Howard Fullerton from Bureau of Labor Statistics, Karen Hamrick from the Economic Research Service, and Debra Gerald from the National Center for Education Statistics.

I want to thank you for your support and guidance during this process.

Last year's Chair, Tom Lienesch, operated under a little bit of a luxury in that he knew who the chairman for the next year's conference would be. I am not yet as honored as Tom was last year at this time. So what I will leave you with is the idea that the associations that I've had over the last six months in planning and working on this conference will be long remembered and pay benefits to me both personally and hopefully professionally. Thanks for the opportunity.

Now, to welcome us to this year's conference, I'd like to introduce you to Larry Pettis, who since January has served as the Acting Administrator of the Energy Information Administration and for many years now has been the Deputy Administrator of EIA. In these roles, he guides a very broad program of data collection, forecasting, and policy analysis. So to kick off the conference, I'd like you to please welcome Larry Pettis.

Mr. Pettis: Good morning. As I was thinking this weekend about meeting with a group of distinguished forecasters, I was reminded of an old saying which is "the most important thing is to know what you don't know."

Forecasting is a very humbling experience and we very quickly learn how much we don't know about the future, and I think it is in part that recognition that brings us all here today together. The Federal Forecasters Conference provides an excellent forum for us to develop professional relationships with other members of the forecasting community and to gain new ideas on how we can improve our forecasting activities.

The theme of this year's conference is the role of forecasting in Anticipatory Government. In their book, *Reinventing Government*, David Osbourne and Ted Gabler devoted a chapter to "Anticipatory Government, Prevention Rather Than Cure." The premise of this chapter is based on two points. One is that old adage that "an ounce of prevention is worth a pound of cure," and the other is that foresight must be built into the decisionmaking process.

With respect to forecasting, I'd like to read a quote from the book. It probably sounds familiar to some things you've heard before.

> Skeptics often argue that it is impossible to predict with accuracy ten years ahead. They are right. But the point is not to make projections that are 100 percent accurate. The point is simply to flag problems that loom ahead.

Most of the successes that are mentioned in this book are at the state and local level, and the authors point out a number of obstacles to Anticipatory Government at the Federal level. Included among those are the one-year budget cycle, accrual accounting which tends to focus on near term rather than long-term outlays, and the short-term political focus in Washington.

Despite these obstacles, as Federal forecasters, I think we can point to a number of successes where our work has allowed decision makers to make informed policy decisions that do have long-term importance to our Nation. Thinking about my own agency, there are a number of instances where our forecasting or analytical work related to those forecasts have been used in developing regulatory or legislative initiatives. Some that immediately come to mind are the tax reform back in 1986, the Clean Air Act amendments, and natural gas deregulation.

Last year at this conference, my agency, the Energy Information Administration, made a presentation in one of the concurrent sessions on an initiative that we've had underway in our agency for the past couple of years: the development of a new National Energy Modeling System, or NEMS, as we fondly call it. I'm pleased to report that at this point in time we're in the final stages of testing on this new modeling system and it will be used in our flagship forecast this year, the Annual Energy Outlook which will be developed later this Fall. We believe that NEMS is going to help us improve our forecasting and more importantly, understand some of the underlying trends behind that forecast.

But the goals of NEMS go beyond forecasting. NEMS has been designed as a policy analysis tool and we're optimistic that this will improve our ability to inform decision makers about how policy changes may change the future.

In closing, I would just like to recognize two groups that I think are in fact under-appreciated and under-recognized. First are forecasters in general. So I would like to acknowledge the fine work that you do as Federal Forecasters that helps both us as agencies and the policy makers in general. The second group which Ron has already alluded to are those people that get the pleasure of planning conferences. And I would like to acknowledge Ron and the coordinating committee for the excellent job they've done in putting together a timely and interesting conference for us today.

So with that, I welcome you to the conference and I hope that you enjoy it as much as I think you will. Thank you.

Mr. Earley: Thank you, Larry. I would like to turn the program back over to some of my colleagues. The contest that we embark upon every year is one of the fun aspects of the conference. I would like to have Karen Hamrick and Debbie Gerald come up and tell us who are the best among us.

Award Presentations

Ms. Gerald: Good morning. Welcome to FFC/93 and to one of its most awaited events. Believe it or not, we receive many calls regarding this contest and several of you take it rather seriously and we're glad to hear that. This is the third consecutive year that Karen Hamrick and I have done the contest. This year's contest has five items to forecast for various times during the month of August. They are: the U.S. civilian unemployment rate, the average bank prime rate, the cash price for an ounce of gold, the temperature, and the number of home runs by the home run leader in the American League.

I will be announcing the runner-ups. Karen Hamrick will announce the winner. Before we continue, we would like to give a special award for the entry with the most direct hits. This special award goes to Allan Eck, Bureau of Labor Statistics. Allan is not here but we'll make sure he receives it.

Next, I will announce the winners who will receive a Certificate of Honorable Mention. Please come up and receive your certificate as I call your name.

From the Bureau of Labor Statistics, Betty Su. This is the second time that Betty Su has received the certificate.

The next award goes to Thomas A. Amirault, Bureau of Labor Statistics.

The next award goes to Joel Green, Economic Research Service.

David Kass, Bureau of Economic Analysis.

The next award goes to Thomas Snyder, National Center for Education Statistics. This is the third year in a row that Tom Snyder has received the certificate.

The next award goes to Kenneth Beckman, Bureau of Mines.

Next is William Hussar, National Center for Education Statistics.

Next is Clifford Woodruff, III, Bureau of Economic Analysis. And this is also the third year in a row that Clifford Woodruff has received the certificate.

Next is Ken Wetzel, Environmental Protection Agency.

And last is Brian Unruh, Energy Information Administration.

Karen Hamrick will announce the winner.

Ms. Hamrick: I know this is the moment everyone's been waiting for.

The winner of our Federal Forecaster's Conference this year -- is John Burgan, from the Bureau of Labor Statistics.

Thanks to everyone who entered and congratulations to all our winners.

Mr. Earley: I'd like to introduce Tom Lienesch. Tom was the chair for last year's Federal Forecaster's Conference, and Tom has the distinct honor of giving some awards today for the papers that were presented at last year's conference.

Mr. Lienesch: This is the last few loose ends to tie up from last year's conference and before I get to the awards for the best technical papers presented last year, I'd like to first thank all the participants who either gave papers, participated in panel discussions, or otherwise contributed to last year's conference. You all materially contributed to what I thought was a successful conference and thank you, very much. I totally appreciate it.

Now to the awards. Borrowing fairly liberally from the criteria that Norm Saunders set out last year in setting up the award for the best technical paper, the general criteria of the selection was first, general relevance. The more people and programs that could use the knowledge, the better. And with that in mind, the paper that's on a generally applicable topic would probably fair a little better than one that was relevant only to a fairly narrow topic of, let's say, a particular program.

The second was the coherence of the paper, did it make sense? The third was the comprehensiveness of the paper, did it cover the topic well? And finally, the knowledge of the topic itself and related research. In addition to that, I thought it was fair to basically discount the entries of members of the organizing committee, so it didn't get like it was a little too incestuous. So under this scheme, two papers were deemed honorable mentions and one paper the award winner. There is no implied ranking of the honorable mentions. I think everything that was given last year was rather fine and it was a difficult task to come up with the winners. But again, I think the most important thing is everybody that participated, thank you very much.

In alphabetical order for the honorable mentions. If you're not here, they go to me by default. Let me get these certificates that Ron and Company drew up.

For Honorable Mention, Best Technical Paper, first honorable mention is for John Hisnanick from the Veteran's Administration for his paper in "Using ARIMA Models in an Unconventional Setting, Forecasting the Demand for Inpatient Hospital Services."

The second Honorable Mention is for Paul Sundell, Economic Research Service, for his paper, "Determinants of Short-Term Agricultural Loan Rates at Commercial Banks."

For the best paper of Federal Forecasters Conference 1992, I would like to award to Jeffrey Butler from Internal Revenue Service for his paper, "Unit Roots and Fractional Differencing of Time Series, Implications for Forecasting."

That wraps up the awards this year. Thank you.

Keynote Address

Mr. Earley: It seems as though, I don't use the term pejoratively, but old conference chairs, just never seem to fade away. They come back as contest winners and award presenters. It's somewhat remarkable to me. I keep on going back to the theme of the people who worked on contributing to this particular conference and you see many of those same names pop up. What was encouraging to me this year is the number of new people who joined organizing committee. I successfully succeeded in doubling its size, which is contrary, perhaps, to the initiatives announced yesterday for the entire federal government. Nonetheless, having more people serve improves the quality of the conference.

I do want to encourage other people to get involved in the Federal Forecaster Conference activities and if anybody is interested in participating in next year's session, let me know. It's fun to be with these people and talk to these people on the phone. And I very much appreciated Tom's efforts over the last year because the strongest communication that you have is the person who just ran the conference last year. Thank's, Tom, for your support and effort.

I want to tell you about the remainder of the day. When we were trying to design what the focus of this year's conference was going to be I had done some reading, basically in Osbourne's book and a few other books talking about reinventing government. When I read the section on the Anticipatory Government, I thought, you can't become an Anticipatory Government unless you are farsighted enough and you are forward thinking enough to anticipate some of the problems.

One of the things that struck me about the comments in the book is that much of what we do or what is perceived to be done in government at all levels, whether it be Federal, State, or local, is now dealing with problems after they've already occurred. In many ways it's much more cost effective and just smoother on the collective psyche if we can anticipate those problems before they occur and attack them before they get out of hand. Unfortunately, I think too many times we are now in a reactive mode and we should be attempting to try and resolve problems in advance.

So in order to accomplish that, I look at this audience of forecasters. In fact, one of the papers this afternoon asks should there be a new occupational category called forecasting? Well, when I saw that abstract, my first reaction was yes, definitely there should be, and they ought to get a one pay grade premium, mainly because of the inherent riskiness and uncertainty of their work. They're always guaranteed to be wrong but you want to try and be in the ballpark so that you can provide to decision makers good information about what likely will happen in the future. Lacking that information, and that information largely stems from people in this audience, lacking that information, bad decisions will be made. If they have good information, at least there's the chance that good decisions will be made. Whether those decisions will be made or not is sometimes more in the political realm and that may be a problem for us as forecasters, but our job is to provide as good a set of information as we can about what likely events are going to be out in the future.

We are involved in all aspects of this process. We are involved in doing the baseline forecast. We're involved in identifying alternatives to the baseline evaluating policy options. Another critical area is evaluating uncertainty.

One thing that has been the thrust of the conversation in this conference in the past and which I think in many agencies not enough effort is spent on, is evaluating our efforts. So we will talk about all of those things during the course of today, both in the keynote, in the panel, and then later on in the afternoon sessions. But the critical thing from my point of view is, in order to do our job well, we've got to make our findings constructive to the decision makers.

With that sort of preparatory comment, I would very much like to at this time welcome Dennis Egan as our keynote speaker for this year's conference. Dennis is a Special Assistant to the Deputy Director of Management and Integration for Vice President Al Gore's National Performance Review. In this role, he has been involved in the overall project coordination, management, production, and strategic communications networking. He is a captain in the U.S. Coast Guard and for the last two years has been a strategic planner on the staff of the Commandant of the Coast Guard specializing in areas of global economics and political and military affairs.

Captain Egan recently served as an economic consultant with the Defense Intelligence Agency in appraising economic forecasts for the People's Republic of China. He was the managing editor for a widely ranging future environmental scan, which was just recently published by the Department of Transportation entitled, "The Road to 2012."

In addition to his numerous degrees, Captain Egan is also a registered professional engineer in the State of Missouri. As you'll note in your program, he's also got a couple other lines on his vita that I would very much like to talk to him separately about. When he came out of the Coast Guard Academy he served with Jacques Cousteau in the Antarctic expedition. And then he also served as a maritime advisor to the Acquino Administration.

Captain Egan has certainly been around and seen a number of different aspects, both economic, political, military, both in the United States and around the world. Captain Egan is now representing his government by serving on the National Performance Review. And it seemed to me that as we look at the role of forecasting in the Anticipatory Government, that it was a natural to have somebody from the National Performance Review here to speak to us today.

So with that in mind, I am honored to present to you Captain Egan from the National Performance Review.

Captain Egan: Thank you. Ron forgot to tell you I've been around a long time. As a matter of fact, I was on Noah's Ark and I was a forecaster on Noah's Ark. It's true. You know, Noah had been going around ranting and raving for a long, long time. "It's going to rain, it's going to rain." Well, they hadn't had any rain in that area for years and he kept building that boat. He was sure of his forecast. Based on the projections at the time, all the king's forecasters said, "You're nuts, Noah. You're nuts. It hasn't rained for years and you say it's going to rain? Forget it." So he couldn't raise any federal Financing for his ark. And so he went out to the private sector and they weren't very interested either. So he kept building and building and he got two of every kind and finally, it started raining. After awhile there was enough rain and slowly the ark rose and we all took off.

Well, he needed two of every kind and so he had a forecaster there too. Normally they called them twocasters then because it was before glasses. But I wear glasses, so I was the first forecaster. Well, here we were floating along for forty days and forty nights. I said, "Noah, it's 8:00 in the morning, time for your daily forecast and here it is. Based on current projections, it's going to rain about another two feet today and tomorrow, more of the same, and I don't see much change out there frankly." Noah said, "Very well, keep it so." And the next day the rain stopped and soon we hit dry land. I got fired so here I am.

That's just to show you in a way the fact that as forecasters in an age of chaos oftentimes we are hit by things we didn't expect. And as things are changing so quickly today as professional forecasters, oftentimes we find ourselves more in the reactionary mode as a matter of circumstance. And to be able to cope with this type of environment, the best way for some of us to survive in that environment is to prepare our leaders and the people they lead for a variety of possible futures. And so they recognize the weak signals when they first start to appear and even though they may be reactionary, they've seen it before and they know how to handle it. That's about the best a lot of us can do right now. And some of us aren't even doing that.

But I had the privilege of serving on Al Gore's National Performance Review and as many of you have probably heard we presented the first of a series of products to the President yesterday. Fascinating, especially when you read the Washington Post. They're the newspapers that say, "Eh, nothing's new under the sun. We've tried it eleven times this century on the average of once every eight and a half years and if you figure that out, that's just a little longer than administrations, and we still haven't got it right."

Well, let me ask you, if you went to Las Vegas and rolled the dice and you finally rolled it eleven times, don't you think you'd come up with a seven? Well, we think we're going to benefit from the confluence of a number of interesting events, most of them not due to our own wisdom.

If any of you noticed since the walls came tumbling down in Eastern Germany and Berlin, we're in the midst of a global revolution. And it's a global revolution that's forced us to look back at the varied pillars of governments. Now what does it mean? A lot of things we took for granted in this country. We were beating a drum and shouting, "Democracy, democracy, whatever the forum may be." We're for democracy and we're against all that's not. And this really drove our foreign policy for years and years and years.

Well, an interesting thing happened. The walls came tumbling down. Eastern Europeans swarmed into the United States graduate schools looking to academia to figure out just what is this democracy? And they started raising questions that we haven't had to grapple with for years. And it raised a greater level of public awareness when Ted Turner's CNN was not only able to broadcast Madonna in little villages out in the jungles of Africa and the Philippines, but the counter effect occurred also. He also was able to bring these images from around the world to show us what's what. And bring them into every house in America, in front of kids. Bring them into the Internet, a global spanning computer network where, even though many of us my age and beyond are computer frightened, we have a problem with computers and the Internet. We ask our teenage son or daughter to please hook us up and we have a wide, wide world around us.

The coup d'etat in the Soviet Union, as things were collapsing. How do you think the people in the Soviet Union found out what was going on? It happened in experiment with the Kosmodrome years before when a group of U.S. astronauts got together in partnership with Soviet Astronauts and they decided they were going to get a computer cable strung between Vienna into the Kosmodrome that would allow the scientists to use the Internet, to be able to share rocket launch information and be able to process it on U.S. super computers and European super computers.

Well, guess what? When the project finished, that cable stayed in. And the underground network of Internet conduit started branching out throughout the old Soviet Union and when things happened, you no longer had to rely on Pravda. Instead, you had packet message after packet message coming out on the Internet going out to the Western World, coming back, beamed by satellite, beamed by TV broadcast, beamed by radio, back to the very citizens that were real, live actors in a drama that unfolded before our own eyes.

So all of a sudden, governments are finding us much more difficult what they do and why, to hide the rationale from the people. And the people are getting smart. And we are seeing this in our own country. This last election was a prime example. These political trends where people all of a sudden developed an acute political awareness of such things as the budget deficit. The fact that we had this growing, growing monster that was increasingly limiting out flexibility to do with government what had to be done. They started limiting what we could do as far as making investments the things that are absolutely essential for our kids and for our own welfare. We just didn't have the financial flexibility to do it unless we fell back on old solutions or just taxed more. Well, the only problem was the economy wasn't healthy enough to just tax more. There had to be a new way of looking at things. There had to be a new way of approaching it.

So you look across the operation and you say, "Well, golly, we've taxed more. We've cut taxes. We've tried to cut government somewhat unsuccessfully by lopping off positions and trying to eliminate departments." We've got all sorts of approaches to this but the one approach we really haven't tried is something that organizations like General Motors, Xerox Corporation, many, many others learned the hard way in the latter '80s, and that's something called competitive awareness, a quality movement empowering employees and a new style of running business lean and mean, and down-sizing.

I don't think the Federal Government has been in this kind of a stark present danger situation before when down-sizing and the environment of down-sizing was so readily apparent to us. A quality movement and the types of management to be competitive and responsive to your customer have never been more clear. And this is an environment that is not just in the United States. It's an environment that's world-wide.

If you take a look from '89 to 1991 something happened that was absolutely amazing. If you go to France, Great Britain, Japan, Canada, or various leading industrialized nations around the world, there are huge vacuum in governance. Not government. There's plenty of that. But a huge vacuum in governance and people started getting angry. People got very upset and right now even you have some of the lowest popularity percentages of leadership across the world, not just in the United States, but across the world. People want government that works better and they want it to cost less.

Is this having your cake and eating it too? Not necessarily. I see some nodding heads in the room. And that is the type of conclusion that somehow or another we've got to grapple with because there are two different attitudes in America. One is very healthy. That's a healthy skepticism. The other is something that's grown and is very strong and apparent right now and that's cynicism. And when you have a healthy cynicism, you've got an oxymoron. You've got something that doesn't work. We call it a dog that doesn't hunt.

There's got to be hope out there and there is hope if you can get out of the way and you destroy all those hurdles that are out there and let Federal workers do what they've always wanted to do. The fact is you can trust Federal employees if you empower them; if you allow them to make the types of decisions they have to make; if you start dismembering this huge gorgon's knot and slice it. You need to remember your Greek mythology. You know that story very well. Alexander the Great came in and he saw this huge, huge knot that nobody could unsnarl before and he said, "I know how to do it" and he sliced right through it. No problem any more. But we're not talking about slicing right through it.

I'd like to rapidly go through some of the initiatives that were announced yesterday. I won't dwell too much on the types of things that you can see in the newspaper, and then I want to conclude by focusing on what does this mean to planners and forecasters, because I think you're going to find that in a chaotic world despite the fact you may not be able to be proactive and really effect how things happen, you can influence both the people at work in Government and the leadership to be more sensitive to weak signals, to understand the type of environment they're functioning in, and give them a framework by which they can make this work. It's vitally important because we only have a few more years. This can't continue this way. We've got to make this work. The time has come.

The principles are pretty straightforward. Basically, a Government that puts people first is what we're seeking, cutting unnecessary spending, serving customers, empowering employees, helping communities to solve their own problems and fostering excellence. The products of the National Performance Review was a report called "From Red Tape to Results." And it meant that there is a way out, there is light at the end of the tunnel. This report contains recommendations to make Government more effective, efficient, and responsive.

At the end of this week you'll have a series of Executive Orders, and the report is one in which probably over half of the initiatives recommended can be initiated by Presidential Action. A series of Executive Orders 1) to mandate customer driven, strategic plans and organizations and 2) to create the President's Management Council, link the Federal Quality Institute into this organization and drive through the chief operating officers of each one of the departments and agencies across the board the initiatives and the spirit of reinvention of government.

There is a package of changes requiring Congressional action and we're working very closely on a bi-partisan basis, trying to get those through. It won't be easy but we've had an incredible ground swell of support in Congress in this measure. The problem was that we had basically industrial era bureaucracies in the information age. But I would like to suggest to you that the information age is over. And as forecasters, I would challenge you to come up with the idea of what's next. We've got enough information out there. And it's coming at us from every quarter.

Now, the question is what's next? And I would submit that it's probably an age of conversation and an age of communication, an age of consensus making solutions, an age of individual empowerment, individuals taking responsibility to make things work instead of coming up with excuses why they can't.

So we looked at a number of entrepreneurial organizations. We found a lot of successes. The Air Combat Command doubled productivity by basically empowering their workers and creating teams across the board, pushing decisions all the way down, removing the regulations that required 20 different signatures just to make routine decisions.

The Internal Revenue Service Center is competing against each other, using customer service performance standards as the marks of competition to determine their own success. Big success story over at IRS.

The Forest Service is streamlining itself again by squeezing middle management, going down to put the emphasis upon the workers in the field and giving them the ability and flexibility to make the decisions and to be able to excise their own budget.

Australia, Canada, New Zealand, Great Britain have all gone through the throes of agony of this type of transition. Canada recently issued their reinvention of government report about six months ago in the Mulroney Administration and it resulted in a 30 percent across the board cut in the operating cost of Canadian Government.

We're being a little less draconian in the United States, really. The winners had these common characteristics: they cut red tape, put employees first, they empowered employees to get results, and they cut back to basics. Are these just generalities? I'd like to get to some more specifics.

First of all, cutting back to basics. Basically requiring a report in 18 months on a base-closing initiative for Federal facilities, not for military facilities. Using more negotiated rule-making where you call it reg neg. And alternative dispute resolution techniques to bring the public, the stakeholders, to the bargaining table and become a good faith member at that bargaining table to hammer out regulations on a consensual basis, rather than a "King Solomon" basis.

Giving the President enhanced rescission authority is an old idea whose time has come. If we are really looking at running a brittle government, a government with the fat squeezed out, the President and his role of leadership has to have more flexibility.

Allowing all agencies greater freedom in setting fees for service and then also letting those agencies use those fees back in their own agency budget for reinvestment, to do their job better and at less cost. I don't know how this last one got thrown in there. That looks like that is probably too difficult.

Well, part of cutting back to basics and part of employee empowerment is also a matter of responsibility. You absolutely will not have the flexibility given the Federal employees for empowerment unless along with that goes flexibility. Right now, for instance, we found several situations where Federal employees had stolen from the Government and were not allowed to be released from their jobs by their supervisors until all avenues of a long, excruciating process were fulfilled, resulting in 18, 24, 36 months after the fact. We need more flexibility. We need more flexibility on the part of management to handle these situations.

We need to improve the process for removing people who are no longer disabled from disability insurance rolls. The process takes incredibly long.

Organizations like Alaska Power that are prohibited by law from going into the commercial sector. Now we have a stabilized market in Alaska, to unload those types of initiatives back and make them competitive again.

Transferring law enforcement functions from the Drug Enforcement Agency to the FBI is just part of an overall initiative to try to reduce duplication and turf battles in the area of law enforcement. There's an awful lot of duplication and turf battles in the area of law enforcement.

Simplify the compliance and certification process, creating a National Spatial Data Infrastructure in cooperation with states and localities. Basically what that is, we've got an awful lot of Landsat information out there. We've got information that can give you images of everything down to your post office number out in front of your house. But what good is this type of information locked up in data banks out in Colorado? We're talking about releasing, getting this information out, getting it managed, getting it linked into state and local government entities where it can really be a tool for economic advancement.

Cutting red tape. I remember in the early '80s when the Federal Acquisition Regulations really came out in reams and reams and reams. And then I remember when the Department of Defense came out with their DAR, which was their own interpretation of that. Their own interpretation was about ten times the size of the original Federal Acquisition Regulations. I remember the Department of Transportation and the Department of Defense interpretation wasn't good enough for us so we came out with another half a dozen volumes called the TAR, and I think it was probably well named. What we didn't have was the feather, but right now we're going out with the feather. We're saying tar and feather the Federal Acquisition Regulation.

We need it simpler. A lot of other countries around the world have done just that. They've thrown out hundreds of pounds of procurement regulations. They made a very, very simple mission objective driven procurement processes with just a lot of authority delegated back to the front line management, front line workers.

Cutting in half the annual cost of headquarters staff and reducing federal government staff by 252,000 employees, which is about 12 percent. Instituting bi-annual budgets and appropriations, two year budgeting cycles rather than a one year budgeting cycle. And getting off of this micro-management by OMB and FTE ceilings that are placed in cost departments and really skew their operations in crazy sorts of ways ending up with lots of temporary employees that don't count against ceiling, all sorts of gyrations caused by FTE management, FTE floors. We're saying we need a process to get away from that and go directly to agencies that are driven by their overall operational budget themselves, and give the agency the flexibility to manage their own people, to hire, to fire, to recruit and these types of things. OMB indeed will stop using FTE ceilings and will instead use caps in operation costs.

We're talking about decentralling personnel authority to departments and agencies to conduct their own recruiting and examinations and simplifying the personnel classifications system, switching to pay bounds, just the personnel regulations. Yesterday it required two fork lifts bringing them out on the White House lawn. It was impressive. The piles were higher than this and that was just for a couple of agencies. I would have liked to seen if we had taken all federal agencies in the Washington, D.C. area and piled them all up. We would have had a mountain.

We want to reduce by half the time it takes to terminate employees. We want the agencies to roll over 50 percent of what they save in operation costs to the next year and establish a process by which agencies can more widely obtain waivers from regulations. And then we need to establish a cabinet level community empowering working group that drives deeply into the Federal relationship with state and local authorities on state local grant aid programs.

Putting the customer first. You're going to see an Executive Order requiring departments and agencies to create customer service programs equal to the best in the business. And this is going to require benchmarking, a skill that some forecasters are very good at. But it's going to be absolutely required to be able to survive as Federal agencies in this new environment.

We're talking about eliminating Government monopolies, such as the Government Printing Office and GSA, and having them have to compete against other alternatives given to the agencies. Create competitive one-stop career development centers. We've got in some states over twelve different employment programs with advisors, lots of field offices, and each one has its own select clientele and anybody else need not apply. And I'll tell you, I have some friends out in Oregon that told me some stories about coming into one and getting this walk around where they finally visited eight different offices, finally to try to untangle what they had coming to and what they didn't. They finally gave up. You can cut a lot of cost by consolidating these types of field offices into one stop shopping.

Restructuring the Air Traffic Control system into a corporation, an initiative strongly supported by the airlines' industry and something that had been tried by several other nations successfully.

Issuing new accounting standards to identify the true unit cost of all government activity. This was an incredible realization as we went through agency to agency. We found we had no common cost accounting system. Nobody knew what the costs of operations really were. But they did know how much money they had to spend and that was what was driving things. No cost consciousness whatsoever. We've got to sort this out. We've got to come to a common cost accounting basis across Federal Government.

And finally, cross Government collaborative efforts to empower communities and strengthen families. You may have heard of worklife. Worklife, does that sound familiar? Putting emphasis in the work place of alternatives for child care? Putting emphasis in the work place for the family focussed type of benefits that are absolutely necessary to level out the working place with the stresses of today.

Enhanced technology, enhancement programs, like telecommuting, are being very heavily emphasized now as not only necessary because the environment and the congestion on the highways, but it's also family friendly. It also allows men or women to work from home and still be able to pull in a paycheck, still be close to their children. A lot of issues here and that's a good one.

Empowering employees to get results. I talked to you briefly about this President's Management Council. It's going to become a very important operative instrument of government. The President's Management Council will have a series of working groups on cross cut issues across agencies. Things like examining a new maritime policy for this nation. Things like rebuilding a new personnel system, a reinvented personnel system for this nation, rebuilding a new procurement system for this nation. These types of things will be driven by the President's Management Council which will be headed by Phil Lader who is currently the Deputy Chief of Management of the Office of Management and Budget.

Establishing performance agreements between the President and Cabinet Secretaries. There is an Act that started out as S-20 that got signed by the President in August, which is going to have a very strong impact upon forecasters and planners because it requires strategic plans to be developed in all agencies across the board. It requires at least every 3 years that these strategic plans be reinitiated, examined, shaken off and redirected. It requires that planning drive budgeting and planning drive budget execution. These are big things. And it requires that the budget execution be measured against the plans in annual reports submitted back to the President's Management Council.

So what we're seeing here is a shift from knowing how much money you have to spend to knowing what you have planned. Then measuring your results from your budgeting against the plan. And I think you can start to see the rationale for shifting to a two year budgetary process because it simply gives you a better horizon. Second of all, eliminating this business that you can't carry money across the fiscal year that results in this surge of spending at the end of the fiscal year. We're talking about eliminating that and allowing money to be carried across the fiscal year.

You saw that reduction in Federal workers. Part of that is due to an initiative to reduce the ratio of managers to employees from one to seven to one in fifteen. This is real cruel shock treatment across the board for middle managers. What it may mean is the whole concept of the middle manager is changed. The whole concept of middle manager is going to be pushed out to the field. A lot more hands on, a lot more closer to the customer contact. Corporation after corporation has found that this has been the necessity of operating in the type of environment that we are headed into in this 21st Century and I'll tell you, government simply has to do the same thing. We can't afford otherwise.

Initiating training at all levels starting with the top. We're talking about sensitizing our leadership with strategic management quality and information technology training because one of the realities we found was that you can't get there without very careful adaptation of information technology and I'm not simply talking about replacing typewriters with computers. I'm talking about conductivity linking up. I'm talking about the ability for employees to take personal empowerment to communicate with each other as Federal managers, Federal employees, to not have to go up and over hierarchial boundaries, to take the initiative themselves to enter the Internet and talk to each other.

Forming a labor management partnership. A labor management coalition work group will be reporting through the President's Management Council, will be taking a careful look at the many issues that affect workers in the Federal work place in this new environment.

Finally, creating a coherent financial management system. I don't want to get into the details on that. That will become more clear when that report comes out.

The schedule of events on what's next. The summary report called "From Red Tape to Results" is now available through the Government Printing Office and the Internet. Interesting, when the President announced acceptance of the report yesterday on the White House lawn, for the first time this report fired up on Internet and was distributed to a potential 6 million networks around the world within minutes of the President announcing it.

The number of copies that were printed by the Government Printing Office were 7,000. Those were gone in an hour and thirty minutes. People were calling like crazy trying to get copies of the report from around the country. And we said, "Talk to your librarian. Your librarian knows how to get on an Internet, if you don't." And we gave them dozens of different sites so they could download that document at no charge and have it in their hands immediately. We gave them the alternative. You can place your order with the Government Printing Office by calling this number and they will have it in your hands within four to six weeks after they have your money. This is truly an example of using technology to have government that works better at less cost.

It's time we had a new customer service contract with the American people. And we guarantee an effective, efficient, and responsive government. It's time we cut the red tape, trim the bureaucracy. It's time we took out of our bureaucracy the words "we've always done it this way." Those were the words of Vice President Al Gore and we're around to try to make this happen.

There's a very interesting follow-on organization called Net Results. That is suddenly appearing on the Internet and it's working its way from one forum, one conference to the next, to the next, just floating out there in cyberspace. And you can become a part of that simply by getting onto the Internet, getting your personal Internet account number from one of numerous ways and entering the discussion that's going to drive these issues deep into local government, state government, individuals all the way up to our leadership and our policy makers.

I want to close by focusing on what does this mean to planners and forecasters. I think you'll find that forecasting and planning has never been more vital to our success. Remember I was talking to you about forecasting for Noah and how I did a simple projection and based on my projection, I thought we were going to get a lot more rain. Well, that's kind of simplistic.

I think in an era of chaotic events, as forecasters to be able to survive and have our bosses survive, we've got to focus on something a little different than our own future. We've got to focus on our customers' future. And the biggest challenge of focusing on our customer's future is first of all knowing who that really is, and knowing that it's not just the people inside our organization as the quality movement in federal organizations have done very well over the last several years pointing out in developing management sensitivity to the fact that there are internal customers. But now we've got to start focusing on our external customers, that other 20 percent of our concerns.

We have to understand how that customer's future may change due to changes in the environment, changes in the economy, changes in the real environment around us. We have to understand how the very nature of a family has changed. We have to understand what this means as far as education and training. We have to think about such issues as health care, and how we're going to bear the burden of health care on this economy and in our organizations and what our responsibilities are, both as employers and also as servants of the public.

I think that focusing on external environmental trends is going to be an easy one for us because we're the best in the field in that but this business of doing GAP analysis, of determining where are we and were did we say we were going to be and being able to map that into the Federal Performance Reports that are submitted back up to the Office of Management and Budget depending on how your organization determines to do this, your services can really come into play in the area of GAP analysis and result measurements.

And then benchmarking, competitive awareness. The fact that you're going to have a lot of initiatives going on around Federal Government for cost servicing agreements. There are going to be a lot of other agencies out there to eat your lunch. You've got to make management aware of that. You've go to let your bosses know that. You've got to know what your competitors are doing. You've got to know who they are. And you have to focus on organizational survival by being the very best in the business.

Making best use of information technology. If you are not on Internet, if your agency is not in Internet, if you have your own internal E-mail communication system that can't talk to anybody, as a Federal employee, do not let that be an excuse for you personally not to enter that world that's out there because it's a tremendous world full of rich resources. You're going to find yourself expanding your capabilities expedientially as forecasters if you learn how to use that Internet. You get out to the rest of the world, you start sharing, providing value added, you'll gain a lot more out of it than you'll put into it. I assure you.

There are organizations like Cap Access here in the Washington, D.C. area that can give you your Internet account free of charge. You can enter E-mail to anybody else that has a Cap Access account personally. We've got lots of details on all sorts of alternatives like this and I would implore you to get involved in this wave of reinvention by personal empowerment.

Finally, Government Performance & Results Act, Public Law 103-62 is the Act that I'm talking about. I don't know if everybody has looked at that or not. It started out as S-20 but it's very important to our collective future. If you haven't read it, take a careful look at it. It's implementation is going to be vital.

And finally, let me just say it's a pleasure to be here. I think you're in the catbird's seat in this new environment that we're entering into. I think it's clear and present danger. I think it's a world wide phenomena and I think we're extraordinarily lucky we live in these times where a rising tide raises all ships whether they intended to be raised or not. So thank you very much. It's a real pleasure.

Mr. Earley: At this time, if there are any questions from the audience for Captain Egan, we'd like to entertain a few.

Question: You've emphasized "customer" in the private sector as the person paying the bill. In the government, that's not the case. How do we deal with that problem?

Captain Egan: The question is that in the private sector, you deal with a customer because the customer is the one that pays the bill. In the government, you have a different situation. We all pay the bill.

Well, that's not entirely correct first of all. We're finding in an awful lot of cases that we have to do things in very specific way for a specific customer who was represented through the Congressional process and then micro-management is written right into our authorization. And whether we like to or not, whether it means good management or not, we've got to do it. And this is just one example of the type of entanglement we're trying to get cooperation from Congress. We're trying to do in an Administration that's got both Democrats and Republicans but in Congress for the first time, a Democrat in the White House plus a Democratic majority for the first time in twelve years. We think we have a rare consensus opportunity with some of the Perotites, in the feeling that Perot represented both sides of the political spectrum. We have strong Republican constituency that wants to have government that costs less and that does things better.

So what I'm trying to say is we've got the forces there that say, look, focus on the segment of the American population that your program was meant to serve. Congress, give as much flexibility as you can to those people who are performing that service and use the quality initiatives, use the professionalism that we all have if we're just allowed to function. And you're going to find a government that does things better, that can recognize a customer and this customer is the people we serve.

Why is it when we've got 2.1 million Federal employees that if I call most agencies in the Washington, D.C. area, I can't get a real person to answer the phone unless I know somebody in the agency. Then I can weed my way in, finally find out who's there. I mean, there's no excuse for that sort of thing. But that's what Americans are finding out about their Government. It's simply not responsive. The Americans are your customer but we're all smart people and I think we can focus pretty much upon who our target is as far as delivery of services go. We've got to know what their needs are. We've got to know what they need and don't need and we've got to have the flexibility to respond to that.

So this is not a matter of who pays. This is a matter of common sense. We've got common sense in our Federal, State, and local government. It's just the fact that we haven't been allowed to use it in the past. I think it's high time that we have that responsibility.

Next question?

Question: (Questioner not on microphone.)

Captain Egan: The question is one of time lines. First of all, over this next month you're going to have a series of agency specific and system specific monographs that get into some detail on what the initiatives are as far as procurement system, as far as finance from the government, management systems and these types of things.

You'll see the Executive Order for the President's Management Council stand up at the end of this week. And during these next two, three, four years, you're going to see the system start falling into place. You're going to see a lot of opportunities for employee involvement, for public involvement and basically this is going to take a process that has a horizon of about four or five years. And depending on the system you're talking about, depending on the flexibility of whether Executive Order or legislation or just simply bringing diverse opinions to a consensus, those dynamics have to have time. They can't be mandated top down. We've got to achieve consensus on it.

So figuratively speaking, I'd say a horizon of five years to get things in place but about 40, 50 percent is going to start the locomotion out of the station by Executive Order and by decision on the part of Cabinet Secretaries and administrators.

Question: (Questioner not on microphone.)

Captain Egan: The question as I understand it is how can we get more cross fertilization, I hesitate to use that word, cross fertilization from the corporate sector, from very good managers and consultants, I might say, into the Federal sector. It's a good question. This is really a question of training. And not only training, but I think probably employment opportunity on the part of those that are in that business.

How many of you have not heard of the Federal Quality Institute? Can I see a show of hands? Probably the majority of us haven't heard of the Federal Quality Institute. And I think that probably was because of the way it was handled by the previous administration, but it's alive and well and it was created a couple of years ago to bring the commercial sector, to bring business and business managers, those associated with the quality movement, these types of management techniques, to bring those types of educational resources to the table as far as agencies across the board that needed that type of assistance. With all the red tape cut, if you please, to be able to bring those types of people on board to help your particular situation out in your agency and be able to measure whether you're really accomplishing what you set out to do with your quality initiatives in your agency.

They're an important resource and their role is going to be considerably enhanced by linking their efforts to measurement initiatives with the President's Management Council and also increasing the involvement of public/private partnerships with the Federal Quality Institute in providing those services to agencies. So that's just one. There are many other initiatives I think that as Federal managers we will be able to bring on line, especially if some of the red tape can be cut to bring this stuff on line.

As far as employment opportunities, I really don't see a very bright picture out there in an environment of down-sizing for a lot of lateral motion in and out of the Federal work place. I see more out and less in but there's going to be some turnover and I think if you have these types of programs and you're emphasizing the importance of your management, at least as far as the training of those that are in government, the money is going to be made available as an investment to get those people trained in the latest management techniques and these types of things. So I see a hot and cold here, okay? The environment is not very conducive but the emphasis of getting to where we have to go requires it so I guess the answer is yes and no.

Question: These various suggestions that you want to reduce duplication and waste, and on the other hand you want to encourage bringing agencies together and doing the same thing. And you cite the law enforcement as one example. So on the one hand to get efficiency you create monopoly and on the other you say we don't want Government monopolies. You're going to have to have both. If you really want to encourage competition, you should encourage more Government agencies to do more things and compete with each other. And of course, if you really want pure and beautiful competition of the classical model, you should have so many of them that none can influence price.

Captain Egan: Thank you. That's great. Seriously, I think you're absolutely right, but it's situational. There are certain things in Government where you want to compete. Absolutely it's healthy. But you know the last thing I want is to have the police in Vienna competing against each other for who can write the most tickets.

The idea of having a police force compete against another police force to try to write more tickets, to enforce the law is not necessarily the competition you're looking for. You're looking for, I think, competition as far as alternatives of either cross servicing agreements, where on service provides a service internally for another agency because they can do it better and they can do it cheaper. And that allows the other agency flexibility, reorganization to use their people better. This is just one example of the types of hybrid arrangements, I think, where competition can be very productive.

In other cases you may be looking at a mean and nasty environment out there where you have certain agencies trying to take over another agency's business as far as operations go. You've got them both doing it and after a period of evaluation or something like that, you say whoever wins will do it, whoever doesn't win has to get out of the business.

But I'm not talking about situations like we've had in the drug czar's office off and on for years where you've got turf battles just raging and unnecessary duplication of effort. I'm not talking about the border patrol and the customs and the types of turf battles that they've had that may not be too familiar to you but when we got into them, they were an incredible waste of money. So it's situational. You've got to take a careful look at it. Competition is great but competition can get out of hand too.

Mr. Earley: One last question.

Question: Over the years we have traded costs for the protection of certain groups. The minorities, women, disabled and so forth in our procurement policy and in our personnel policies. Are you willing now to give up all of those advances in order to conserve cost alone?

Captain Egan: No, absolutely not. There is a social cost of justice and there's an ethics of good government. And we're not going to turn our backs on that, absolutely not. Instead we've got to take a look at it from their point of view and their accesses to services in eliminating all the unnecessary steps they've got to do to get what their program should be offering them as a matter of law.

Unfortunately, a person that's in distress, a woman who may have children is out of a job, her house is going to be foreclosed on her, she has all these things coming down on her head in this environment and now she's got to go out and she's got to weed her way through 17 different programs just to get well baby money so she can keep her child healthy? I mean something's wrong there and there's just so much red tape. I think that if we can consolidate situations like that for one stop shopping on these types of programs along with the other service delivery programs we have, we're going to be doing a much, much better job.

Mr. Earley: I'd like to have us all thank Captain Egan for speaking to us today.

Panel Discussion

Mr. Earley: In thinking about the next part of the program, one way to view the discussion in terms of the role of forecasting in the Anticipatory Government is to take kind of a slice in time in the analytical time it takes to evaluate certain programs, to making a decision, and then eventually acting these decisions into law.

We all, as I mentioned earlier, have responsibilities for creating the initial baseline, evaluating various options, coordinating the analysis of these options across agency and across governmental lines and then working, of course, with Congress in terms of the ultimate resolution of these and the packaging these into programs that are put in Law.

So with that in mind, I would like to introduce to you Mark Rodekohr. Mark is presently the Director of the Energy Demand and Integration Division within the Energy Information Administration. He manages approximately 35 economists and other professionals. Dr. Rodekohr is responsible for mid-term and long-term modeling within the Office of Integrated Analysis and Forecasting.

He is also responsible for working on the EIA's Annual Energy Outlook and the International Energy Outlook. His major responsibility right now is the development of the demand and integration components of the National Energy Modeling System, an effort we've been working on within EIA for the better part of two years now. Prior to this current position, Mark was responsible for the International and the Contingency Information Division where he was responsible for much of the international data and analysis functions within EIA.

So to talk generally about the development of baseline and the analysis of policy options, I'd like you to welcome Dr. Rodekohr.

Dr. Rodekohr: Thank you. I'm afraid my talk is going to be a little bit more mundane than the previous one. I'm not going to attempt to reinvent government today. Although I probably will have a few things to say about that because I just read part of Mr. Osbourne's book last night, and I found some interesting corollaries between what I was going to talk about and some things he said in his book.

I'm going to talk about two things this morning. One is creating a credible baseline projection, and that word credible is a dangerous word. And secondly, creating a policy analysis tool and I'll rely on our recent experience with reinventing our models. It's always a good idea to reinvent one's models. We seem to do it about every ten years and it's driven by necessity of changes in the policy environment as much as anything else.

So first I'll talk about establishing a baseline. What I'm going to do here is list a series of rules and talk about the rules and why we found that they appear to work. We've been issuing a projection of the energy economy since 1973 on an annual basis so we've got about 20 years of these projections under our belt at this point. Unfortunately I've been there for most of those 20 years so I also have a lot of experience there.

But nevertheless, a few things come to mind. One, unless you're dealing with short-term forecasts, just concentrate on trends. Trying to predict the timing and amplitude of cycles in the economy and energy prices is dangerous. It leads to a lot of controversy and you're seldom right. And it drags the discussion into an area where you'll find yourself ignoring the basic issues or policy issues that are of interest to people into talking about things of short-term nature that really policies can have little influence over.

So the moral of the story here is to concentrate on the trends and I'll show some graphs about that in a few minutes.

Secondly, explain differences from the conventional wisdom. Conventional wisdom is an amorphous thing. It changes every year but nevertheless, it's what many of the policy makers are thinking about, are being told is the way the future is going to be and right or wrong, if you're different from the conventional wisdom, you owe it to them, you owe it to the establishment of accredibility to explain why you're different and why the conventional wisdom is perhaps in error.

The third point is to always compare your forecast with those of others. I hate to say it, but quite frankly, people don't have a lot of confidence in federal forecasters. Their errors are noted more than their successes and therefore, it becomes imperative to understand and compare your forecast with those of others made by whatever party. But policy makers see others. They want to understand why you're different. They want to understand what the implications of those differences are for their policies. Does it make a policy a high risk option, a low risk option, whatever? And lastly, and this is sometimes over emphasized but uncertainty analysis can be very successfully used to incorporate other views.

A long term forecast of, let's say the economy, if you incorporate all the uncertainties, you explode your error bands on out, the policy maker will take one look at that and they'll say well, the economy can be between $5 and $15 trillion in the year 2010. Well, that doesn't tell them very much. That's not very helpful. It doesn't give them a sense of what they can do to enact a change or get to a position they want to be in. It just says nobody knows what's going on. So you have to be a little bit more specific than that.

Now, I want to talk about a phenomena. Being wrong is sometimes being right, or helping the situation. I haven't picked on our projections of world oil prices. I've used the International Energy Workshop poll, which is a group of people who project world oil prices. As you know, this is a very volatile market. The history has shown here the price shock of '73, of 1980, the resulting collapse in prices.

Nevertheless, projections made in the early '80s, the conventional wisdom was, well, oil prices have been going up and they're going to go up forever. And these projections were made in '81, '83, '85, '86. All that period of five years prices were going to go up, we had to do something about it. Well, the reality was, of course, the prices didn't go up and in fact, I've often said that if I walked into a forecasting conference in the early 1980s and said what actually was going to happen to prices did happen, I'd be out of a job. And I expect that's true.

Nevertheless, part of the reason prices didn't go up, is because these projections were saying, if something wasn't done, the status quo would be for this situation to happen. Policy makers saw that, said well, we can't live with that. That's too much of our GNP going to energy and governments in the U.S., outside of the U.S. Policies were enacted which helped to bring about the moderation in prices.

So it's an example of where being wrong is being right. You've pointed out a possible problem, policy makers reacted to that problem, and other events happened to make that not become a problem. And so it's important that you establish the reasons why you think things are going to go the way, if nothing else happens. Then the debate can take place. Then the debate can focus on, well, is this really going to happen if I don't do anything? And what do I need to do to effect it? And that's the important part of the policy maker.

Now, on to uncertainty. Everybody who's an oil market analyst has an opinion about what's going to happen to world oil prices. It's probably one of the most widely focused on variables. We've dealt with this historically by using uncertainty to wrap a band around the projections of plausible prices and then talking about what happens if prices try to move outside of this range. And we've argued, and I think successfully, that market forces will tend to push them back in here. In reality, prices may bounce around in the future. I don't know. I can't predict the timing or the political events that will lead to supply disruptions or other things that will control the actual stability or instability of prices. But I can say with a fair amount of certainty it's unlikely the prices will lie outside of that band. That, to policy makers, I believe is useful.

The uncertainty range helps establish your credibility of your baseline. Even though people might not believe it, they're much more comfortable with the range that you see here and it helps add another dimension to your problem.

Now I'm going to turn to the area of policy analysis. And then again, what are the keys here to what we're doing. There's been a lot of talk this morning about customers and clients and so forth. I've used the term 'users' here, but you'll find that spending a lot of time identifying who uses your product and for what is very important. Then you've got to ask yourself as a manager, these views are sometimes competing and you can't make everybody happy. And so in a sense, part of your job is to decide where the emphasis should be, where the real policy focus is? At the federal level usually, because we are in the Federal Government. We constantly get comments, I need state forecasts or I want a forecast at the county level. Well, we can't do all of that.

So we have to make it clear who our users are and what is the goal of the analysis. And that's not so obvious and in fact, it takes a lot of work and a lot of thinking to define the goals of the analysis. For example, in the energy market our model probably projects thousands and thousands of different things. Only a few of those are really important to most policy makers. The level and price of oil imports is one of them that rightly or wrongly is a big focus. The cost of energy, the ability of the energy segment of the economy to damage the rest of the economy. Those are things that we've developed and started to look at with a lot of care and a lot of time.

That is, what should be measured is very important. Not everything should be measured or can be measured in your model.

Next, as an economist I have a bias and that is to focus on measuring the cost and benefits of proposed policies. Oftentimes this is not done. I was very dismayed last night when I was reading Mr. Osbourne's book on reinventing government. He noted several successes. One of them was a fire department in the west. They privatized the fire department. They made it mandatory that all homes have sprinkler systems and they said, you know it's a success. I mean, not as many buildings catch on fire. The fire department costs the city less. And that's exactly the problem with that kind of analysis. It's focused on the benefits. It hadn't focused on the cost at all. You don't put sprinkler systems in buildings for nothing. I mean, this is a cost. It may not be a cost to the government, but it's a cost to the consumer. No discussion of cost and benefits.

I find it extremely useful if in any analysis, people can step back and start to think about things in terms of costs and benefits. To whom is it a very important political issue and to society is it a very important economic issue.

Carefully explain the results in a simple, understandable manner. That can't be emphasized too much. Black box answers seldom make for a credible forecast or policy analysis tool. The killing words are "well the model says." You just can't use that as a descripter. I get in more trouble every time I say that and I keep trying not to.

The last thing I'm going to mention here is our reinvention of our model, the National Energy Modeling System and walk through an example of how we've been developing this and the kind of things we've done differently. And hopefully, this will be a little bit useful.

First of all, we took a lot of time in developing a simple statement of the objective of the model. What are we all doing this for? And this was to illustrate the energy, economic, environmental, and security consequences on the United States, not every country in the world, of various energy policies and assumptions by providing forecasts of alternative energy futures in the mid and long-term using a unified modeling system. That may seem rather simple and rather obvious but it does help to break the problem down and define it into something that's manageable. It's not designed to do everything for everybody. It simply can't. No model can do that.

Now, how did we go about doing this? Well, we started two years ago and we started with a small internal group that developed a series of working documents, sort of overall design documents. We also got the National Academy of Sciences involved. They gave us a lot of support and direction as to what the overall make up of the model should look like. What was good from our existing model? What should we keep, what should we throw away? How should we do it?

We then spent a tremendous amount of time and effort, and this cannot be over-emphasized, identifying your clients, to work with your client through your customers. And we started off with a broad-base. There is a tremendous industry in this country and around the world in energy. I think it's $400 to $500 billion industry in the United States. Something on the order of seven or eight percent of our GNP is devoted to energy related activities.

We then broke that industry down into the various components, the Governmental components, the state and local components, the industry itself, and those who study the industry. We attempted to go very systematically through each one of those components and to develop groups who review and coordinate wherever we could. In the department we had monthly working group meetings. We developed very detailed design documents before we actually built the model. And staff tells me we over designed it in a stack of reports about that high. It doesn't take a forklift to bring them in, but close to it. We developed very detailed design reports, which we distributed widely. I think the mailings were up to 8,000 or so the last time I looked, all over the place.

We developed an independent expert review program where we got outside experts in the various fields to come in and give us advice on the design and on the implementation of the models. We used a group of academic, business, and other associates, called the Energy Modeling Forum, to help us with part of the design and review effort. And we've had academic review as well, so we've tried to hit all parts of the community that have studied energy situations or energy issues and have an interest and an input into how they think it should be modeled.

We then developed outside users' groups. We went to trade associations. We went to, then again, back to academics. And we went to industry as well. And we've had a conference, of course, and we'll have another one, I'm sure. But anyway, as you might imagine, all of this took a tremendous amount of time, effort, but it was probably worth it in terms of we think we've got a model that has a fairly broad-base of consensus about what it should be.

Now we've made mistakes. We've probably gone off in a few wrong directions based upon some very strong comments from people that are very articulate but you're going to have some of those, no matter what you do. Anyway, the model is near completion, at least in its first version. It appears to be working as we think it should although it's certainly not without a lot of problems. But we do believe that it is a model that will be able to look at the issues of the '90s and not rehash the old wars of the '80s and the '70s.

But there has taken a tremendous amount of change in the energy industry in the last two decades. Some of that from a modeling sense, you can put behind you. And then you have to be forward looking. And we think that we've developed a way of doing that by carefully going out and talking to all of the constituents, people that use or have opinions about the forecasts and analysis and have incorporated their concerns into our effort.

I'd like to thank you for your time.

Mr. Earley: Next I'd like to introduce Heather Ross. Heather Ross serves as the Special Assistant to the President for Economic Policy at the newly formed National Economic Council. I somewhat pride myself in getting two of the new organizations in town on the dais today. I think it's good to hear from them. They represent some change here in Washington.

Dr. Ross has been involved in a lot of coordination efforts in her recent tenure here in the National Economic Council working on the Btu tax that was part of the original package submitted by the President. She is currently working on the initiative of reducing greenhouse gas emissions.

Dr. Ross performed public policy research at the Woodrow Wilson School and the Urban Institute. She joined the federal government working at the Senate Budget Committee, and then at the Department of Interior. During the '80s she moved to overseas assignments with British Petroleum and now she has returned to Government service since 1993 at the National Economic Council.

Please join me in welcoming Dr. Ross.

Dr. Ross: Thank you. My topic is coordination and from my arrival here over the last few months, I have to say that one of the things I've learned is that coordination is more than the mechanical coming together of people. We have a lot of working groups. There is a lot of convening and talking about where to go. There's not as much shared vision about where that is and that's what I'd like to talk about.

If we have a view of where we want to go, we can reinvent government to get there, but we really do need to start with a view of the future that people can rally around. And so when I saw the topic on Anticipatory

Government, I thought there really is in the country a new element seeping into our anticipation of the future. It's pessimism and it's fear.

This is not fear of an external enemy. We've coped with that. It's a concern about how our own economy is functioning, how society is functioning. And how if you have winners and losers in the new challenges to us, right among our own compatriots, how do we reach agreement on a way forward? The answer to that must be that we find some common ground in how to go ahead together. That is where people who look to the future come in. There are many people plying their trade. The importance now is to find some way to bring our views together into one that we have in common about what the future looks like and how we can prosper in it.

The opinion polls that we see show pessimism. Substantial majorities of people are concerned that we're on the wrong path. Confidence indices are down. And we have 20 year lows in insurance rates but not a surge of buying of houses or other major commitments. People are worried and there's a tremendous irony here because the world that we face is one that is better than we ever anticipated.

During the Cold War, we hoped to get people to develop along our model of democracy in markets and we've succeeded remarkably. There's major liberalizing trends throughout the world. Other nations are in fact becoming more like us and that is the challenge. They're coming our way and they're coming our way very rapidly.

The prospect for growth now in the Third World is six percent a year, maybe three times what the OECD is. And so we have an engine here of global growth. We have a center of gravity that is moving toward the developing countries. In 25 years they'll have perhaps 70 percent of the entire world GDP. This is an enormous change and it's enormous good news for us. It's a fact about the future that is central to our opportunities for success.

But we have to change in order to be part of that world. We have to move ahead as it catches up with us. It's hard to compete with a Mexican peasant coming down from the mountains and working in a plant. Maybe he can increase his productivity ten times over. We can't do those same things any more. We have to do new things.

We know that markets are really major drivers of those new things that we will do. But markets are in fact what are concerning people now with layoffs, with downsizing. We have to get through this period. We have to see into a future that looks beyond the transition that we are in now.

And so the idea that we will, in the work that we do, find a view of the future that we can share and that we can convince people is open to us and that we have the capacity to do things both in the private and the public sector to get there, is the most important decision that we face right now.

There is a wave out there that we have to ride, rather than trying to hold back those waters. And not only the actuality of the experience of that in the economy but the policy judgments to support that, are occurring not in some future time but right now.

The example I'd like to use is NAFTA, which is not this week but next week, as the harbinger. This growth in the Third World. Mexico is certainly in the forefront of that in terms of liberalization and in being right next door it is the instance before us. It's the event that we have before us right now. It's four percent of our economy, the Mexican economy, and pooling our future with theirs will be net gains for us in terms of growth of exports to their growing economy.

Those are net gains. That is, there will be some jobs that will not continue. Those have been estimated in the 100,000 to 150,000 level. This is almost nothing compared to a giant economy of our own and so it's not a make or break for us. This is an important juncture in the policy direction that we take, in the view that we have of the future. The symbolic importance of it is exceptional. We can choose a path that says we can handle what the world is becoming. We can be part of a virtual circle where we encourage that, where we are benefitted by the growth in the world generally and that we make the changes we need to in our own economy in order to bring everyone along so that they can be part of those gains and so we recognize a new world and we are going to stay in the lead in that new world.

The other side of that coin is a negative feedback loop where we say we can't handle the opportunity or the possibility of free trade with Mexico, which is an extraordinary statement for us to make, that we don't see how we can make adjustments in our own economy that will allow us to welcome the world that is coming toward us. We won't put in place the programs that we need to put in place in order to be able to stay ahead of the wave.

So we're looking at a pretty major struggle on NAFTA between these two visions of what the future holds and of our place in it. And this is not a conflict only of people's current interest. It's a conflict of world views. It's a conflict of how we see where the world is heading and where we are heading.

It's easy as we go into this to play on people's fears. In an economy where we've had real wage declines for a substantial part of the population for 20 years, alarming people is not difficult to do. The challenge is to see how to have this be an experience where people who have some constructive thoughts about the future and coordinate ways of facing the future can convince people that there are benefits for us here that are not worn off, that the path will be chosen, which is critical to a world that we're moving into, that there are gains and that the winners can compensate those who do not win initially. So that we don't have to fight it out over our current circumstances.

This may not seem to be the obvious thing to say to a group of forecasters. It's the essence of what looking at the future is, I think, for all of us, that we look out at the world and we see fundamentally how favorable it is to us. This is an enormously promising world. It's built in our image. It's trying to be like us. And so as we go about our work and try to look at what it is that we face and to develop a common view about it for all of us whose work is to try to anticipate, it's critical now that we see how we can make a new era of success for ourselves in this rather than a fight over the current concerns that we have about our future, that we still have a lot of opportunity to shape.

The only solutions to these things will be done collectively and so the set of people who are in the forefront of this have to be the set of people in the government who are working ahead within their own spheres to try to convince people about what the future is and how we can prosper in it. And so I am happy to be here and to try to put what appears to be these days a counter spin on what the future looks like and to wish all of you well in trying to convince people of an affirmative view of what we can do now together.

Thank you.

Mr. Earley: Dr. Ross brought up NAFTA. I think we've all heard the radio ad that plays on a lot of people's fears. I concur with her comments that it is our job to try to be analytically sound and to provide information to people who are making the decisions, to get rid of the hysteria of fear.

I would like to introduce Mike Woo who is our last speaker. Mike is a professional staff member for the Committee on Energy and Commerce, a position he has held since 1981. In this capacity he is primarily responsible for economic and regulatory analysis of a number of issues primarily focusing on energy, health and environment. In each of these efforts Mr. Woo is essentially involved in the development and analysis of policy options and probably as important if not more important, the development of consensus about what is a desirable policy option to put in place.

Mr. Woo served as a consultant to the Department of Energy through the firm of Sobottka and Company. He coordinated and assisted in the Interdepartmental Study of Contingency Planning Options to deal with severe oil shortfalls. He also served as a consultant to the Board of Education of the Chicago Public School System on the development of a desegregation plan there.

With all of the panelists today I've had some connection with them cutting through the issues and so it's with great pleasure that I introduce now, the legislative perspective on this Anticipatory Government. Please join me in welcoming Mike Woo.

Mr. Woo: Thank you. Let me first start out with a few comments about forecasting and then talk about the legislative view.

Earlier in my career, and I've been here almost as long as Mark, I focused more on economic analysis. I thought I was pretty good at what I was doing. However, I was put to the test with a group of assorted people who were interested in energy; Danny Boggs, Mike German, Dan Dreyfuss -- a whole bunch of us energy policy junkies. Each year during the early part of the Reagan Administration we would make a forecast of what the price of energy would be at the end of the year, how much imports would be, etc.

It was very sobering to contrast how good I thought I was when I actually had to make forecasts that were benchmarked against reality. I always ended up in the bottom half. However, even as bad as I was, others were worse. In the early '80s I also worked on the policy of whether or not we should legislate certain special provisions for the then potential Alaskan Gas Pipeline. An economic analysis I prepared concluded that if there was anything less than three percent real growth in the price of oil starting in 1980, as Mark showed was at the top of his chart, this pipeline was not going to be economic.

And as Mark showed, the forecasts of most people were way up there. I made the bald statement that it couldn't possibly be sustained at that level. The proponents of the pipeline got together a number of eminent energy forecasters and did a forecast of what the price was going to be. They showed "conclusively" that the price was going to be $100 a barrel in 1990 and that the lowest it could possibly be would be $50 a barrel. That was in 1980 dollars and, of course, it wouldn't even make us close even in nominal dollars. So anyway, I've had some good and bad experiences in forecasting.

When I told a colleague of mine, who is a lawyer, I was going to speak at the Federal Forecasting Conference he said, "Wait a minute. Isn't that a contradiction in terms? Isn't that an oxymoron? Isn't that like military intelligence? Isn't there something wrong?" I said, "Wait, wait, wait.

Let's explore what he was talking about. One, maybe he was talking about Federal forecasting. I think it goes without saying that Federal forecasting is as good as private forecasting. Private forecasters use Government data and forecast. Government forecasters use private data and forecast. I don't buy the concept that the Federal Government is any worse than the private sector.

Maybe it's Federal forecasting. I think that may be where the problem is. Everyone remembers the forecast. No one remembers the caveats. The main caveat in all forecasts is there are no facts about the future. A very good Washington radio sports call-in-host is often asked to predict the outcome of games and he always says he can only evaluate strengths and weaknesses. And, then he makes a little forecast. However, he also says if he knew what the future was going to be, he would make one bet, retire, and become the biggest sports junkie in the world. Similarly, I think those who work in the Administration and Congress are often asked to forecast the future and most people remember the forecast but not the caveats.

However, saying that, the Congress depends on forecasts in much of its legislation -- from the Federal budget, to the environment, to health care. Let me give you some examples of what the dependency is and interaction between how good those forecasts are and what the policy turns out to be. The Federal budget is one of the more obvious examples. There is always a forecast of what the Federal deficit is going to be and the pressure always is to reduce the Federal deficit.

Now, if you can trust forecasts, that is, if you can trust what you think the baseline is going to be and you can trust what you think the policy changes are going to achieve, then often you depend on individual policies to achieve the overall goal that you're trying to achieve. In other words, you're trying to bring the deficit down so you enact a series of individual initiatives that you forecast will bring the deficit down. What has been shown, though, is that these forecasts have tended to be wrong on the low side, i.e., that the deficit keeps growing even though the forecasts would lead you to believe that if you achieve certain policies the deficit should start coming down.

What that leads to in terms of policy is a desire by a number of Members to put absolute limits on budgets. That was one of the battles that recently occurred in the House. Some of the conservative Democrats said what does CBO forecast health care entitlement spending to be? Okay. We won't cut that spending. All we'll do is limit spending to that level. The forecast acts as a spending cap. That in fact made it into the House bill and it was dropped eventually in conference because it violated the Byrd rule. But that type of analysis may come back again when we consider health care.

A different type of forecasting occurs in global warming. It's more physical, although, the underpinnings are still basically economic. It's essentially forecasting fuel use and efficiency of use. It often misses central questions like what are the emissions and a similarly difficult question about the future -- what are the environmental impacts of these emissions?

President Clinton committed to achieving a reduction in global warming emissions to 1990 levels by the year 2000. We face the same types of questions we faced in terms of trying to reduce the Federal deficit. A number of people will try to focus on individual policies. They say: here's what the baseline is and if you do these individual policies, you will reduce the emissions down to this level. Therefore, we should carry out these individual policies. There are those, however, who want to say let's put an absolute cap on emissions and reduce everybody's emissions to achieve 1990 levels in the year 2000.

I think the battle of forecasting has yet to be resolved. I believe the Administration will come up with a series of individual initiatives to try to achieve some reductions and the environmental community will be unsatisfied. Perhaps, the two potential options will battle it out next Congress.

A final area that I want talk about is how forecasts are used in the health care reform legislation. One of the main issues is the rapid escalation of private sector health costs. Continued growth in health care costs at the rate that has occurred over the last decade or two is unsustainable. If you do the graph, you go back 50 years, and project forward with a straight line, by the middle of next century, 100 percent of our money is spent on health care. While absurd, it nevertheless raises a series of questions about how sustainable this is and whether or not we need to change policies. My hope would be that we would be as lucky as we were in energy. We had that same forecasts of unrelenting price increases. If it turns out that actual health care prices go down like actual energy prices went down, our policies will be a rousing success.

But anyway, the same questions are going to occur in health care that I talked about earlier. What's the baseline? What is that health care growth going to be? Is it going to sustain its double digit increase in cost? How do the policy changes that occur as a result of health care legislation change that baseline and again, the questions of relative effect and absolute effect. Should we try to achieve individual policies which will affect that result or put on absolute caps to try to keep that health care cost at a certain level.

It's also interesting to note that in terms of the overall policy, if health costs continue to go up at anywhere near the way they're going up now, the policies are probably not sustainable even with the changes that the President is talking about. If you talk about enacting an employer mandate to cover health care costs and health care costs keep going up as a percentage of payroll, it just won't work.

On the other hand, if the costs are contained and if anything like what happened to energy starting in 1980 happens to health care, then any policy will be thought of as a rousing success.

Let me summarize by saying everybody complains about Federal forecasting and, as I said, everybody remembers the forecasts and not the caveats. But having said that, it also is true that Federal forecasting is very integral to policy making in the Congress. Finally, I would like to say that I really liked Mark's presentation. I have been here all through that period of time. My chairman, Mr. Dingle, is considered to be the father of EIA and I have handled EIA issues for him. The way Mark described the NEMS model as being generated is the way that forecasting and policy models ought to be generated. The NEMS model will be used in the context of many policy discussions including global warming.

Thank you.

Mr. Earley: I want to thank the panel for their comments. The afternoon sessions will begin at 1:00 p.m.

FFC-93 Survey Results

Debra Gerald, U.S. Department of Education, National Center for Education Statistics and Karen S. Hamrick, U.S. Department of Agriculture, Economic Research Service

During FFC-93, we conducted a survey of conference participants. Our intent was to get basic demographic statistics on Federal forecasters. For FFC-93, 45 out of 315 conference registrants completed the survey form. Because there is no guarantee that this is a representative sample, we cannot apply the results to all Federal forecasters, or even to all FFC-93 registrants. We were pleased that so many people participated in the survey, and that most of the survey questions were answered by all of the respondents.

Over 60 percent of the respondents at FFC-93 had a degree in economics. Other fields represented at the conference were accounting, demography, geography, information science, mathematics, operations research, and sociology.

Nearly 30 percent of the respondents were female. Also 30 percent of the FFC-93 respondents were managers. On average, the FFC-93 respondents had nearly 12 years of federal service. This group of respondents had been forecasting for 9 years. Thirty-six percent had a Master's degree and 41 percent had a Ph.D. Of the FFC-93 respondents, 84 percent published their forecasts. Also 84 percent of the FFC-93 respondents did evaluations of their forecasts.

Most of the respondents include national forecasts in their scope of work. Regional/state forecasts concerned 36 percent of the FFC-93 respondents and local forecasts concerned nearly 20 percent of the respondents. Slightly more than 20 percent of the FFC-93 respondents were concerned with international forecasts.

In terms of primary forecasting techniques, respondents cited a variety of methods, including trend analysis, regression models, time series models, macroeconomic models, and judgement.

Among the issues facing Federal forecasters listed by respondents were data consistency and quality, forecast accuracy, budget resources, reliability of forecasts, and modelling uncertainty.

		CHARACTERISTICS BY OCCUPATION TOTAL RESPONDENTS	MANAGERS	NON MANAGERS
YEARS OF GOV'T SERVICE	(AVERAGE)	11.9	15.9	10.1
DISTRIBUTION:	0-4 YEARS	27	8	35
(PERCENT OF TOTAL)	5-14 YEARS	34	31	35
	15-24 YEARS	32	54	23
	25+ YEARS	7	8	6
PERCENT MALE		71	77	69
AVERAGE GRADE (EXCL. EXECUTIVE SERVICE)		GS/GM 13.0	GS/GM 14.2	GS/GM 12.5
PERCENT GS/GM 13		22	15	30
ED HIGHEST DEGREE:	ASSOCIATE'S	2	0	3
PERCENT TOTAL	BACHELOR'S	20	15	23
	MASTER'S	36	38	35
	PhD	41	46	39
YEARS OF FORECASTING	(AVERAGE)	9.2	11.9	7.5
DISTRIBUTION:	0-4 YEARS	41	23	48
(PERCENT OF TOTAL)	5-14 YEARS	30	23	32
	15-24 YEARS	27	54	16
	25+ YEARS	2	0	3
PERCENT WHOSE FORECASTS PUBLISHED		84	92	81
PERCENT, FORECASTS EVALUATED		84	83	84
OF WHICH, PERCENT EVALUATION PUBLISHED		41	55	35

		CHARACTERISTICS BY SEX TOTAL RESPONDENTS	MALE	FEMALE
YEARS OF GOV'T SERVICE	(AVERAGE)	11.9	12.4	10.4
DISTRIBUTION:	0-4 YEARS	27	15	58
(PERCENT OF TOTAL)	5-14 YEARS	36	42	17
	15-24 YEARS	31	39	8
	25+ YEARS	7	3	17
PERCENT MANAGER		30	31	25
AVERAGE GRADE (EXCL. EXECUTIVE SERVICE)		GS/GM 13.0	GS/GM 13.4	GS/GM 12.0
PERCENT GS/GM 13		25	30	9
ED HIGHEST DEGREE:	ASSOCIATE'S	2	0	8
PERCENT TOTAL	BACHELOR'S	20	12	42
	MASTER'S	36	39	25
	PhD	42	48	25
YEARS OF FORECASTING	(AVERAGE)	9.2	9.9	7.1
DISTRIBUTION:	0-4 YEARS	40	30	67
(PERCENT OF TOTAL)	5-14 YEARS	31	36	17
	15-24 YEARS	27	30	17
	25+ YEARS	2	3	0
PERCENT WHOSE FORECASTS PUBLISHED		84	82	92
PERCENT, FORECASTS EVALUATED		84	84	83
OF WHICH, PERCENT EVALUATION PUBLISHED		42	39	50

Note for tables: percent distribution figures may not add to 100 due to rounding.

FORECASTING TECHNIQUES

SESSION I

MODERATOR Peg Young, Department of Veterans Affairs

PRESENTATION 1	John Parker, Economic Research Service
PRESENTATION 2	Alberto Jerado, Economic Research Service
PRESENTATION 3	Andreas Muller, Department of Health Services Administration, University of Arkansas at Little Rock

SESSION II

MODERATOR Sandra Absalom, Bureau of Mines

PRESENTATION 1	Ramesh Dandekar and Nancy Kirkendall, Energy Information Administration
PRESENTATION 2	Frederick Joutz, Energy Information Administration and the George Washington University, and Harry Vroomen, Economic Research Service
PRESENTATION 3	Kay Adams and Gus Mastrogianis, U.S. Patent and Trademark Office

Forecasting Techniques

Use of a Trade Matrix System for Data on Saudi Arabia and Iraq

John Parker, Economic Research Service

Adequate trade statistics from the official sources in Iraq are difficult to obtain. Thus, a matrix of data from suppliers provides a tool to estimate Iraqi trade activity. Through July 1990, statistics for Iraqi trade in major agricultural commodities were available through the USDA Agricultural Trade Office reports from Baghdad. Since August 1990, economists and business people have not had the advantage of that source, and a computerized tabulation from trading partners provides the best estimate for Iraqi agricultural trade.

While trade books are published by Saudi Arabia, they are on a delayed basis, and 1991 was the latest official trade book published in late 1993. Saudi statistics for imports of many commodities are far below those reported by trading partners exporting agricultural commodities to Saudi Arabia. The matrix tables help economists to obtain more current and realistic estimates for Saudi Arabia which is the world's leading importer of barley and sheep.

It is possible to get an approximate idea of the quantity of food imported by a given country through compilation of export statistics from trading partners. This is a useful tool when the official statistics are either not available, or they report a number far below the expected range. The development of data which is in the appropriate range is a necessary step in building a history for production, trade, and consumption used in forecasting. FAO (Food and Agriculture Organization of the United Nations) in Rome, Italy, and USDA in Washington use the trade matrix system to obtain estimates for imports of specific agricultural commodities by certain countries. This paper reviews the use of trading partner data for obtaining estimates for imports of selected commodities by Saudi Arabia and Iraq, with a strong focus on cereals, vegetable oils and sugar.

The first step is to gather all the trade information which has been compiled from United Nations trade tapes and printed for the format compiled by the Economic Research Service of USDA. This provides trade information through 1990 for over 60 countries and for 1991 for over 33 countries. Second, manually add in statistics from available trade books of countries not yet compiled in the United Nations system, including FSU, Iran, and Yemen.

Third, add in data for countries not in either the UN system, or with proper official trade books, but with reliable unofficial reports on exports to Iraq, like Vietnam's rice. Fourth, make estimates for the transit trade through neighboring countries. This includes an evaluation of the transit trade through Dubai, UAE to Saudi Arabia. Many shipments destined for Aqaba, Jordan are known to be completely for Iraq. Fifth, review various publications using the trade matrix system, especially statistical bulletins of the International Wheat Council.

Saudi Arabia's Import Statistics and Matrix Comparison

Saudi Arabia was the leading importer of agricultural commodities among developing countries during 1990-92, with imports averaging nearly $5 billion annually (table 1) and (figure 1). Average prices paid by Saudi Arabia are near the world average, or slightly below because of the excellent Saudi banking system and modern inexpensive unloading facilities (table 2). Despite wheat exports which averaged about 2 million tons during 1990-92, Saudi Arabia's grain imports were over 7 million tons annually. Corn imports trended upward, with larger purchases from the United States, Argentina and China (figure 3). Feed imports allow local producers to provide about 40 percent of the poultry meat sold in stores. Imports of a wide range of products allow Saudi supermarkets to have the best selection of food available to customers in the Mideast at relatively low prices. Intense competition among supermarkets means low prices.

Statistics reported in the official Foreign Trade of Saudi Arabia through 1991 came within 25 percent of the range of the tabulation obtained from a matrix of suppliers for livestock products and horticultural items. However, the official statistics for imports of barley, corn, rice, and sugar were less than half the matrix total during 1988-91.

This means that the matrix estimates must be used to develop an appropriate data base for forecasting future Saudi Arabian grain imports. Saudi Arabia is expected to remain a growing importer of feed grains, and in the next decade it is likely to be one of the largest importers of barley in some years, along with Japan and FSU. Some forecasts indicate that FSU feed grain imports are expected to decline as domestic production rebounds. This will cause Saudi Arabia's share to total world imports of feed grains, particularly barley, to rise.

During 1992, Saudi Arabia was the leading world importer of barley, purchasing nearly 4 million tons from the European Union (EU) alone (table 3). Imports of over 2 million tons from other suppliers pushed total Saudi barley imports to over 6 million tons in 1992 (figure 2). U.S. barley exports to Saudi Arabia declined from 1 million tons in 1991 to 669,000 tons in 1992. Canadian barley exports to Saudi Arabia averaged about 1 million tons annually during 1990-92. During 1988-91, the official statistics reported for barley imports were less than a fifth of the amount reported shipped to Saudi Arabia by trading partners.

The Foreign Trade of Saudi Arabia reports imports of corn and sorghum together, but the numbers are usually less than half the matrix number for exports of corn to Saudi Arabia. As a matter of fact, U.S. corn exports to Saudi Arabia are usually greater than the officially reported imports from all sources. In 1992, Argentina exported 296,000 tons of corn to Saudi Arabia, while U.S. exported 497,000 tons. Shipments by China, EU and other suppliers led to a combined matrix total of over 950,000 tons for shipments of corn to Saudi Arabia in 1992. Saudi Arabia's official statistics tend to record over 90 percent of the imports of basmati rice reportedly shipped by India and Pakistan, but the number for imports of rice from the United States and Thailand is less than half the expected level.

Saudi Arabia imported about 625,000 tons of rice in 1992, and during 1988-91, its imports exceeded 520,000 tons annually, about double the official import numbers. Saudi official statistics for imports of rice from Thailand and the United States are usually well below the amount shipped by these two suppliers.
In the last several years, India emerged as a leading supplier of Saudi rice imports (figure 4).

One reason that the statistics for Saudi Arabia's imports of specific commodities may be that they are recorded in a miscellaneous category, rather than that for the specific commodity. Efforts to use computers to classify where imports of a given commodity should be recorded may have may have aggravated the problem. Through 1986, official trade statistics for some commodities like barley were in an acceptable range, but then the import numbers plummeted in the late 1980's. The number of personnel to review the trade statistics for accuracy were apparently reduced, and errors in commodity classification of import invoices for the trade books increased.

Detail statistics for many commodities are provided in Foreign Trade of Saudi Arabia according to the Harmonized Statistical System. This includes statistics for imports of items like peanut butter and chocolates which are in an acceptable range. However, imports of frozen potatoes reported in the book are less than a third the shipments to Saudi Arabia reported by the Netherlands.

Matrix Data Used to Review the Rise and Fall of Iraqi Imports

The trading partner matrix data indicates that Iraqi agricultural imports increased sharply in the 1980's and peaked at about $2.8 billion in 1989 (table 5) and (figure 5). Imports were close to the same pace in the first half of 1990, but declined steeply after August 2. The sharp decline resulted in a setback in consumption of most food items, particularly livestock products. The trade matrix indicates that Iraqi agricultural imports declined about a third in 1990, and fell steeply to less than $800 million in 1991, before making a rebound in 1992, with larger imports of wheat, flour, rice, sugar, and potatoes.

The decline was virtually complete for Iraqi imports of mixed animal feed, corn, and soybean meal. This resulted in a crash for Iraqi broiler meat output (figure 6). An unusual surplus and intense competition among rice producers in Southeast Asia worked to Iraq's advantage, with very low rice prices and a willingness of exporters to provide credit. Iraqi rice imports declined in 1991, but rebounded in 1992 (figure 7), and reportedly reached a record in 1993 with Thailand and Vietnam as major suppliers (table 4). Wheat imports have remained far below the peak levels of the late 1990's. Although Australia has resumed shipments on a reduced level, no U.S. wheat shipments have been reported since the summer of 1990, causing total wheat and flour imports to lag (figure 8). This combined with only token feed grain imports caused total grain imports in recent years to remain far below the peak of 5.4 million tons in 1989 (figure 9).

Nearly a third of the agricultural imports came from the United States in the late 1980's, but other suppliers were available. Total Iraqi imports declined sharply during late 1990 and fell below $800 million in 1991 (table 5). no U.S. food exports to Iraq were reported in 1991, compared with $808 million in 1988.

Regardless of where the basic food commodities came from, Baghdad's planners were determined to reduce shortages and to provide a reasonably adequate diet for the people through 1990. Yet, imports luxury items like processed foods were restricted soon after the war with Iran started in September 1980, and particularly after 1983, with an exception of limited imports from Turkey for foreign technicians and hotels.

Iraqi agricultural production could not keep up with the rising demand for food in expanding urban areas as imports lagged in 1991. Through 1990, an artificial boom evolved in the cities related to government spending for projects and for the military. Efforts were made to greatly improve housing, creating a lot of jobs in construction. In addition to the rising population, the average diet improved as incomes increased and income for the poorer urban population improved. Regulations and disputes over land ownership tended to disturb plans for expanding food production by Iraqi farmers. After 1984, programs focused on raising producer prices helped to boost production of poultry products and vegetables helped increase output in central Iraq.

Trade statistics from matrix tabulations are used to prepare food balances for Iraq, in combination with data on production of specified commodities. Statistics from the Iraq Statistical Abstract through 1991, provide information on grain production by province and production statistics for many specific crops provide an indication of what was happening in the 1980's. For 1992, official data is not yet available from Iraq Statistical Abstract, although some preliminary estimates for cereal crops harvested were made by the FAO office in Baghdad. The data indicates because of modest grain yields, and less than half the normal imports, that the average Iraqi diet in 1992 contained about a third fewer calories than during 1990.

The use of the official Iraqi agricultural production numbers with estimates for imports developed by FAO from trading partner information clearly show the growing reliance on imports in the late 1980's, and especially in 1989 when drought reduced Iraq cereal yields, which lagged behind neighboring countries.

Neighbors Maintain Significant Trade with Iraq

Exporters in neighboring countries have been more aggressive in obtaining United Nations waivers for exporting food to Iraq. Many small traders have found Iraq a convenient market. Since Turkey and Jordan depend completely on imports for petroleum, and the low prices and convenient location of Iraqi supplies is a major reason why traders are willing to cope with many problems encountered in trading with Iraq. Jordan has continued to buy most of its gasoline needs from Iraq, while Turkey has turned to other suppliers, particularly FSU, GCC, and Libya.

Turkey is the leading source of Iraqi food other than wheat during most of the summer months. The border trade with an exchange of seasonal fruits and vegetables from Turkey moving by truck through Silopi Pass to Iraq rises in the summer, and declines by January. The trucks return to Turkey with as many drums of gasoline as possible. The price for gasoline in Iraq is less than 30 cents per gallon, compared with $4 in Turkey. Kurds have tried to get a toll for trucks passing through their territory, since Saddam Hussein has his own boycott of trade with the Kurds, and forbids the sale of gasoline to Kurdish truck drivers.

Turkey made large exports of many items to Iraq during 1981-83, but reduced deliveries during 1984-89, prior to the sharp rebound in 1991. Prior to the Gulf Crisis, the barter arrangement which allowed Turkish traders to deliver many commodities to Iraq and then receive payment at Turkish banks handling accounts for Iraq's petroleum exports ran into problems. The Turkish currency steadily declined in value. Slowness in paying them through the barter arrangement sharply reduced their profits. The traders shifted heavily to exports to less troublesome markets, where prompt payments in foreign exchange were made. Turkey's agricultural exports to Iraq declined from $346 million in 1983 to only $61 million by 1989. Turkey's deliveries of food to Iraq increased sharply in 1991 and remained strong in 1992. Official Turkish data show large exports of wheat and flour to Iraq through Aqaba in 1991 and early 1992. Turkey's exports of sugar and tomato paste to Iraq were strong in 1991 and 1992.

Since Iraq delivers about $300 million worth of petroleum products annually to Jordan, and gets Jordanian currency as payment, food purchases in Jordan are easier to obtain than elsewhere. Jordan's agricultural exports to Iraq trended upward in the late 1980's, consisting mostly of wheat flour, eggs, and processed foods and beverages. Jordan's agricultural exports to Iraq reported in its official statistics rose to $64.2 million in 1991, compared with an average of $22 million annually during 1987-89. A striking increase in exports of vegetables from Jordan to Iraq occurred in 1991, when the quantity reached 100,000 tons, including 57,000 tons of tomatoes. While Kuwait banned vegetables from Jordan and 90 percent of the Saudi business was lost in 1991, Jordanian vegetable exports brought $275 per ton, compared with $119 per ton in 1989 when Kuwait and Saudi Arabia were major markets.

For political reasons, Iraq had little trade in food with Iran and Syria during 1980-89. However, imports of apples, pistachios, raisins, and dry apricots from Iran increased in 1991 and 1992. This occurred because Iraq suffered a severe shortage of luxury food imports at the same time that Iran was increasing exports of these commodities. Kurds found it advantageous to begin some trade with bordering Syria in 1991, since the ban on trade with them by Saddam Hussein for the rest of Iraq became a problem, and they no longer had to abide with his ban on trading with Syria. Kurdish farmers derived considerable benefit from imports of wheat and vegetable seed from Syria, allowing them to increase yields.

Conclusion

The use of a matrix for trading partners will allow economists and statisticians to make estimates for imports of specific commodities which are in the expected range. The estimates are subject to change, because it may be discovered that some significant suppliers were left out of the beginning matrix. Yet, the use of the matrix estimate contributes to the development of historical data to make reliable forecasts.

This helps to get a better measure of world imports demand for selected commodities, and contributes to a forecasting system which is more polished than data systems using official statistics only.

Since a number of countries are late in reporting trade data, the matrix is adjusted with manually added numbers used to provide numbers for imports in recent years not yet recorded in the electronic series, which would show up as a 0 until the country data is put into the system. For example, Vietnam provided good data on rice exports to Iraq through 1992, but its data is not yet properly recorded in the UN trade data system. Another major tool in filling the gap for Iraq trade during 1980-90 was the trade book from Moscow reporting trade between Iraq and FSU, which was never a part of the electronic data base.

Figure 1

SAUDI ARABIAN AGRICULTURAL IMPORTS

By Specified Commodities, 1987-92

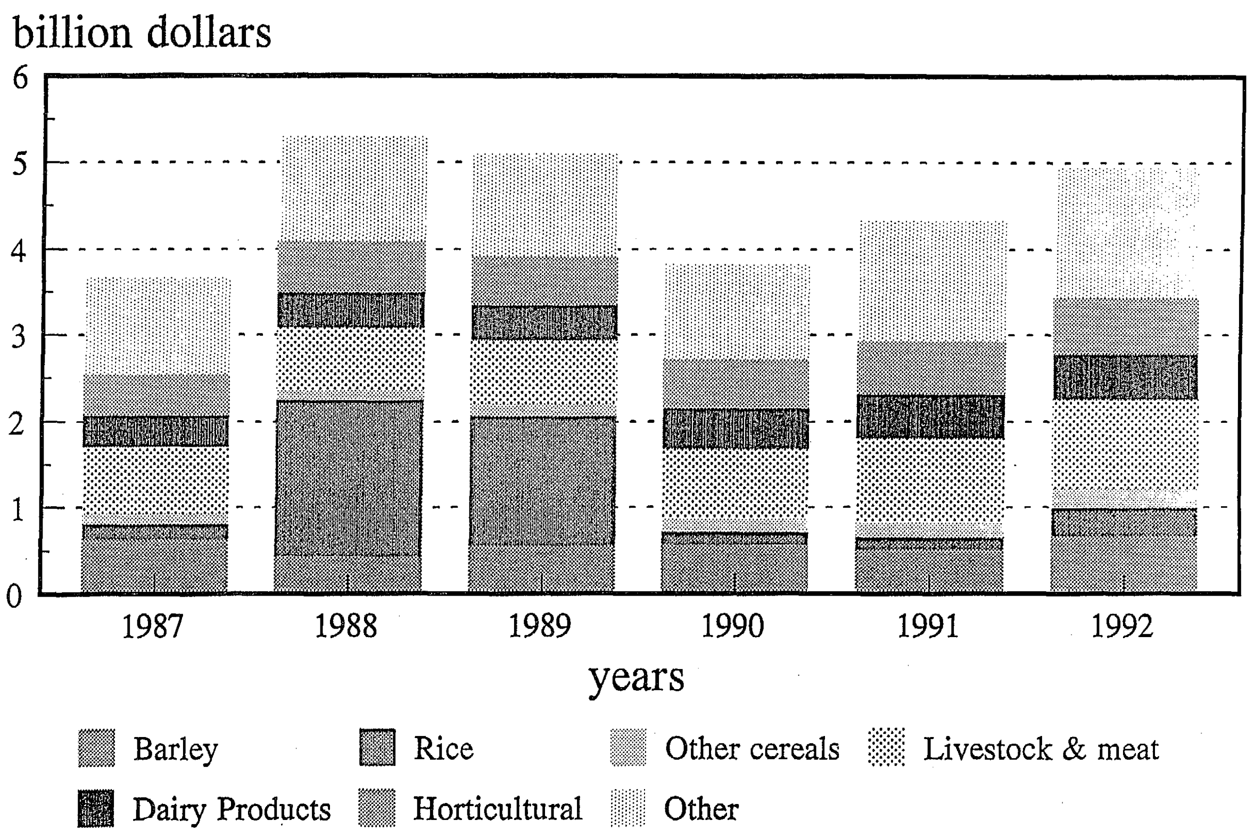


Figure 2

SAUDI ARABIAN BARLEY IMPORTS

From Specified Suppliers 1985-92

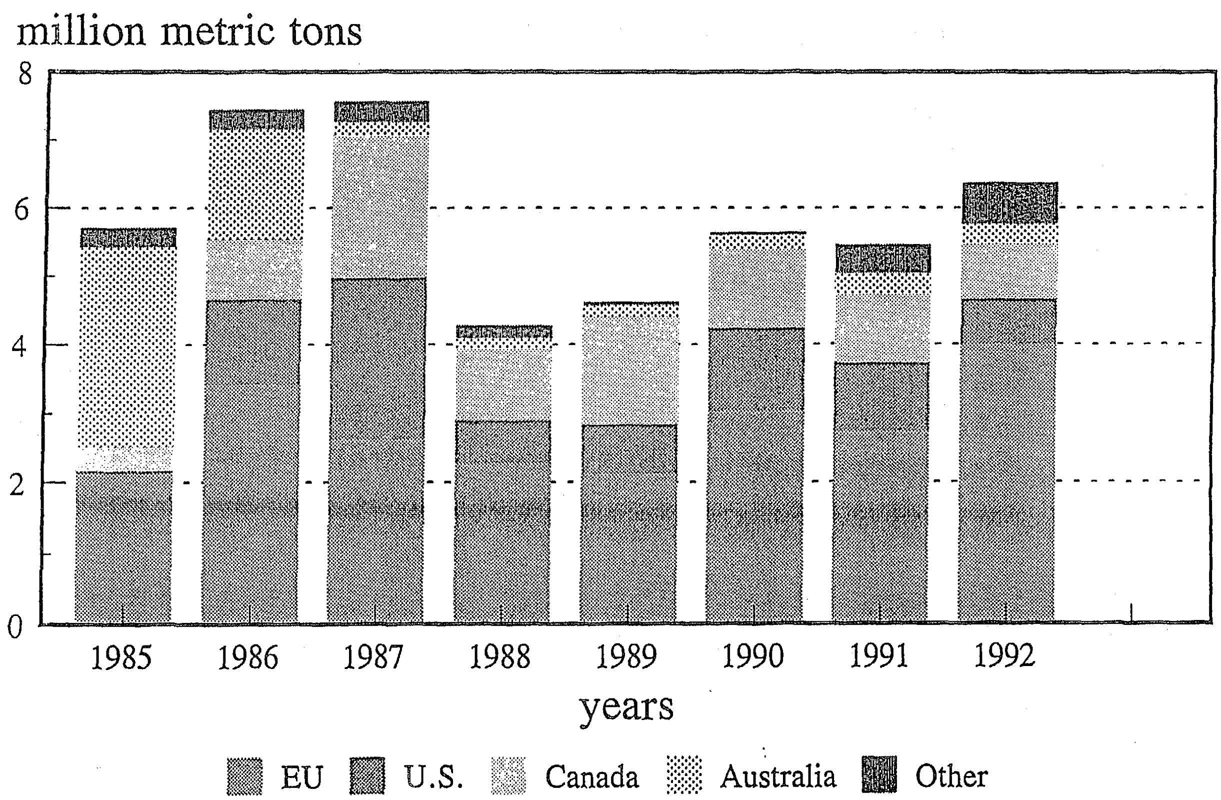


Figure 3

SAUDI ARABIAN CORN IMPORTS

From Specified Suppliers, Annual 1985-93

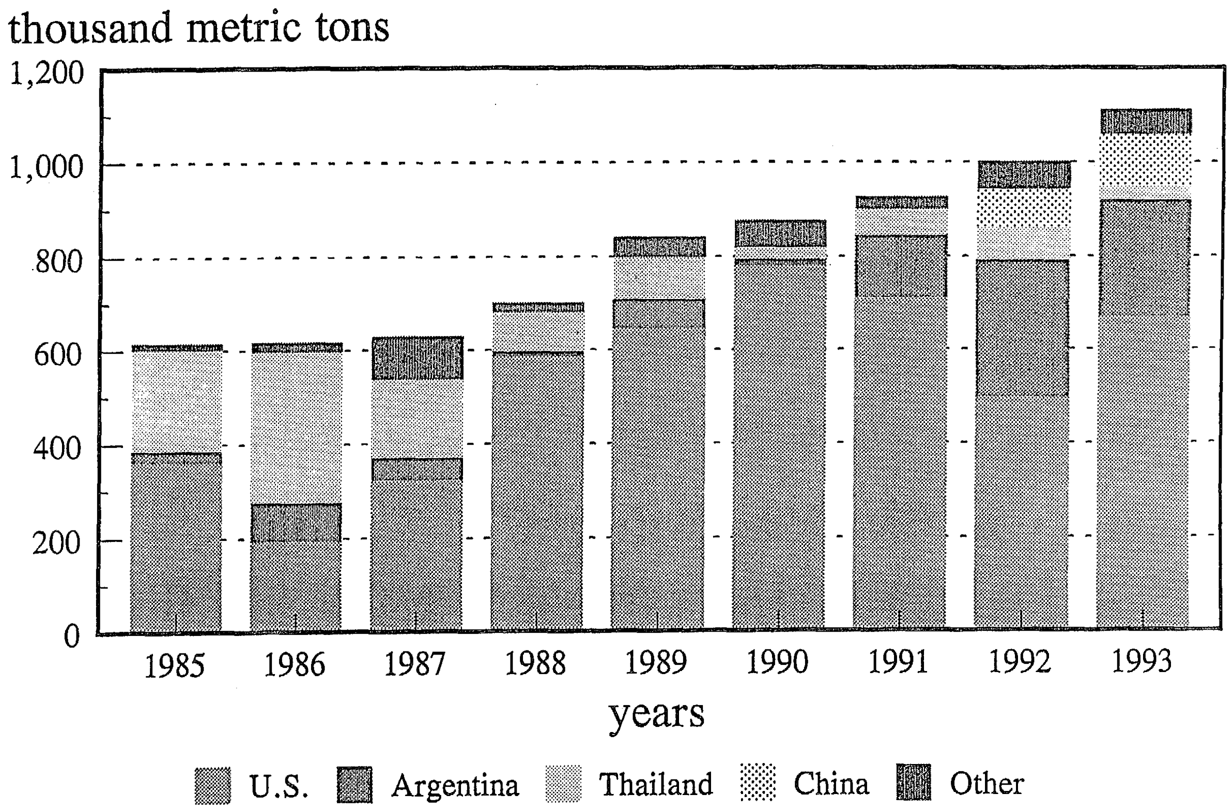


Figure 4

SAUDI ARABIAN RICE IMPORTS

From Specified Suppliers

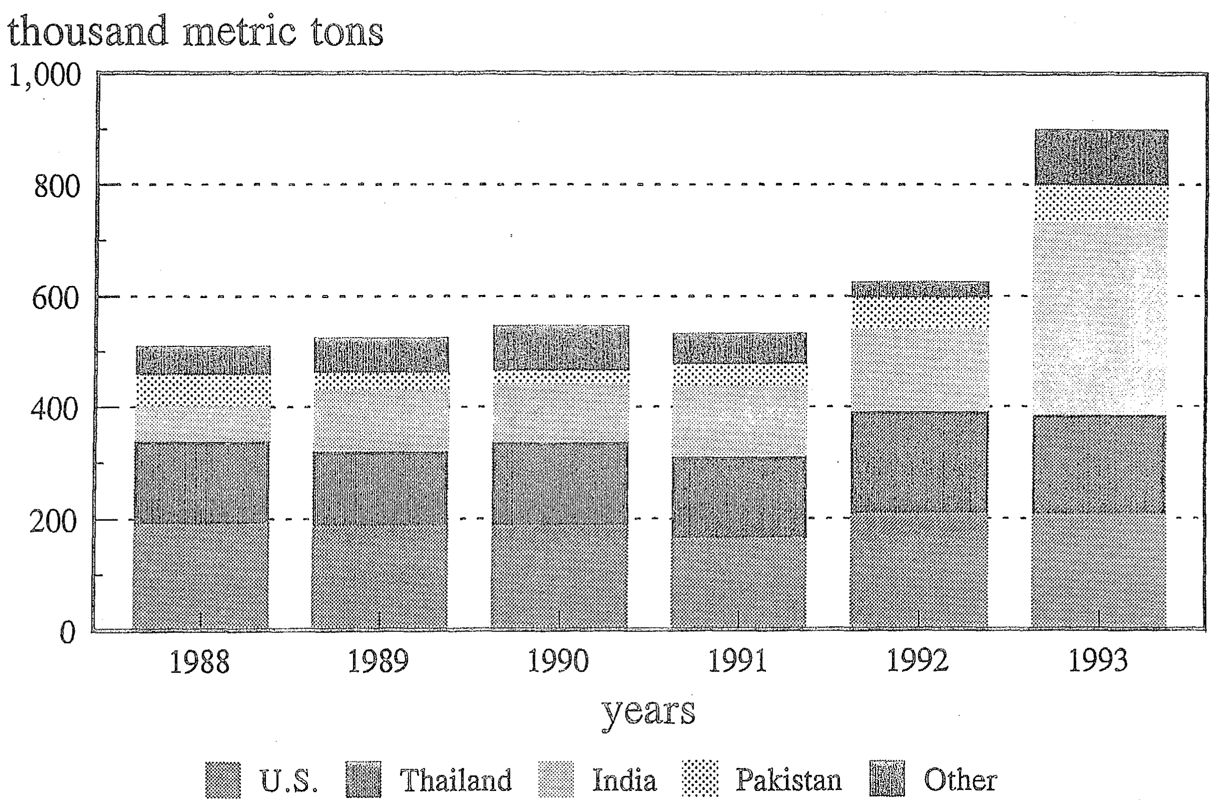


Figure 5

IRAQI AGRICULTURAL IMPORTS

BY COMMODITY AND VALUE, ANNUAL 1988-92

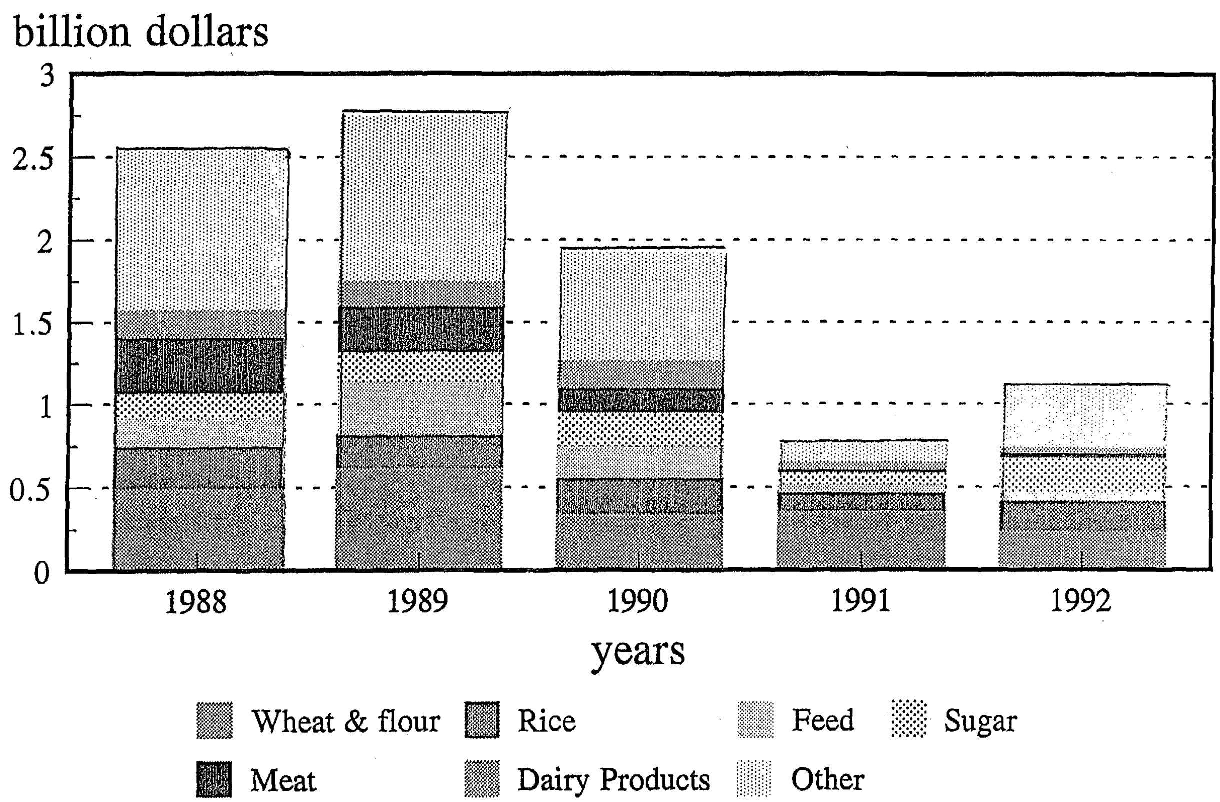


Figure 6

IRAQI POULTRY MEAT SUPPLY

OUTPUT AND IMPORTS, 1980-92

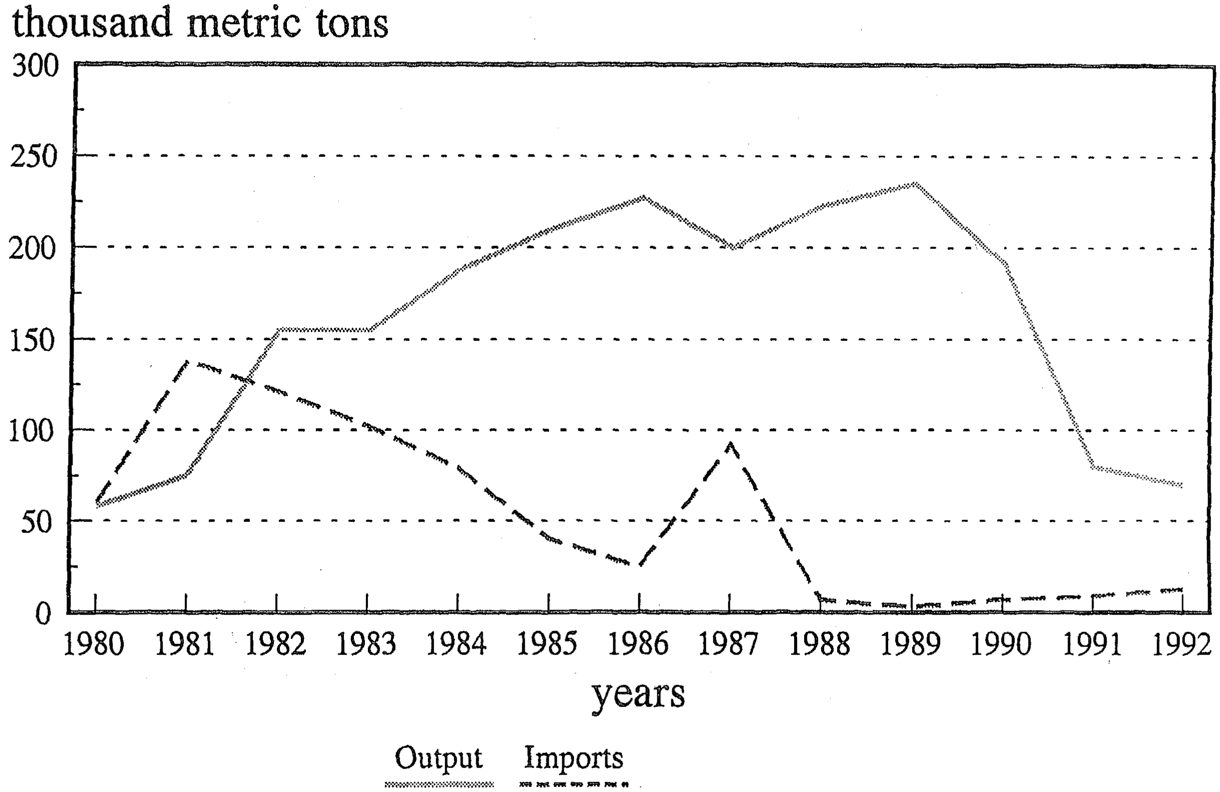


Figure 7

IRAQI RICE IMPORTS

From Specified Suppliers

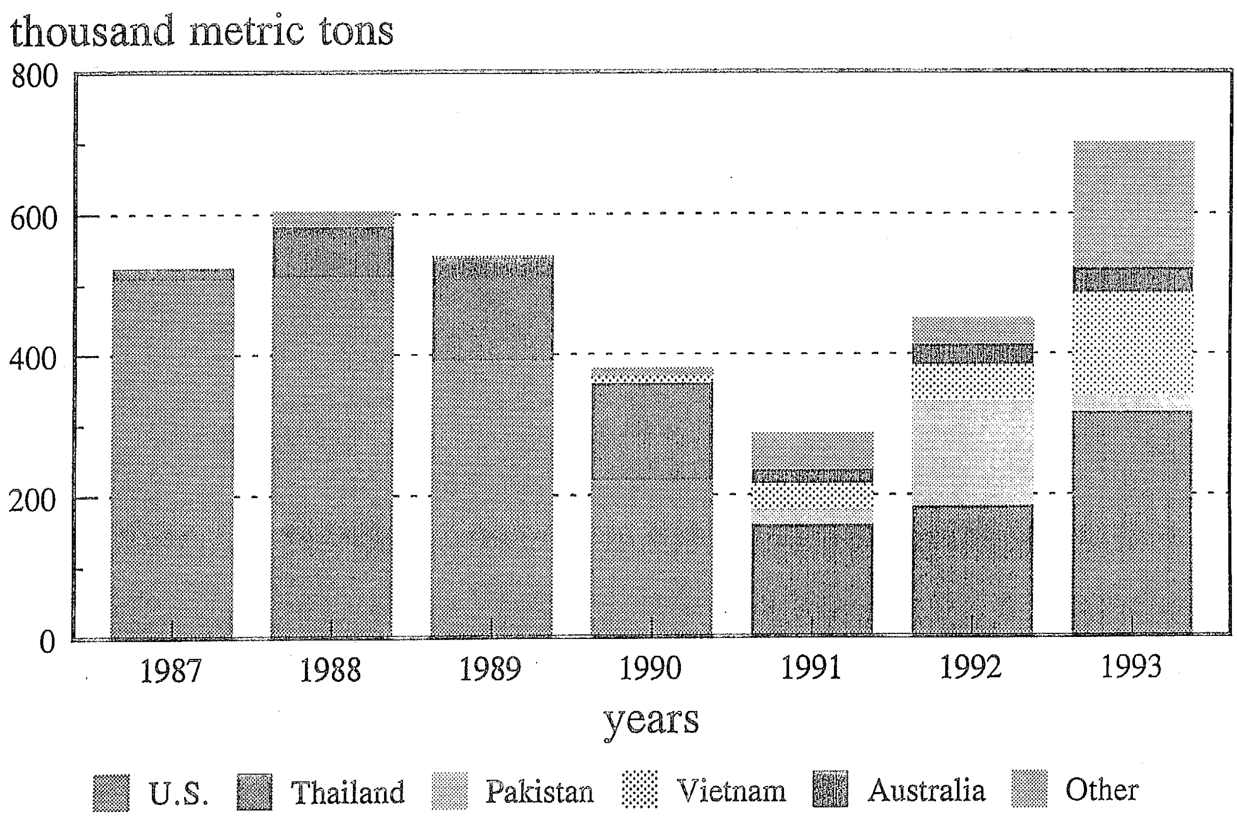


Figure 8

IRAQI IMPORTS OF WHEAT AND FLOUR

FROM SPECIFIED SUPPLIERS, 1988-92

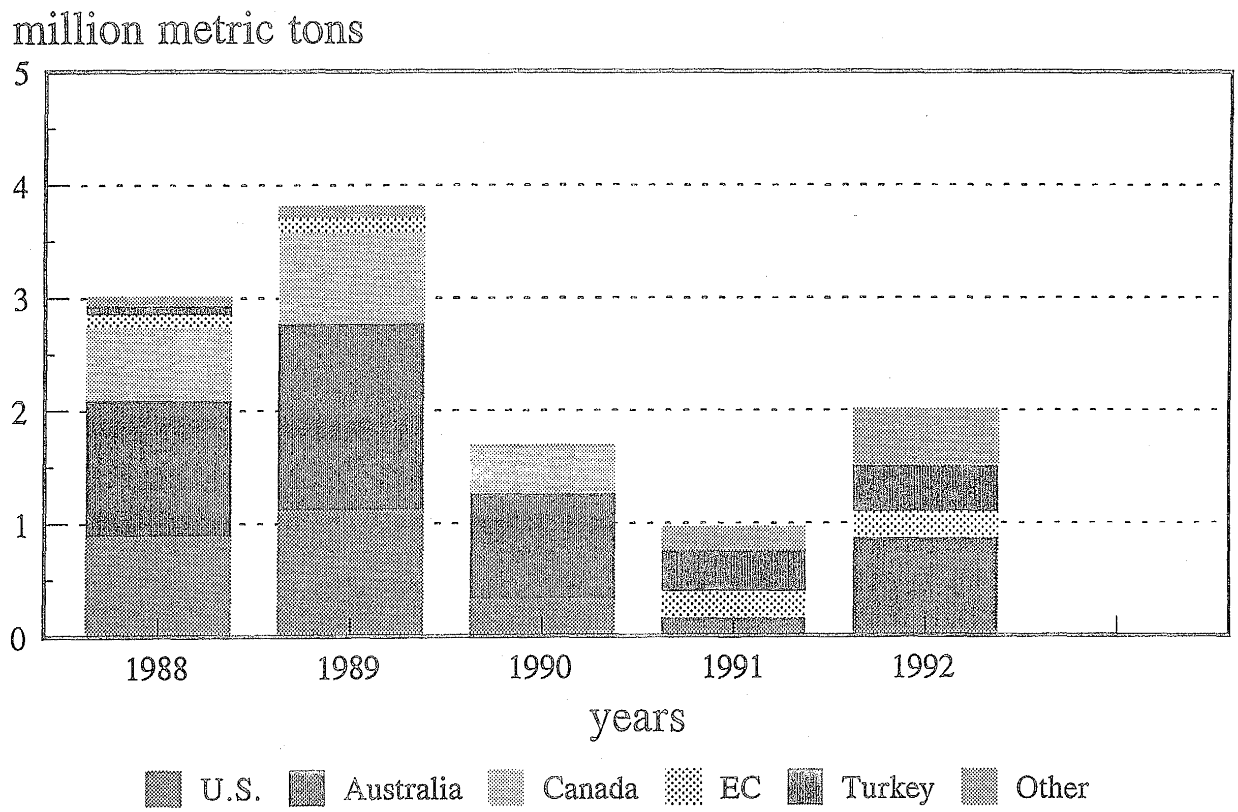


Figure 9

IRAQI GRAIN IMPORTS

BY TYPE, ANNUAL 1987-92

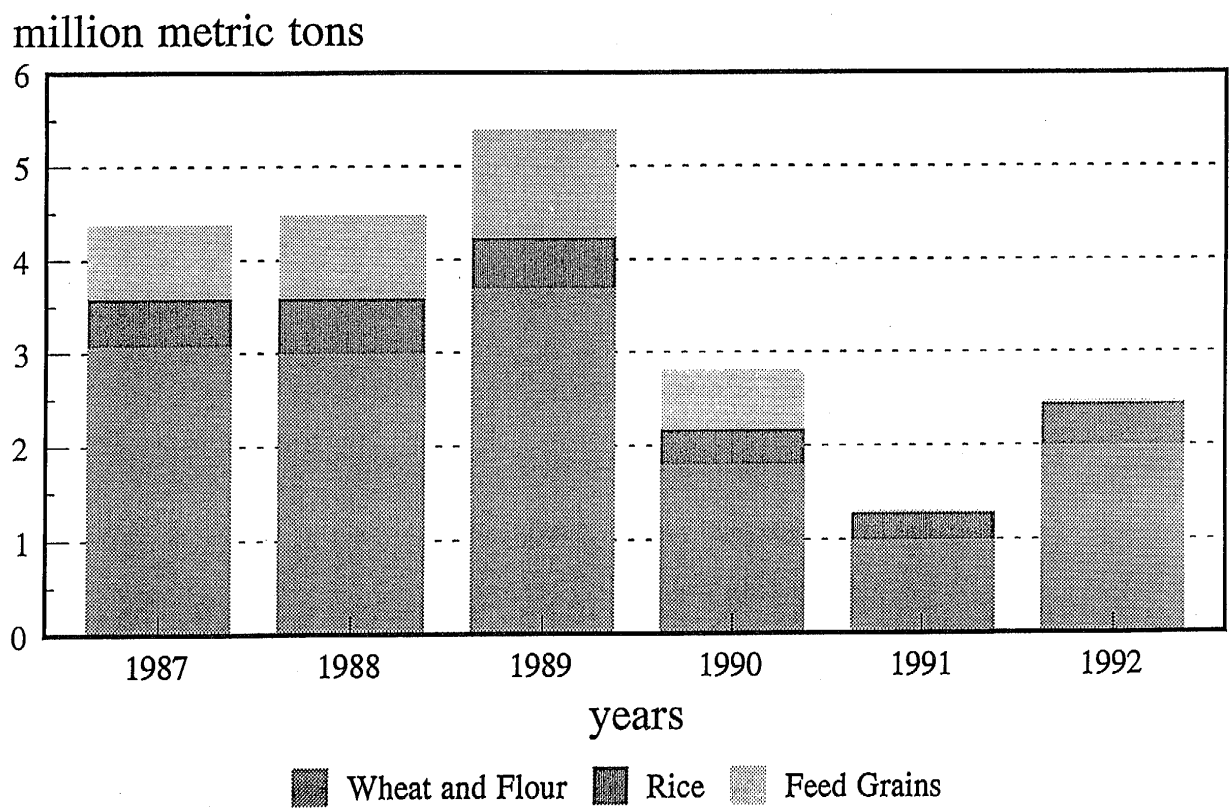


TABLE 1
SAUDI ARABIA: AGRICULTURAL IMPORTS BY COMMODITY, QUANTITY AND VALUE, ANNUAL 1987-92. 1989-92

COMMODITY	1989	1990	1991	1992	1989	1990	1991	1992
	metric tons				thousand dollars			
WHEAT	129737	127915	121906	173000	46917	45578	43644	90000
FLOUR OF WHEAT	7256	10276	22052	21000	1590	3170	7198	7400
MACARONI	22340	23777	21280	22300	19541	21973	21151	21151
PASTRY	41285	28854	29620	31200	73061	54977	60305	60305
MILLED PADDY RICE	524990	546770	532610	625000	146588	152166	147065	320000
BREAKFAST CEREALS	4976	4604	6427	6400	9818	10985	16595	16595
BARLEY	4611000	5622100	5243000	6356000	572000	576000	519000	680000
BEER OF BARLEY	28348	31431	42468	50000	16209	19846	28014	30000
MAIZE	839700	875000	925000	990000	89000	96000	117000	117000
OIL OF MAIZE	46000	60000	43495	56000	30000	44000	34101	38000
SORGHUM	9840	527	35000	20000	1951	210	7000	4500
POTATOES	135225	127074	129975	130000	24375	28717	33503	33500
FROZEN POTATOES	7500	9800	9500	8700	7700	9300	8400	8900
SUGAR REFINED	98560	101827	69935	100000	39349	46956	26587	37000
MOLASSES	2702	4615	8178	8178	1304	1173	2513	2513
SUGAR CONFECTIONERY	14435	12173	14936	12000	26063	23063	31018	28000
BEANS, DRY	2404	3138	2697	3000	1148	1789	1732	1600
BROAD BEANS, DRY	13395	16478	14700	14000	4141	4433	4700	6000
CHICK-PEAS	12503	10709	10720	12000	3778	4188	3815	5400
LENTILS	16249	16675	14037	16000	5068	7194	5677	8000
ALMONDS	1168	1057	1086	1086	2117	2073	2033	2033
PISTACHIOS	2614	3329	3382	3382	3811	4813	4642	4642
ALMONDS SHELLED	1541	1549	1494	1770	3757	3843	3004	4800
OIL OF SOYA BEANS	4011	4267	5134	5500	2549	2929	3947	4000
CAKE OF SOYA BEANS	248598	248500	262600	311700	79825	69497	60996	74000
GROUNDNUTS IN SHELL	1132	1015	1550	2000	506	677	1835	2000
PALM OIL	11925	53879	81598	159500	5522	17242	42908	90000
OLIVE OIL	5932	6641	6681	7000	10392	13483	13736	17000
OLIVES, PRESERVED	12310	10997	15811	14500	13101	10741	18859	17700
SESAME SEED	8203	16194	13105	15000	4626	16831	12021	13200
MELONSEED	7237	9294	8788	8788	1849	2303	3091	3091
CABBAGES	12849	11405	4673	6600	2330	2297	1586	2100
TOMATOES	143093	156165	136545	170000	23572	28899	32918	36000
TOMATO PASTE	25578	18902	10338	18000	21753	14727	8526	14500
ONIONS, DRY	154973	154665	146178	161000	26759	26939	27864	22000
GARLIC	8130	12681	10201	11900	3049	5696	5858	7000
VEGETABLES PREPARED NES	56420	59461	38579	43000	32286	35286	24994	27800
VEGETABLES FROZEN	7524	8364	6297	9100	6142	7810	6758	11100
BANANAS	114568	129332	124940	125800	31430	31412	47764	47000
ORANGES	197441	210803	192352	234000	58286	56732	54211	71000
TANG.MAND.CLEMENT.SATSMA	39884	42533	39652	39000	9286	10386	10694	10000
LEMONS AND LIMES	49742	54167	37356	46000	13615	14700	11664	18500

Saudi Arabian Agricultural imports continued

Table 1 continued

Saudi Arabian Agricultural imports continued

COMMODITY	1989	1990	1991	1992	1989	1990	1991	1992
	metric tons				thousand dollars			
(0515) APPLES	138989	134602	134588	110000	52026	44657	51787	48000
(0518) APPLEJUICE SINGLE STRENG	4917	3511	7728	7728	3384	2821	6406	6600
(0521) PEARS	13440	14382	13836	15000	5889	6001	5885	6000
(0534) PEACHES AND NECTARINES	25961	22925	9480	20000	8476	6345	3647	7600
(0560) GRAPES	31993	27812	27066	29900	16846	13038	14464	16000
(0561) RAISINS	5706	5170	4496	4400	5144	4995	3781	4100
(0567) WATERMELONS	15867	11086	9753	12000	4077	2980	3288	3500
(0568) CANTALOUPES OTH MELONS	10241	8790	13900	13900	2117	2292	6882	6882
(0571) MANGOES	3800	5557	7883	7883	2792	5450	8673	8673
(0575) PINEAPPLES, CANNED	9824	8204	6053	9600	8118	7951	4722	6100
(0577) DATES	8198	4346	300	100	4531	2588	180	60
(0583) MANGO JUICE	12356	10549	4956	4956	10219	8536	4661	4661
(0631) WATERS,ICE ETC	19874	15374	38084	38084	5964	4325	8585	8585
(0633) BEVERAGES NON-ALCOHOLIC	12839	8173	7916	7916	7402	5544	6343	6343
(0656) COFFEE, GREEN	6049	11761	16065	14000	18956	25905	37005	28000
(0666) CHOCOLATE PRODUCTS NES	16427	13821	21916	20000	45003	42533	69342	70000
(0667) TEA	9051	17890	19031	20000	30755	75865	74545	75000
(0702) NUTMEG, MACE, CARDAMONS	2755	6112	6678	7000	11753	32541	29991	32000
(0828) CIGARETTES	15949	18565	19362	19000	234345	261259	327434	343000
(0831) TOBACCO PRODUCTS NES	19190	15775	16369	16000	13594	12339	14488	14000
(0850) FEED SUPPLEMENTS	6775	12240	9547	16000	4262	7301	7955	12000
(0867) BEEF AND VEAL	15314	32534	47474	49500	38404	69638	88488	97000
(0874) SAUSAGES BEEF AND VEAL	1250	1381	5580	1500	2383	3224	14708	4500
(0882) COW MILK, WHOLE, FRESH	1669	1008	16108	14000	1421	1236	20392	23000
(0886) BUTTER OF COW MILK	16147	19097	17406	18000	33105	40439	36465	37000
(0894) WHOLE MILK, EVAPORATED	43522	44141	10662	8000	45580	46239	18395	14000
(0897) DRY WHOLE COW MILK	11958	13451	56218	61600	23570	27884	142546	165000
(0901) CHEESE (WHOLE COW MILK)	26733	30746	44483	43000	75015	91286	119523	115000
(0976) SHEEP	4930700	4775442	4618516	5100000	364711	341447	386435	410000
(0977) MUTTON AND LAMB	15640	22612	31159	36000	31072	45208	57088	70000
(1058) CHICKEN MEAT	194332	210076	244157	255000	229289	262482	333024	335000
(1182) HONEY	1450	838	2282	2700	3929	3301	6996	8000
OTHER	NA	NA	NA	NA	975959	1067751	1010383	1057666
(1882) AGRICULT PRODUCTS,TOTAL	NA	NA	NA	NA	3677496	3985395	4312343	4901000

Sources: Foreign Trade of Saudi Arabia, FAO and ERS matrix.

TABLE 2

SAUDI ARABIA: AVERAGE PRICE FOR SPECIFIED IMPORTS

	1989	1990	1991	1992
	dollars per metric ton			
(0015) WHEAT	361.63	356.31	358.01	520.23
(0016) FLOUR OF WHEAT	219.13	308.49	326.41	352.38
(0018) MACARONI	874.71	924.13	993.94	948.48
(0022) PASTRY	1769.67	1905.35	2035.96	1932.85
(0031) MILLED PADDY RICE	279.22	278.30	276.12	512.00
(0041) BREAKFAST CEREALS	1973.07	2385.97	2582.08	2592.97
(0044) BARLEY	124.05	102.45	98.99	106.99
(0051) BEER OF BARLEY	571.79	631.41	659.65	600.00
(0056) MAIZE	105.99	109.71	126.49	118.18
(0060) OIL OF MAIZE	652.17	733.33	784.02	678.57
(0083) SORGHUM	198.27	398.48	200.00	225.00
(0116) POTATOES	180.26	225.99	257.76	257.69
(0118) FROZEN POTATOES	1026.67	948.98	884.21	1022.99
(0164) SUGAR REFINED	399.24	461.14	380.17	370.00
(0165) MOLASSES	482.61	254.17	307.29	307.29
(0168) SUGAR CONFECTIONERY	1805.54	1894.60	2076.73	2333.33
(0176) BEANS, DRY	477.54	570.11	642.20	533.33
(0181) BROAD BEANS, DRY	309.15	269.03	319.73	428.57
(0191) CHICK-PEAS	302.17	391.07	355.88	450.00
(0201) LENTILS	311.90	431.42	404.43	500.00
(0221) ALMONDS	1812.50	1961.21	1872.01	1872.01
(0223) PISTACHIOS	1457.92	1445.78	1372.56	1372.56
(0231) ALMONDS SHELLED	2438.03	2480.96	2010.71	2711.86
(0237) OIL OF SOYA BEANS	635.50	686.43	768.80	727.27
(0238) CAKE OF SOYA BEANS	321.10	279.67	232.28	237.41
(0242) GROUNDNUTS IN SHELL	447.00	667.00	1183.87	1000.00
(0257) PALM OIL	463.06	320.01	525.85	564.26
(0261) OLIVE OIL	1751.85	2030.27	2055.98	2428.57
(0262) OLIVES, PRESERVED	1064.26	976.72	1192.78	1220.69
(0289) SESAME SEED	563.94	1039.34	917.28	880.00
(0299) MELONSEED	255.49	247.79	351.73	351.73
(0358) CABBAGES	181.34	201.40	339.40	318.18
(0388) TOMATOES	164.73	185.05	241.08	211.76
(0391) TOMATO PASTE	850.46	779.12	824.72	805.56
(0403) ONIONS, DRY	172.67	174.18	190.62	136.65
(0406) GARLIC	375.03	449.18	574.26	588.24
(0472) VEGETABLES PREPARED NES	572.24	593.43	647.87	646.51
(0473) VEGETABLES FROZEN	816.32	933.76	1073.21	1219.78
(0486) BANANAS	274.33	242.88	382.30	373.61
(0490) ORANGES	295.21	269.12	281.83	303.42
(0495) TANG.MAND.CLEMENT.SATSM	232.83	244.19	269.70	256.41
(0497) LEMONS AND LIMES	273.71	271.38	312.24	402.17

SAUDI ARABIA: AVERAGE PRICE

	1989	1990	1991	1992
	dollars per metric ton			
(0515) APPLES	374.32	331.77	384.78	436.36
(0518) APPLEJUICE SINGLE STREN	688.22	803.47	828.93	854.04
(0521) PEARS	438.17	417.26	425.34	400.00
(0534) PEACHES AND NECTARINES	326.49	276.77	384.70	380.00
(0560) GRAPES	526.55	468.79	534.40	535.12
(0561) RAISINS	901.51	966.15	840.97	931.82
(0567) WATERMELONS	256.95	268.81	337.13	291.67
(0568) CANTALOUPES OTH MELONS	206.72	260.75	495.11	495.11
(0571) MANGOES	734.74	980.75	1100.22	1100.22
(0575) PINEAPPLES, CANNED	826.34	969.16	780.11	635.42
(0577) DATES	552.70	595.49	600.00	600.00
(0583) MANGO JUICE	827.05	809.18	940.48	940.48
(0631) WATERS,ICE ETC	300.09	281.32	225.42	225.42
(0633) BEVERAGES NON-ALCOHOLIC	576.52	678.33	801.29	801.29
(0656) COFFEE, GREEN	3133.74	2202.62	2303.45	2000.00
(0666) CHOCOLATE PRODUCTS NES	2739.58	3077.42	3163.99	3500.00
(0667) TEA	3397.97	4240.64	3917.03	3750.00
(0702) NUTMEG, MACE, CARDAMONS	4266.06	5324.12	4491.02	4571.43
(0828) CIGARETTES	14693.40	14072.66	16911.17	18052.63
(0831) TOBACCO PRODUCTS NES	708.39	782.19	885.09	875.00
(0850) FEED SUPPLEMENTS	629.08	596.49	833.25	750.00
(0867) BEEF AND VEAL	2507.77	2140.47	1863.93	1959.60
(0874) SAUSAGES BEEF AND VEAL	1906.40	2334.54	2635.84	3000.00
(0882) COW MILK, WHOLE, FRESH	851.41	1226.19	1265.95	1642.86
(0886) BUTTER OF COW MILK	2050.23	2117.56	2094.97	2055.56
(0894) WHOLE MILK, EVAPORATED	1047.29	1047.53	1725.29	1750.00
(0897) DRY WHOLE COW MILK	1971.07	2073.01	2535.59	2678.57
(0901) CHEESE (WHOLE COW MILK)	2806.08	2969.04	2686.94	2674.42
(0976) SHEEP	73.97	71.50	83.67	80.39
(0977) MUTTON AND LAMB	1986.70	1999.29	1832.15	1944.44
(1058) CHICKEN MEAT	1179.88	1249.46	1363.97	1313.73
(1182) HONEY	2709.66	3939.14	3065.73	2962.96

TABLE 3. SAUDI ARABIA: GRAIN IMPORTS BY TYPE, ACCORDING TO SUPPLIELIERS

Year	WHEAT AND FLOUR (1,000 metric tons) U.S.	EC	Australia	Argentina	Other	TOTAL	RICE (1,000 metric tons) U.S.	Thailand	Pakistan	India	EC	Other	Total	CORN U.S.	Argentina	Thailand	Other	TOTAL
1980	338	406	113	0	134	992	224	43	94	18	1	6	386	21	0	133	9	163
1981	518	71	126	46	11	773	236	143	111	10	1	7	507	27	0	201	17	245
1982	480	83	208	32	2	805	313	99	82	45	0	7	547	31	0	318	12	361
1983	310	114	58	60	57	599	281	71	112	38	9	60	572	22	0	448	33	503
1984	290	64	111	11	4	479	268	84	56	63	7	22	500	144	5	377	74	600
1985	65	53	34	3	0	155	195	145	60	67	10	15	492	362	26	215	12	615
1986	108	23	5	0	93	229	193	184	55	56	4	12	504	195	82	322	18	617
1987	137	10	7	0	0	154	201	150	50	75	10	24	510	323	49	167	90	629
1988	140	12	5	0	3	160	193	148	55	63	11	40	510	587	12	83	18	700
1989	133	3	2	0	12	150	189	133	27	113	12	51	525	646	63	90	41	840
1990	156	4	2	0	13	175	190	148	23	105	2	79	547	781	15	24	55	875
1991	161	3	2	0	6	172	167	145	39	127	3	52	533	713	132	55	25	925
1992	160	23	2	0	9	194	209	184	55	150	2	25	625	497	296	70	127	990
1993	36	35	3	0	26	100	208	179	65	348	10	90	900	670	270	30	135	1,105

Year	BARLEY (1,000 metric tons) U.S.	EC	Canada	Australia	Other	TOTAL	SORGHUM (1,000 metric tons) Thailand	Sudan	Argentina	U.S.	Other	TOTAL	OTHER GRAINS	GRAIN IMPORTS MATRIX
1980	16	773	4	145	291	1,229	137	301	0	0	70	508	2	3,259
1981	59	1,722	0	335	159	2,275	107	242	0	0	4	354	3	4,145
1982	444	1,497	0	695	1,224	3,860	223	349	0	0	6	578	5	6,150
1983	206	865	70	270	474	1,885	137	252	0	0	17	406	9	3,975
1984	318	2,475	110	953	2,020	5,876	207	3	4	0	0	214	7	7,670
1985	42	2,145	361	2,888	264	5,700	140	1	0	0	7	148	8	7,049
1986	1,236	3,434	879	1,615	282	7,446	80	288	0	0	3	371	9	9,182
1987	2,345	2,641	2,103	197	290	7,576	13	286	0	0	2	301	11	9,170
1988	634	2,284	1,033	162	170	4,283	0	310	0	0	0	310	14	5,963
1989	715	2,134	1,558	117	140	4,664	0	275	0	9	1	285	15	6,464
1990	1207	3,040	1,155	653	20	6,075	0	10	0	0	5	15	17	7,687
1991	985	2,753	997	243	610	5,588	0	5	0	10	2	17	18	7,235
1992	669	3,998	289	379	698	6,033	0	4	41	7	5	57	20	7,899
1993	335	3,370	450	90	455	4,700	0	3	40	5	2	50	22	6,855

Sources: Foreign trade of Saudi Arabia and matrix of suppliers trade run

TABLE 4. IRAQ: GRAIN IMPORTS BY TYPE AS REPORTED BY SUPPLIERS, ANNUUAL 1980-93.

WHEAT AND FLOUR (thousand metric tons)

Year	U.S.	EC	Australia	Canada	Argentina	Turkey	Other	TOTAL
1980	312	23	787	429	0	120	32	1,703
1981	95	90	208	453	134	97	760	1,837
1982	178	128	860	259	280	200	252	2,156
1983	1,137	339	459	409	0	280	73	2,697
1984	1,064	206	1,290	492	0	311	70	3,432
1985	532	212	964	231	0	184	257	2,380
1986	816	70	980	552	0	0	240	2,658
1987	1,038	0	1,183	745	0	89	25	3,080
1988	895	122	1,200	640	0	78	68	3,003
1989	1,120	120	1,663	809	0	0	4	3,716
1990	335	0	930	423	0	0	112	1,800
1991	0	233	165	0	0	373	210	981
1992	0	221	879	0	0	412	512	2,024
1993	0	332	135	7	0	130	340	944

RICE (thousand metric tons)

Year	U.S.	Thailand	Pakistan	EC	Australia	VN	Other	Total
1980	269	53	26	0	0	0	0	348
1981	74	54	76	0	0	0	146	350
1982	235	129	0	0	0	0	56	420
1983	282	134	0	1	0	0	23	440
1984	448	38	0	1	0	0	9	495
								0
1985	407	113	0	0	0	0	5	525
1986	373	125	0	0	52	0	0	550
1987	509	15	0	0	0	0	0	524
1988	513	60	0	0	11	0	19	603
1989	392	120	0	0	28	0	2	542
1990	222	138	0	0	0	10	5	375
1991	0	160	20	0	20	37	50	287
1992	0	185	150	0	30	50	35	450
1993	0	318	25	0	36	145	176	700

BARLEY (thousand metric tons)

Year	U.S.	EC	Canada	Turkey	Other	TOTAL
1980	148	63	16	6	233	466
1981	0	78	0	37	115	230
1982	0	0	30	135	0	165
1983	125	0	129	36	0	290
1984	276	0	254	17	10	557
1985	100	0	132	29	0	261
1986	0	0	10	0	10	20
1987	67	35	111	6	1	220
1988	153	25	20	24	0	222
1989	187	60	283	0	40	570
1990	104	40	56	0	19	219
1991	0	0	0	5	40	45
1992	0	0	0	10	20	30
1993	0	0	0	0	0	0

CORN (thousand metric tons); OTHER GRAINS and GRAIN IMPORTS (thousand metric tons); From U.S. / Share U.S.

Year	Corn U.S.	Corn Thailand	Corn TOTAL	OTHER GRAINS	GRAIN IMPORTS	From U.S. Quantity	Share U.S. %
1980	0	83	83	1	2,601	730	30.8
1981	0	6	6	1	2,424	169	7.3
1982	0	70	70	1	2,812	412	14.7
1983	57	25	82	1	3,510	1600	45.6
1984	339	0	339	1	4,825	2127	44.1
					0		
1985	240	30	270	1	3,437	1279	37.2
1986	323	7	330	2	3,560	1512	42.5
1987	542	4	546	2	4,372	2156	49.3
1988	647	3	650	3	4,481	2208	49.3
1989	565	10	575	2	5,405	2264	41.9
1990	413	2	415	3	2,803	1074	38.3
1991	0	5	5	4	1,322	0	0.0
1992	0	4	4	5	2,513	0	0.0
1993	0	2	2	5	1,651	0	0.0

Sources: UN trade runs, U.S. Dept. of Commerce, FAO matrix
Sources: Grain and Feed Report of ATO Baghdad and matrix estimates

TABLE 5. IRAQ: AGRICULTURAL IMPORTS OF SPECIFIC AGRICULTURAL COMMODITIES, ACCORDING TO SUPPLIERS, 1989-92.

COMMODITY	1989	1990	1991	1992	1989	1990	1991	1992
	metric tons				thousand dollars			
(0015) WHEAT	3500000	1800000	500000	1100000	627000	342000	62000	165000
(0016) FLOUR OF WHEAT	101000	50000	237000	340000	25000	13000	49000	71000
RICE, MILLED	542000	375000	287000	450000	196000	220000	112000	181000
(0044) BARLEY	385362	210000	45000	30000	58254	35000	6500	3800
(1892) FEEDINGSTUFFS	257000	72000	58000	70000	117529	71000	44000	45200
(0049) MALT OF BARLEY	2000	9000	2500	6000	900	3600	850	2100
(0051) BEER OF BARLEY	14000	15000	5700	5000	14700	17000	4000	3500
(0056) MAIZE	570000	415000	800	1100	68500	52000	125	170
(0109) INFANT FOOD	6100	15000	3715	7800	18000	50000	10488	10488
(0116) POTATOES	31800	2500	140000	229000	10000	1600	28500	33000
(0164) SUGAR REFINED	656275	432400	166500	529000	283644	200000	73000	196000
(0191) CHICK-PEAS	28068	27000	10000	12000	14000	13000	6500	5000
(0201) LENTILS	20000	30000	19000	10000	16000	20000	11000	6000
(0211) PULSES NES	10000	10000	17000	28000	5341	6000	7900	9000
(0221) ALMONDS	1244	1244	400	900	4536	4536	1100	2300
(0238) CAKE OF SOYA BEANS	319500	101000	50	3300	77000	28000	20	1200
(0257) PALM OIL	257100	210000	21500	43100	133000	130000	13000	27000
(0258) OIL OF PALM KERNELS	10500	8100	300	450	7300	4500	220	350
(0268) OIL OF SUNFLOWER SEED	25400	22000	5700	3000	16300	14000	4500	2300
(0388) TOMATOES	25000	50000	57500	1400	7500	15300	15000	300
(0391) TOMATO PASTE	30000	25000	12257	14500	32000	24000	7624	7624
(0403) ONIONS, DRY	3500	400	65000	60000	800	200	12500	9400
(0472) VEGETABLES PREPARED NES	9000	6500	2000	2300	7800	5500	1400	1700
(0515) APPLES	10000	4500	88000	3300	5000	1800	23000	1200
(0561) RAISINS	2100	700	50	200	2900	1200	50	240
(0564) WINE	170	220	450	890	420	600	780	2300
(0653) FOOD WASTES	123000	60000	60000	67000	92000	44000	44000	44000
(0659) COFFEE EXTRACTS	300	200	200	230	2600	1800	1800	2100
(0666) CHOCOLATE PRODUCTS NES	450	1400	1340	1450	2000	4000	3000	3400
(0667) TEA	39973	32000	2600	1600	65185	70000	6500	1900
(0767) COTTON LINT	27259	50	50	11000	35605	58	64	25000
(0826) TOBACCO LEAVES	24108	13000	6000	5500	97552	55000	27000	24800
(0828) CIGARETTES	4000	5000	2500	3200	55000	75000	30000	35600
(0867) BEEF AND VEAL	78384	45000	18000	9800	232827	120000	21000	11600
(0886) BUTTER OF COW MILK	8942	3700	5600	12700	10492	5300	6000	13300
(1935) MILK COND EVAP	48700	39000	11000	3000	106918	115000	27000	6000
(0901) CHEESE (WHOLE COW MILK)	2313	10000	600	560	8353	32000	2100	1000
(0977) MUTTON AND LAMB	15581	10000	100	600	27477	14500	200	600
(1058) CHICKEN MEAT	5000	6000	500	2700	7000	8000	1000	1900
(1062) HEN EGGS	8080	6000	5000	6000	34797	21000	10000	11000
(1242) MARGARINE SHORTENING	28000	20000	57000	95000	26000	18000	43000	70000
OTHER	NA	NA	NA	NA	ERR	ERR	ERR	ERR
TOTAL AGRICULTURAL	NA	NA	NA	NA	2771000	1907900	772000	1096000

NA Not Applicable

Sources: FAO Agrostat, Trade Data of Suppliers, and ERS.

TABLE 6. IRAQ: ESTIMATED AVERAGE IMPORT PRICES, 1989-92

COMMODITY	1989	1990	1991	1992
		U.S. dollars per metric ton		
(0015) WHEAT	179.14	190.00	124.00	150.00
(0016) FLOUR OF WHEAT	247.52	260.00	206.75	208.82
RICE, MILLED	361.62	586.67	390.24	402.22
(0044) BARLEY	151.17	166.67	144.44	126.67
(1892) FEEDINGSTUFFS	457.31	986.11	758.62	645.71
(0049) MALT OF BARLEY	450.00	400.00	340.00	350.00
(0051) BEER OF BARLEY	1050.00	1133.33	701.75	700.00
(0056) MAIZE	120.18	125.30	156.25	154.55
(0109) INFANT FOOD	2950.82	3333.33	2823.15	1344.62
(0116) POTATOES	314.47	640.00	203.57	144.10
(0164) SUGAR REFINED	432.20	462.53	438.44	370.51
(0191) CHICK-PEAS	498.79	481.48	361.11	416.67
(0201) LENTILS	800.00	666.67	578.95	600.00
(0211) PULSES NES	534.10	600.00	464.71	321.43
(0221) ALMONDS	3646.30	3646.30	2750.00	2555.56
(0238) CAKE OF SOYA BEANS	241.00	277.23	400.00	363.64
(0257) PALM OIL	517.31	619.05	604.65	626.45
(0258) OIL OF PALM KERNELS	695.24	555.56	733.33	777.78
(0268) OIL OF SUNFLOWER SEED	641.73	636.36	789.47	766.67
(0388) TOMATOES	300.00	306.00	260.87	214.29
(0391) TOMATO PASTE	1066.67	960.00	622.01	525.79
(0403) ONIONS, DRY	228.57	500.00	192.31	156.67
(0472) VEGETABLES PREPARED NES	866.67	846.15	700.00	739.13
(0515) APPLES	500.00	400.00	261.36	363.64
(0561) RAISINS	1380.95	1714.29	1000.00	1200.00
(0564) WINE	2470.59	2727.27	1733.33	2584.27
(0653) FOOD WASTES	747.97	733.33	733.33	656.72
(0659) COFFEE EXTRACTS	8666.67	9000.00	9000.00	9130.43
(0666) CHOCOLATE PRODUCTS NES	4444.44	2857.14	2238.81	2344.83
(0667) TEA	1630.73	2187.50	2500.00	1187.50
(0767) COTTON LINT	1306.17	1160.00	1280.00	2272.73
(0826) TOBACCO LEAVES	4046.46	4230.77	4500.00	4509.09
(0828) CIGARETTES	13750.00	15000.00	12000.00	11125.00
(0867) BEEF AND VEAL	2970.34	2666.67	1166.67	1183.67
(0886) BUTTER OF COW MILK	1173.34	1432.43	1071.43	1047.24
(1936) MILK DRY	2195.44	2948.72	2454.55	2000.00
(0901) CHEESE (WHOLE COW MILK)	3611.33	3200.00	3500.00	1785.71
(0977) MUTTON AND LAMB	1763.49	1450.00	2000.00	1000.00
(1058) CHICKEN MEAT	1400.00	1333.33	2000.00	703.70
(1062) HEN EGGS	4306.56	3500.00	2000.00	1833.33
(1242) MARGARINE SHORTENING	928.57	900.00	754.39	736.84

Source: ERS calculations from trade matrix table.

References

"Country Data Sheets For Iraq "Food and Agricultural Organization, Rome, Italy, September 1993.

Foreign Trade of Saudi Arabia for 1991, Ministry of Finance and Economy, Central Statistics Department, Riyadh, November 1992.

"International Trade Matrix For Iraq", Economic Research Service, U.S. Dept. of Agriculture, Washington, DC. July 1993.

Exchange Rate Forecasts Using Eurocurrency Interest Rate Differentials

Alberto Jerardo, Economic Research Service

Abstract

The Eurocurrency market, in conjunction with the forward exchange market, represents the largest international pool of short-term funds tapped by arbitragers, hedgers, and speculators in financing and covering their currency positions. The forward covering of arbitrage positions in Eurocurrencies brings about covered interest rate parity. Empirical evidence of the correspondence of the forward premium and the interest rate differential between two currencies, especially when the U.S. dollar is involved, attests to the cogency of the interest rate parity theorem. The integration of the Eurocurrency market with the forward exchange market provides a framework in which interest rate parity can be exploited as a tool in generating exchange rate forecasts. In this paper, state space--a multivariate time series model, is employed in projecting interest rate differentials, which, by definition, estimate the forward premia. Taking the forward rate as the predictor of the future spot exchange rate, exchange rate forecasts are generated from the estimated forward premia. Notwithstanding satisfactory forecast results, the influence of interest differentials on exchange rate projections was not significant.

Keywords: Interest rate parity, exchange rate forecasts, Eurocurrency market, forward premium, state space, multivariate time series model.

Introduction

A country's export competitiveness depends to a significant extent on changes in currency exchange rates as they affect prices of traded goods and services. Prices, along with changes in income, in turn help determine the volume of exports and imports. While trade flows do influence exchange rates, their effect has increasingly been overshadowed by massive flows of funds between financial markets, which are driven principally by relative interest rates. Given current massive trade activity in short-term money assets, this paper attempts to forecast exchange rates from international interest rate differentials, specifically between Eurocurrency rates. The role of the forward exchange market in achieving covered interest rate parity (i.e., exchange risk is eliminated) is exploited in generating the forecasts.

Eurocurrency interest rate differentials largely determine the forward premium or discount that equates effective returns or costs--a broad characterization of interest parity. However, it is also the case that Eurocurrency rates adjust to changes in the forward premium (Herring and Marston, 1977). The joint determination of the forward exchange rate and the respective Eurocurrency rates is inherent in the way transactions are conducted--traders in one market base their quotations on rates established in the other market, and vise versa.[1] This interaction suggests the need for an integrated model, as opposed to independent models, of the Eurocurrency and forward exchange markets.

The linkage between forward exchange premia and international interest rate differentials brought about by covered interest arbitrage is the theoretical basis for the theory of international money market equilibrium. Interest rate parity--when the covered interest differential between convertible currencies is zero--provides the framework for simultaneous equilibrium in the forward and Eurocurrency markets. The linkage between these two markets was often not explicitly specified in past models of the forward market.[2] Interest parity provides a model for integrating these markets and for predicting exchange rates from interest rate differentials.

U.S. dollars, followed by Deutsche marks, and Japanese yen dominate the Eurocurrency and forward exchange markets. As such, the exchange rates for which forecasts are generated are the mark per dollar and yen per dollar. In general, a currency will achieve interest rate parity more quickly and with greater frequency the larger its external money and forward exchange markets.

---

[1] p. 80, R.J. Herring and R.C. Marston.

[2] pp. 56-57, 80, ibid.

Exchange Rate Forecasting Models

The three broad approaches to forecasting exchange rates use either structural models, arbitrage models, or pure prediction models. In structural models, spot or forward rates are derived from the relative size and direction of funds flowing between sources (supply) and users (demand). Alternatively, arbitrage models employ interest rate parity such that the (discounted) interest rate differential equals the anticipated percent change of the spot rate, as reflected in the forward premium if interest arbitrage is covered. That is, the forward premium depends on the size of the interest differential. Pure prediction models typically identify the underlying statistical structure of the spot rates after filtering out trend and seasonal patterns.

Whereas structural models show how underlying economic variables (including interest-rate differentials) jointly determine the market exchange rate, arbitrage models base forecasts on forward exchange rates, which is equivalent to adopting the interest parity relationship as the implicit reduced form of the true structural model (which is derived below). Arbitrage models, however, assume that interest rates are exogenously determined, at least independently of spot exchange rates, which is not consistent with Eurocurrency transactions where deposit or loan rates are jointly determined with forward rates.[3] A more realistic model for exchange rate determination can be constructed by making a Eurocurrency rate a function of the forward premium.

Time series models, which inherently cannot explain economic linkages, are nevertheless able to capture lag structures; and in the case of multivariate models, can correlate variables which are causally related. Estimating an arbitrage (interest parity) model by time series analysis will take care of the simultaneous bias from the linkage between Eurocurrency rates and the forward premium. Also, serial correlation in the residuals as well as multicollinearity problems are dealt with by the requirement of joint stationarity of the variables. This approach avoids potential econometric problems such as multicollinearity, simultaneity, and serial correlation in estimating a model that sets out to link forward rates and Eurocurrency rates--variables which are simultaneously determined and highly autocorrelated.

Interest Parity Theory of the Forward Exchange Rate

This theory postulates that equilibrium in the forward exchange market is reached when the covered return on a foreign asset is equal to the return on similar home assets--the forward premium is at interest parity with respect to home and foreign interest rates.[4] That is, if domestic and foreign assets, covered for exchange risk, are perfect substitutes, then interest arbitrage determines the forward premium. Risks such as capital controls, which are not generally covered in the forward market, are assumed to be zero. Exploiting interest rate differentials with the protection of forward cover (covered interest arbitrage), helps create, along with speculators, supply and demand for forward exchange that determines the forward premium.

The development and expansion of global financial markets under floating exchange rates have tied a country's monetary policy inextricably to its exchange rate policy. In the short run, the linkage between interest rates and exchange rates dominates the linkage between trade flows and exchange rates. When domestic interest rates rise, the home currency appreciates as capital flows in. However, the consequent loss in trade competitiveness usually takes some time to feed back into the exchange rate. In other words, changes in the real sector do not translate immediately into expectations about the currency's short-term value. As monetary policy instruments or targets, interest rates are more relevant for explaining short-term exchange rate movements than long-term trends such as trade flows or changes in income.

Floating exchange rates and highly mobile capital have also increased the interrelationship between international capital flows and interest rates and exchange rates. Decisions involved in holding financial assets, specifically those transacted on a covered basis, are based on both interest rate and forward exchange rate levels. Hence the direction of capital flows depends on relative interest rates as well as on the forward premium. Uncovered positions, on the other hand, may influence interest differentials when amounts are large enough, but will only affect the forward premium through any induced changes in the spot rate.

---

[3] op. cit.

[4] p. 60, ibid.

National money markets are subject to regulations that may inhibit interest arbitrage. In this case, forward exchange rates will diverge from interest parity with respect to national interest rates. In contrast, the Eurocurrency markets have been free of capital controls, and because of extensive interbank arbitrage, forward exchange rates quickly converge toward interest parity with respect to Eurocurrency rates. Tests have shown that interest parity holds with much greater accuracy and frequency between Eurocurrencies than between national money market deposits.[5] Figure 1 shows close adherence of the Deutsche mark's and the Japanese yen's forward premium (against the U.S. dollar) with the respective interest differential against the Eurodollar deposit rate.[6]

A model of the forward exchange market and the Eurocurrency market shall be constructed in the following sections with the objective that simultaneous market equilibrium results in the joint determination of equilibrium Eurocurrency interest rates and corresponding forward exchange rates. Covered interest arbitrage, in integrating the two markets, will equate the forward premium with the interest rate differential between two currencies. This relationship implies that once the interest differential is determined, the forward premium is determined as well, and conversely.

Interest Rate Parity in the Eurocurrency Market

Covered interest arbitrage is by nature transacted concurrently in the forward and Eurocurrency markets. An increase in demand for a particular Eurocurrency will affect the forward exchange rate, and an increase in demand for a particular forward currency will affect Eurocurrency rates. Equilibrium in both markets should be reached at the same time. But the interdependence of these markets has often been ignored in past studies, with conditions in one market generally assumed to be exogenously determined. To address this shortcoming, the structure of the exchange rate forecasting model will combine the two markets such that the forward premium and the Eurocurrency rates are simultaneously determined in the implicit reduced-form equation.

The Eurocurrency market can be described by a model in which all Eurocurrency assets (covered deposits and loans) are assumed to be perfect substitutes and perfectly mobile. It is further assumed that domestic interest rates are exogenously determined (by monetary authorities) so that arbitrage affects only Eurocurrency rates and the forward rate. Equilibrium in the Eurocurrency market must reflect the net (or excess) currency positions of all participants--Eurobanks as well as nonbank borrowers and investors.

The structure of the model is based on the equilibrium condition of equality between the net demand and net supply of Eurocurrency funds. The Euromark is used as the case currency, and Eurodollars as the foreign currency in formulating the model below, although a similar model for the yen can also be constructed. The demand for Euromark loans by nonbank borrowers is denoted by M^D, and M^S is the supply of Euromark deposits from nonbank investors. Eurobanks are not mere brokers, matching deposits and loans in each Eurocurrency. They typically take net loan or deposit positions in different currencies. That is, Euromark loans will equal the sum of Euromark deposits and the net Euromark position (NP) of the Eurobanks:

$$M^D = M^S + NP \tag{1}$$

These endogenous variables depend accordingly on the Euromark deposit rate (i_M), the Eurodollar rate ($i_\$$), and the forward premium (fp), in terms of marks per dollar:

$$\begin{aligned} M^D &= M^D(i_M^-,\, i_\$^+,\, fp^+) \\ M^S &= M^S(i_M^+,\, i_\$^-,\, fp^-) \\ NP &= NP(i_M^+,\, i_\$^-,\, fp^-) \end{aligned} \tag{2}$$

The demand for Euromark deposits is expected to increase in response to a rise in the forward premium since the cost of forward cover (in dollars) is lower. When banks cover their net Euromark positions in the forward

[5] p. 153, R.T. Baillie and P.C. McMahon (1989).

[6] The mean differential from 1980 to 1991 is not significantly different from zero at the 5 percent level (t test).

Eurocurrency Interest Rate Differentials
and Forward Premiums: Dollar and D-Mark

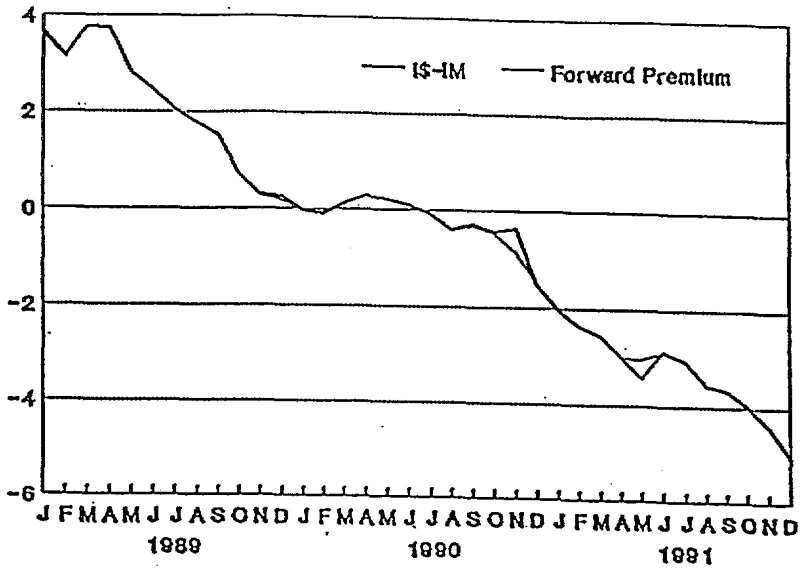


Figure 1a

Eurocurrency Interest Rate Differentials
and Forward Premiums: Dollar and Yen

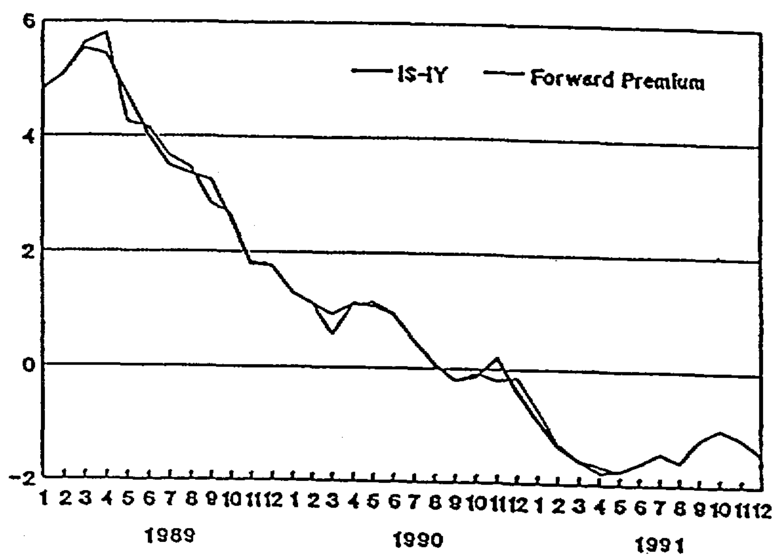


Figure 1b

market, they create a net demand (equal to NP) for forward dollars, the forward currency in this case. Competitive bidding leads to a point at which the net supply of forward exchange (FX) from speculators and hedgers must equal the Eurobanks' net demand for forward dollars. That is, at equilibrium,

$$FX = M^D - M^S \; (= NP) \tag{3}$$

Equation 3 can be expressed as an implicit function ($FX-M^D+M^S=0$), whose total differential equals zero:

$$(M_m^D - M_m^S)\, di_M + (M_\$^D - M_\$^S)\, di_\$ + (M_{fp}^D - M_{fp}^S)\, dfp = 0 \tag{4}$$

The subscripts indicate partial derivatives with respect to i_M, $i_\$$, and fp. Solving for the change in the forward premium, dfp, after dividing through by M_m^S:

$$dfp = \frac{\left(\frac{M_m^D}{M_m^S} - 1\right) di_M + \left(\frac{M_\$^D}{M_m^S} - \frac{M_\$^S}{M_m^S}\right) di_\$}{-\left(\frac{M_{fp}^D}{M_m^S} - \frac{M_{fp}^S}{M_m^S}\right)} \tag{5}$$

Under the condition of perfect capital mobility, where financial assets become perfect substitutes on a covered basis, these partial derivatives either approach infinity or negative infinity:

$$M_m^D = M_\$^S = M_{fp}^S = -\infty \quad and \quad M_\$^D = M_{fp}^D = M_m^S = \infty$$
$$Hence, \quad \frac{M_m^D}{M_m^S} = \frac{M_\$^S}{M_m^S} = \frac{M_{fp}^S}{M_m^S} = -1 \quad and \quad \frac{M_\$^D}{M_m^S} = \frac{M_{fp}^D}{M_m^S} = 1 \tag{6}$$

These conditions correspond to an infinitely elastic arbitrage schedule--a flat M^D curve in Figure 2 below, thus reducing the equilibrium equation to

$$dfp = di_M - di_\$ \tag{7}$$

which is the interest parity condition in differential form. This relationship represents all combinations of fp, i_M and $i_\$$ that will clear the forward market as well as the Eurocurrency market. Originally developed by Keynes, the interest-parity theory of forward exchange as described by Equation 7 holds that interest-rate gaps are always eventually matched by the forward premium or discount, making the covered interest differential zero.[7]

Market Equilibrium

The interest parity condition implies that changes in forward exchange rates have the same effect on capital flows as changes in the difference between Eurocurrency interest rates. Eurobanks offer loans and deposits at an interest rate of $i_M = i_\$ + fp$. Each combination of the Euromark and Eurodollar rates corresponds to a unique fp* in Figure 2 below, where M^S is infinitely elastic. Conversely, once fp* is determined, $i_M - i_\$$ is determined as well.

[7] John Maynard Keynes, A Tract on Monetary Reform (Macmillan, 1924), pp. 125-51. A more direct derivation of interest parity is described in Appendix A.

The degree of integration of the Eurocurrency and forward markets through covered interest arbitrage depends critically on the level of capital mobility. The speed of adjustment to new market interest and exchange rates is generally the measure of market efficiency, and capital mobility is required for these rates to clear the markets.[8] Any divergence of Eurocurrency rates or forward rates from parity should elicit instant realignment through new forward quotations, Eurocurrency quotations, or both.

Market equilibrium coincides with interest parity between Eurocurrencies as well as the equilibrium (or parity) level of the forward premium. In Figure 2, the equilibrium points in both the Eurocurrency and forward exchange markets are influenced by the level at which the net supply equals the net demand for forward currency, as well as by the degree of substitutability between domestic and foreign financial assets (which determines the slope of the net supply schedule for forward exchange, M^S).[9] Risks such as capital controls and other market imperfections that delay or limit the flow of funds cannot be covered in the forward market. They inhibit arbitrage by making covered forward assets less than perfect substitutes for domestic assets. In this situation, M^S is upward sloping.

Even in the absence of restrictions on capital flows, arbitrage does not bring domestic interest rates into relative parity. This is because national monetary policies bear heavily on domestic interest rates, and transaction costs between national money markets are contrastingly higher than in the Eurocurrency market. If there were perfect substitutability (M^S is flat), central banks will lose control over domestic interest rates and inevitably over capital flows. Since this is not the case, the forward premium adjusts to the differential between domestic and foreign interest rates as well as to the differential between the respective Eurocurrency rates. In Figure 2, the forward supply of D-marks simultaneously determines fp^* and i_M, given M^D and $i_\$$.

Forecasting Exchange Rates

Structural exchange rate models are usually specified by supply and demand equations, or their reduced form, representing the principal sources and users of funds of foreign exchange. Alternatively, exchange rate forecasts can be generated from observed prices of money market assets, which, if covered in the forward market, give the corresponding forward premium and interest rate. In efficient markets that reflect all available information, expectations about the future spot exchange rate are generally based on observed spot rates, forward rates, and interest rates. The Eurocurrency market, where the largest volume of (covered) money assets are intermediated internationally, represents an ideal structure for an exchange-rate forecasting model.

Because of the absence of capital controls, the Eurodollar and Euromark markets are tightly linked to domestic money markets in the U.S. and Germany. Consequently, deposit rates in those Eurocurrency markets correspond closely to domestic short-term rates for large deposits. Figure 3 below shows the close tracking of negotiable CD rates (for 3-month deposits) in the U.S. by the Eurodollar rate.

Domestic money markets in the United States and Germany are large in comparison to the size of the forward market for their currencies. An incremental change in supply or demand for forward dollars or marks will not ordinarily lead to a change in U.S.-German interest rate differentials. Rather, the change in forward supply or demand will usually be reflected in a corresponding change in the spot rate, leaving the forward premium at the unchanged interest-rate differential.

The interest parity relationship, supported by Figure 1, can be employed as the basis for forecasting exchange rates. Eurocurrency rates and the forward premium serve as the price variables that describe the market's structure. Models of the Eurodollar, Euromark, and Euroyen markets and their forward exchange markets, representing the bulk of global short-term (covered) financial flows, will provide the framework from which forecasts for each currency's exchange value can be generated.

State Space Model

Univariate time series estimation techniques--usually Box-Jenkins, do not distinguish between endogenous and exogenous variables. Multi-variable time series techniques, however, like econometric estimation, capture some

[8] Covered interest parity is said to be the only "unalloyed criterion" for capital mobility (p. 197, Frankel, J.A.).

[9] The Eurodollar rate is usually assumed as exogenously determined, since U.S. dollars make up about three-quarters of the Eurocurrency market. As such, the Eurodollar rate is influenced largely by U.S. monetary conditions.

Joint Equilibrium in the Forward and Eurocurrency Markets

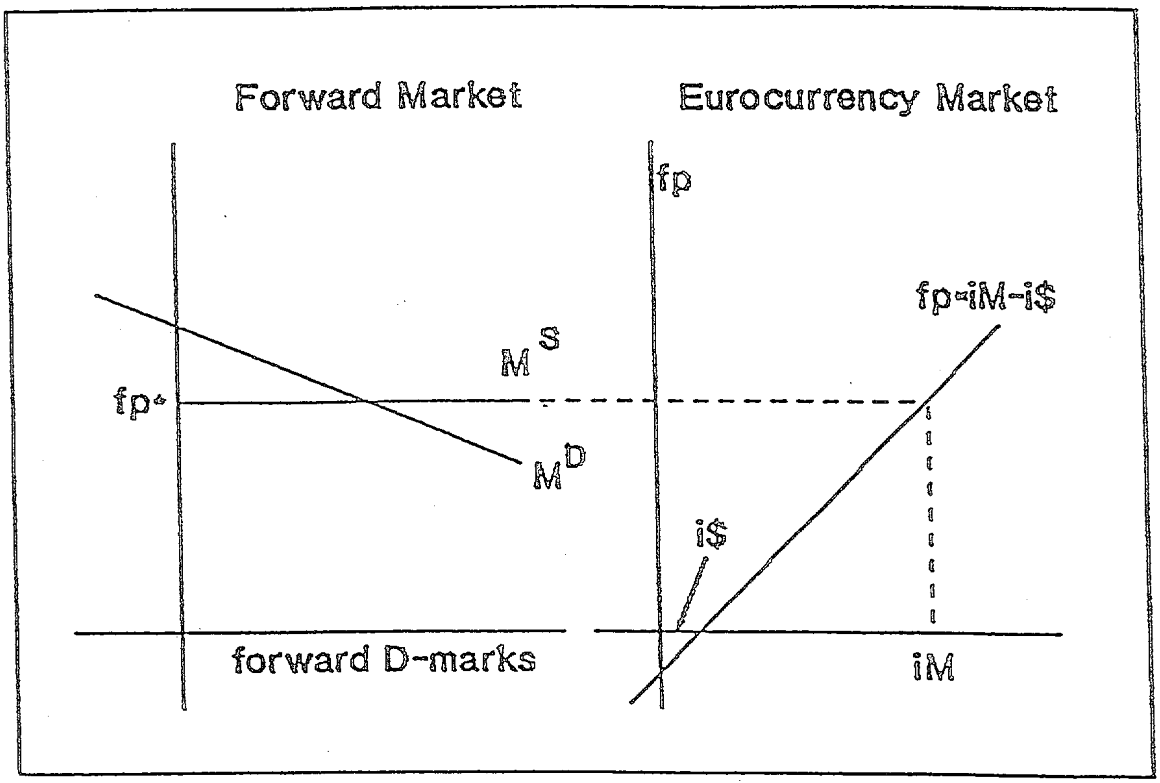


Figure 2

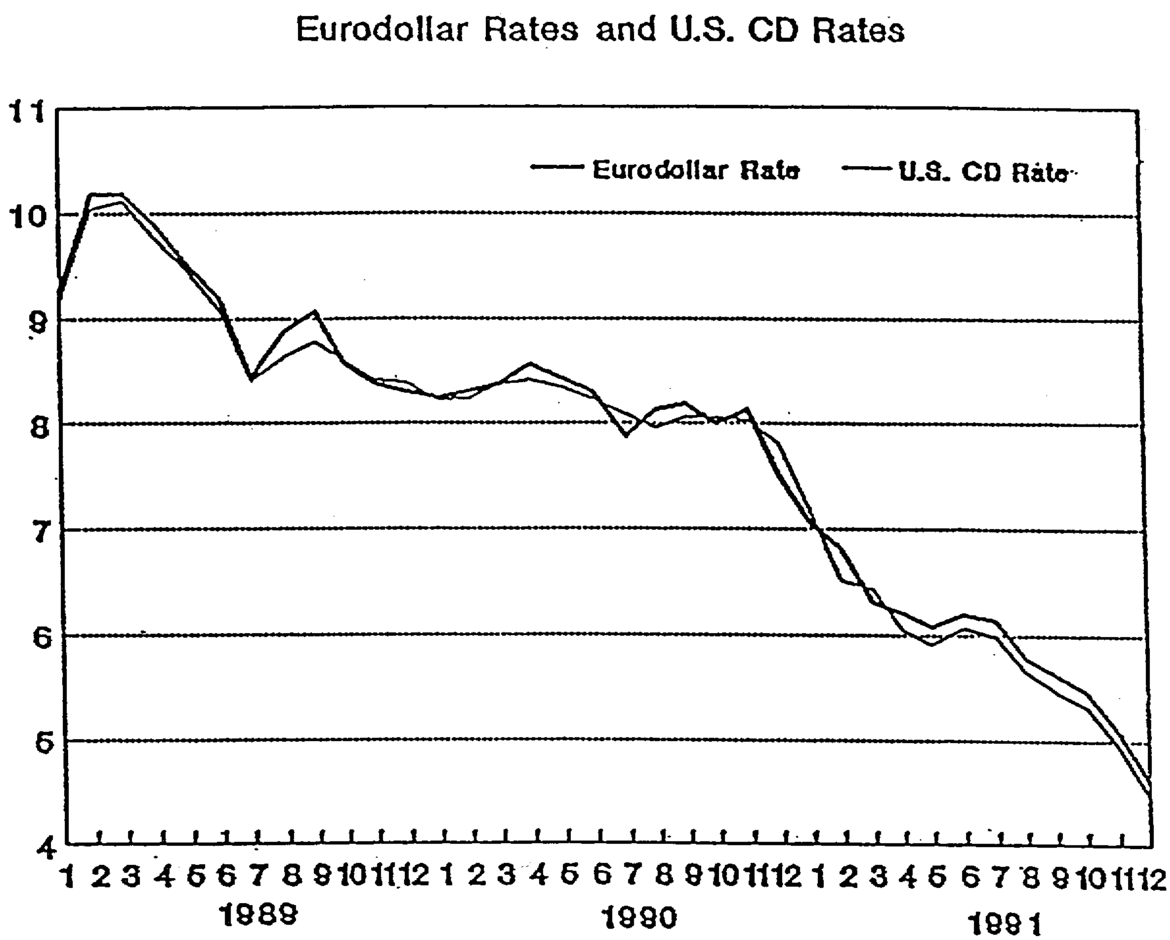

Eurodollar Rates and U.S. CD Rates
Eurodollar Rate
U.S. CD Rate
11
10
9
8
7
6
5
4
1 2 3 4 5 6 7 8 9 10 11 12 1 2 3 4 5 6 7 8 9 10 11 12 1 2 3 4 5 6 7 8 9 10 11 12
1989
1990
1991

Figure 3

causal relationship between the economic variables by differentiating between output and input variables. State space estimation, a generalized time series procedure, accounts for the effect of lead (input) variables on the output (dependent) variables.

A state space model is a multivariate time series forecasting model whose structure is made up of output and input equations in which values of the input variables are required in generating forecasts of the output variables. State space is a generalized form of Box-Jenkins and other statistical time series methods for estimating correlations and lag relationships in equations. As in Box-Jenkins, the output and input variables in the state space model must be stationary--i.e., have an unchanging mean, variance, and nonsignificant autocorrelations. Otherwise the residuals will not be white noise (when the autocorrelation function is not statistically significant for all lags).

A multivariate state space model is represented by two sets of linear equations:

Output equations: $\mathbf{y}_t = \mathbf{A}\mathbf{x}_t + \mathbf{B}\mathbf{w}_t + \mathbf{v}_t$

Input equations: $\mathbf{x}_t = \mathbf{C}\mathbf{x}_{t-1} + \mathbf{D}\mathbf{z}_t + \mathbf{E}\mathbf{u}_t$

where: $\mathbf{y}_t$ = vector of output (endogenous) variables to be predicted
$\mathbf{x}_t$ = vector of input variables needed to predict $\mathbf{y}_t$
$\mathbf{w}_t$, $\mathbf{z}_t$ = vector of explanatory (exogenous) variables
$\mathbf{v}_t$, $\mathbf{u}_t$ = uncorrelated error terms
A,B,C,D,E are parameter matrices

As the model's structure suggests, the estimated values of the input variables are substituted into the output equations in order to generate forecasts of the output variables. The input variables are taken as exogenous when plugged into the output equations. Error terms of the input equations are assumed to have an autocorrelation structure (E) that accounts for the autoregressive effect of the lagged dependent variable $\mathbf{x}_{t-1}$.

Model Specification

Covered interest parity between two Eurocurrencies implies that when one Eurocurrency rate and the forward premium are determined, the other Eurocurrency rate is determined as well. Since Eurobank liabilities are largely denominated in U.S. dollars, the Eurodollar rate and the forward premia (in terms of the dollar) would be the endogenous variables in the output equation of the state space model. Once they are determined, the Euromark and Euroyen rates would then equal the sum of the Eurodollar rate and the respective forward premia. That is, given that arbitrage between Eurobanks ensures interest parity with the Eurodollar rate:

Euromark/yen rate = Eurodollar rate + mark/yen forward premium

Euromark and Euroyen rates differ from the Eurodollar rate only by the cost of forward cover.

In the state space format outlined below, the Eurodollar rate and the forward premia are the output variables ($\mathbf{y}_t$). The Eurodollar rate is a function of the U.S. CD rate and the forward premia for the mark and yen. The forward premia in turn depend on the differential between respective deposit rates as well as the percentage change in relative export prices. The domestic deposit rates are the input variables ($\mathbf{x}_t$), specified as functions of the treasury bill rate and CPI inflation rate. The explanatory variables in each equation were chosen because they normally lead or predict movements in the corresponding dependent variables. While in practice the forward premia for the mark and yen are determined simultaneously with their Eurocurrency rates, in the absence of capital controls, they can alternatively be estimated from interest differentials between national money market rates. They are also influenced to some extent by relative export prices in which exchange risk generally requires traders to cover exposed currency positions through the forward market. The following equations define the model's structure:

Output equations:

Eurodollar rate = f{U.S. CD rate, mark/dollar forward premium, yen/dollar forward premium}

Mark/dollar forward premium = f{U.S. CD-German deposit rate differential, percent change in relative export prices}

Yen/dollar forward premium = f{U.S. CD-Japanese time deposit rate differential, percent change in relative export prices}

Input equations:

U.S. CD rate = f{U.S. treasury bill rate, U.S. CPI inflation rate}

German deposit rate = f{German treasury bill rate, German CPI inflation rate}

Japanese deposit rate = f{Japanese time deposit rate, Japanese CPI inflation rate}

In forecasting exchange rates, the forward rate F is generated from the ratio of the Euromark/yen yield to the Eurodollar yield multiplied by the base spot rate S (the last actual value in the sample period), which follows directly from the assumption of interest parity (see Appendix A). That is, after equating the Eurodollar yield to the covered Euromark yield, the forecasting equation for F is as follows:

$$F = S^*(1+i_M)/(1+i_\$)$$

$i_\$$ and i_M are the forecast values for the Eurodollar and Euromark rates. This relationship does not, however, imply any unidirectional causality. Rather, with the money and foreign exchange markets responding continuously to one another, the forward rate is determined simultaneously with the Eurocurrency rates (and the spot rate, which is assumed to be constant). Hence an exogenous disturbance causing a divergence from the parity condition induces simultaneous equilibrating adjustments in either the spot rate, the forward rate, or the two interest rates, or any combination thereof.

The model's specification depicts the linkage of national money markets to the international money market, with the presumption that cross-border capital movements are unrestricted. Domestic monetary conditions as represented in domestic interest rates are influenced by government monetary policy (through the treasury bill rate) and by price inflation. Since U.S. and German treasury bill rates and the Japanese time deposit rate normally lead their domestic deposit rates, they are used in place of the lagged domestic deposit rate to add explanatory power in the input equations. The model's recursive structure captures this progression of causality--from consumer and export prices to domestic interest rates and the forward premia, and then to Eurocurrency rates and the expected future spot rate, represented by F above.[10]

The order of estimation starts with the input equations for the national deposit rates, which are fed into the output equations to generate projections for the Eurodollar rate and the forward premia. The latter are then used to calculate the Euromark and Euroyen rates, whose covered returns when equated to the Eurodollar return provide forecasts for the forward exchange rate F in the forecasting equation. This process is equivalent to forecasting F, given the last actual spot rate, from projections of the interest differential, which equals the forward premium.

Projections of the exogenous variables in the input equations--Treasury bill rates and time deposit rate, inflation rates, and relative export prices--were generated by univariate Box-Jenkins. All short-term interest rates are monthly averages for 3-month deposits. The 10-year sample period covers January 1983 through December 1991, a total of 120 monthly observations; the forecast period is from January 1992 to June 1992.

[10] In contrast to the simultaneous estimation of parameters of the endogenous variables, a recursive structure sequentially estimates parameters.

Estimation Results

The equations were individually estimated and checked for autocorrelation in the residuals.[11] The requirement that variables be jointly stationary in each equation was relaxed as long as the Ljung-Box chi-squared statistic for error autocorrelation was statistically not significant. Large values of the Ljung-Box statistic indicate that the model does not fit the data. All dependent variables were identified as endogenous and explanatory variables as exogenous in the estimation procedure. The estimated parameters and test statistics of each equation are shown in Appendix B.

The Box-Jenkins models chosen for the inflation rate and relative export prices were poor fits, with large root mean square errors, but have nonsignificant Ljung-Box chi square statistics (the test for serial correlation). The forecasts generated by these univariate models generally exhibited constant patterns. This lack of variability in the forecast values of the exogenous variables should, however, produce more stable (or linear) behavior in projections for the endogenous variables. State space forecasts of the domestic deposit rates show linear trends with minuscule slopes. The Japanese deposit rate exhibited the least variability, which can be attributed to the relatively insulated money market in Japan. Japan's financial markets are not completely deregulated.

The root mean square forecast error for the Eurodollar rate was relatively small in comparison to the forward premia and the domestic deposit rates. This is because Eurodollars dominate the Eurocurrency market, and therefore are less vulnerable to external forces than other currencies. While the respective contributions of the forward premia with respect to the D-mark and the yen were statistically significant in explaining the variability of the Eurodollar rate, they paled in comparison to the U.S. CD rate's effect. The forward premium equations also show a significantly greater sensitivity to the domestic interest rate differentials than to relative export prices. Similarly, in the input equations, the domestic deposit rates responded more to the treasury bill rates than to consumer price inflation.

Evaluation of Forecasts

The out-of-sample projections for the exogenous (inflation rates and relative export prices) and input (domestic deposit rates and forward premia) variables are shown in Appendix C. The resulting Eurocurrency rate forecasts, when expressed as differentials, give the following exchange rate forecasts of the dollar's value in terms of the mark and the yen:

a. German marks per dollar:

1992	diff. of interest rate forecasts	diff. of actual interest rates	exchange rate forecast (DM/$)	actual exchange rate (DM/$)	absolute forecast error
January	5.6836	5.355	1.647	1.579	4.3%
February	5.7879	5.457	1.650	1.619	1.9%
March	5.8600	5.353	1.651	1.662	0.66%
April	5.9089	5.486	1.652	1.649	0.18%
May	5.9417	5.955	1.652	1.623	1.8%
June	5.9632	5.771	1.653	1.573	5.1%

Average forecast error: 2.32%

[11] The multivariate state space procedure of FORECAST MASTER PLUS (Version 1.01), a statistical forecasting package, was used in estimating the parameters of the equations.

b. Japanese yen per dollar:

1992	diff. of interest rate forecasts	diff. of actual interest rates	exchange rate forecasts (yen/$)	actual exchange rate (yen/$)	absolute forecast error
January	-1.089	1.074	126.63	125.46	0.93%
February	-1.100	1.077	126.69	127.70	0.79%
March	-1.104	0.628	126.68	132.86	4.65%
April	-1.106	0.458	126.74	133.54	5.09%
May	-1.107	0.779	126.68	130.77	3.13%
June	-1.107	0.635	126.68	126.84	0.13%

Average forecast error: 2.45%

The relatively low average forecast errors for both exchange rates are encouraging despite the narrow range of the exchange rate projections. The smaller interest differentials between Euroyen and Eurodollars accounted for the dominant influence of the base spot rate used in calculating the yen per dollar forward rate (the implicit future spot rate). This becomes more apparent by noting the large gap between the actual and projected interest rate differentials. To further test the accuracy of these forecasts, actual values of the explanatory variables in the forecast period were substituted instead into the output equations. Both the mark per dollar and yen per dollar ex post projections produced a 2.3 percent average forecast error (projections of variables are in Appendix C). Thus the out-of-sample forecasts in which predicted values for the explanatory variables were individually generated compare favorably with the ex post projections. These satisfactory results demonstrate the predictive ability of the interest parity theorem and the applicability of the state space model. The forward rate in this case provided an acceptably accurate predictor for the future spot exchange rate.

One advantage of a multivariate time series model over a univariate model derives from the ability to simulate various assumptions about domestic interest rates, inflation rates, or exchange rate policies and observe the impact on the output variables. When joint stationarity holds and residuals are white noise, econometric complications from multicollinearity, serial correlation, and simultaneity are avoided. For example, lagged dependent variables and serial correlation would have posed a serious problem had an econometric estimation method been applied instead. The state space model of relating Eurocurrency rates to lead variables lends itself uniquely to forecasting.

Summary and Conclusions

The role of forward exchange in covering short-term Eurocurrency deposits and loans provides a means for determining the expected future spot rate. Since the forward premium is the foreign exchange market's measure of a currency's expected appreciation, it can be used in predicting the future spot rate. By invoking the interest parity relationship, forecasts of Eurocurrency rates can be derived from estimates of the forward premia and the Eurodollar rate. Interest arbitrage leads to an equilibrium condition in which covered yields between Eurocurrencies converge, and where the forward rate represents the expected future spot rate. The Eurocurrency rates are determined as soon as the forward premia are determined. This way of forecasting exchange rates was suitably accommodated in the state space model for time series analysis. A recursive structure was constructed for the purpose of projecting Eurocurrency interest rates from forward premia. While satisfactory exchange rate forecasts were obtained from this model, the base spot rate was highly influential in determining the forecast values, overshadowing the effect of the interest differential.

This approach to forecasting exchange rates implicitly makes exchange rates a function of price variables--relative interest rates, inflation rates, and export prices, as opposed to underlying market demand and supply variables. That is, domestic monetary and price conditions are reflected in the external market by their transmission through domestic interest rates to the forward premia. The assumptions that changes in domestic interest rates lead (by one month) changes in Eurocurrency rates and that price inflation affects domestic interest rates were built into the recursive structure of the model. Also, the forward premium is linked to changes in relative export prices and projected domestic interest differentials. Univariate Box-Jenkins forecasts of treasury bill rates, CPI inflation rates, and changes in the relative export price indexes were used to generate out-of-sample projections for the exogenous variables. The forecast results were relatively accurate because the actual exchange rates did not significantly diverge from the base exchange value. The marginal contribution to the forecast values of the interest differentials was beneficial in the case of the yen, where the discrepancy between the actual and projected interest differentials was particularly wide. These results, however, point out the model's vulnerability to exchange market volatility, especially when speculative activity is high. Had the exchange rate projections been extended to the second half of 1992 when the European Monetary System unraveled, the model's predictive ability would have suffered. But then, when projections based on one initial exchange value are made beyond a few months, the base exchange rate's influence always tends to recede.

The assumption that the forward rate represents the best guess of the future spot rate, as was made here, stems simply from the way the forward premium is computed--as the percentage difference over the spot rate. A forward rate bias is introduced by this assumption only if over time the average forecast error between the forward rate and the future spot rate is not zero. Similarly, if the forward premium and the interest differential consistently vary from each other, the forward rate would likely be a biased predictor of the future spot rate. While the simultaneous interaction of the forward premium and the interest differential due to the interdependence of the Eurocurrency and forward exchange markets may introduce a simultaneity, or endogeneity, bias in the forecast values of the Eurocurrency rates, taking differentials would minimize it. The unpredictable relationship between short-term interest rates and price inflation because of active government intervention in the money markets compounds some lead-lag assumptions in the model. Fortunately, if stationarity is achieved between these variables, state space estimation generates forecasts that are statistically sound. Last, not least, by purposely excluding exchange rates from the model, the problem of joint stationarity was avoided since exchange rates follow a random walk, thus nonstationary, process. Moreover, the inability by previous studies to find cointegration--a linear combination that is stationary--between nominal exchange rates and relative money supplies suggests that any structural linkage of exchange rates and interest rates, other than indirectly through the interest parity theorem, would not have been feasible in a time series model (see p. 548, Adams and Chadha).

From the premise that hedged and nonhedged holders of forward contracts on foreign-currency assets conduct their bidding by equating real rates of interest (adjusted for price inflation), exchange rates can also be predicted from inflation rate differentials (see p.891, Giddy). In practice, however, real interest rates tend to equalize only in the long run and observed exchange rates differ only partly because of price disparities. The forward premium, on the other hand, reflects the broadest range of market expectations, including inflation, and hence is a superior indicator of currency preference.

Appendix A: Covered Interest Parity Equation

The parity equation can also be derived by comparing expected returns between domestic and foreign investments. This simpler model depends directly on the price variables as opposed to structural variables. Under the same conditions of perfect competition and free currency convertibility, domestic and foreign yields (i_h and i_f) will tend to converge in equilibrium. Assuming continuous compounding of interest,

$$e^{i_h t} = e^{i_f t}\frac{r_t}{r_0}$$

$$\frac{r_t}{r_0} = (i_h - i_f)t$$

r_t is the forward rate of exchange for delivery at time t (in fractions of a year), and r_0 is the spot rate at time 0. Expressing the forward discount on domestic currency in percent per annum,

$$fp = \frac{r_t - r_0}{r_0 t}$$

$$\frac{r_t}{r_0} = 1 + fp*t$$

Therefore,

$$1 + fp*t = e^{(i_h - i_f)t}$$

After taking logs, the first-order approximation with respect to t gives

$$fp = i_h - i_f$$

which is the familiar equivalence between the interest differential and the cost of the forward contract. Market equilibrium is reached when the forward premium equalizes effective interest rates.

Hence in equating 1-year yields between Eurodollar and Euromark deposits, $i_\$$ and i_M,

$$1 + i_\$ = (1 + i_M)*\frac{r_t}{r_0}$$

which, if i_M is small, can be simplified to

$$i_\$ = i_M + fp$$

Appendix B: Estimation Statistics

A. State Space Diagnostics of Output Equations:

1. Eurodollar rate
   - Transformation for stationarity: differencing order=1; power=log
   - U.S. CD rate coefficient: 0.8789
   - D-mark/$ forward premium coefficient: 0.0052
   - Yen/$ forward premium coefficient: -0.0003
   - Adjusted R-square: 0.995
   - RMS (root mean square) error: 0.014
   - Ljung-Box chi square: 17.72 (0.526)
   - Schwarz criterion (BIC): 0.127
   - Error autocorrelation function: no significant autocorrelation (except at lag 1)

2. Mark/dollar forward premium
   - Transformation for stationarity: degree of differencing=1
   - Deposit rate difference coefficient: -.0989
   - (transformation for stationarity: degree of differencing=1; power=log)
   - Percent change in relative export prices coefficient: .007
   - Adjusted R square: 0.965
   - RMS error: 0.471
   - Ljung-Box chi square: 12.87 (0.201)
   - Schwarz criterion (BIC): 0.514
   - Error autocorrelation function: no significant autocorrelations

3. Yen/dollar forward premium
   - Transformation for stationarity: degree of differencing=1; power=inverse
   - Deposit rate difference coefficient: 0.19
   - (transformation for stationarity: degree of differencing=1; power=inverse)
   - Percent change in relative export prices coefficient: -.064
   - Adjusted R square: 0.00
   - RMS error: 4.757
   - Ljung-Box chi square: 1.98 (0.00)
   - Schwarz criterion (BIC): 24.946
   - Error autocorrelation function: no significant autocorrelations

B. State Space Diagnostics of Input Equations:

1. Domestic deposit rates

a) United States
   - transformation for stationarity: degree of differencing=1; power=log
   - Treasury bill rate coefficient: 7.102
   - CPI inflation rate coefficient: 0.0015
   - Adjusted R square: 0.980
   - RMS error: 0.027
   - Ljung-Box chi square: 26.78 (0.917)
   - Schwarz criterion (BIC): 0.2283
   - Error autocorrelation function: significant autocorrelations at lags 3, 11 (at 5% level)

b) Germany
   - Transformation for stationarity: degree of differencing=1; power=inverse
   - Treasury bill rate coefficient: 0.353

CPI inflation rate coefficient: 0.0018
Adjusted R square: 0.989
RMS error: 0.007
Ljung-Box chi square: 22.08 (0.771)
Schwarz criterion (BIC): 0.1724
Error autocorrelation function: no significant autocorrelation

c) Japan

Transformation for stationarity: degree of differencing=1
Treasury bill rate coefficient: 0.5806
CPI inflation rate coefficient: 0.00455
Adjusted R square: 0.992
RMS error: 0.123
Ljung-Box chi square: 24.95 (0.874)
Schwarz criterion (BIC): 0.134
Error autocorrelation function: significant autocorrelation at lag 3 (at 5% level)

C. Box-Jenkins Diagnostics of Exogenous Variables

1. Treasury bill rates

a) United States

Transformation for Stationarity: differencing order=1; power=inverse
first-order autoregressive coefficient: 0.493 (t statistic=5.743)
Adjusted R square: 0.967
RMS error: 0.0057
Ljung-Box chi square: 9.28 (0.047)
Schwarz criterion (BIC): 0.312
Error autocorrelation function: no significant autocorrelation

b) Germany

Transformation for stationarity: differencing order=1; power=inverse
first-order moving average coefficient: -0.2224 (t statistic=-2.489)
Adjusted R square: 0.961
RMS error: 0.0134
Ljung-Box chi square: 13.63 (0.247)
Schwarz criterion (BIC): 0.4076
Error autocorrelation function: no significant autocorrelation

c) Japan

Transformation for stationarity: differencing order=1
First-order moving average coefficient: -0.459 (t statistic=-5.607)
Adjusted R square: 0.971
RMS error: 0.2364
Ljung-Box chi square: 12.17 (0.162)
Schwarz criterion (BIC): 0.241
Error autocorrelation function: no significant autocorrelation

2. Consumer price index inflation rates

a) United States

Transformation for stationarity=none
Constant: 0.195
First-order moving average coefficient: -0.556

First-order seasonal autoregressive coefficient: 0.393 (t statistic=4.688)
Adjusted R square: 0.296
RMS error: 0.216
Ljung-Box chi square: 24.52 (0.861)
Schwarz criterion (BIC): 0.2247
Error autocorrelation function: no significant autocorrelation

b) Germany

Transformation for stationarity: seasonal differencing order=1
First-order autoregressive coefficient: 0.447 (t stat.=4.677)
First-order seasonal moving average coefficient: 0.844 (t statistic=21.173)
Adjusted R square: 0.299
RMS error: 0.203
Ljung-Box chi square: 2.548 (0.827)
Schwarz criterion (BIC): 0.2115
Error autocorrelation function: significant autocorrelation at lag 9 (at 5% level)

c) Japan

Transformation for stationarity: order of differencing=1; power=inverse
First-order moving average coefficient: 0.986 (t statistic=101.789)
Adjusted R square: 0.00
RMS error=2.50
Ljung-Box chi square: 15.42 (0.367)
Schwarz criterion (BIC): 0.9347
Error autocorrelation function: significant autocorrelation at lag 10 (at 5% level)

3. Change in relative export prices

a) Germany/United States

Transformation for stationarity: degree of differencing=1; power=log
First-order moving average coefficient: 0.6667 (t statistic=3.894)
Second-order moving average coefficient: 0.2698 (t statistic=3.118)
Adjusted R square: 0.00
RMS error: 0.808
Ljung-Box chi square: 11.90 (0.146)
Schwarz criterion (BIC): 1.2006
Error autocorrelation function: no significant autocorrelation

b) Japan/United States

Transformation for stationarity: degree of differencing=1; power=log
First-order moving average coefficient: 0.9218 (t statistic=25.424)
Adjusted R square: 0.00
RMS error: 0.805
Ljung-Box chi square: 10.19 (0.074)
Schwarz criterion (BIC): 1.0159
Error autocorrelation function: no significant autocorrelation

Appendix C: Projections

A. Out-of-Sample Forecasts

1. Input variables:

a) Treasury bill rates

		U.S.	Germany	Japan
1992	Jan	3.9185	9.1172	5.9903
	Feb	3.8263	9.1172	5.9903
	Mar	3.7824	9.1172	5.9903
	Apr	3.7612	9.1172	5.9903
	May	3.7508	9.1172	5.9903
	Jun	3.7457	9.1172	5.9903

b) Consumer price index inflation rates

		U.S.	Germany	Japan
1992	Jan	0.3196	0.5482	0.4505
	Feb	0.2577	0.2320	0.4505
	Mar	0.2577	-0.0016	0.4505
	Apr	0.2574	0.2798	0.4505
	May	0.2888	0.2175	0.4505
	Jun	0.3195	0.2746	0.4505

c) Domestic deposit rates

		U.S.	Germany	Japan
1992	Jan	4.1480	7.9632	5.9444
	Feb	4.0007	7.9588	5.9258
	Mar	3.9305	7.9688	5.9213
	Apr	3.8972	7.9490	5.9202
	May	3.8817	7.9315	5.9199
	Jun	3.8750	7.9005	5.9198

2. Output variables:

a) Percent change in relative export prices (relative to the U.S.)

		Germany	Japan
1992	Jan	2.1441	1.3922
	Feb	1.5792	1.3922
	Mar	1.5792	1.3922
	Apr	1.5792	1.3922
	May	1.5792	1.3922
	Jun	1.5792	1.3922

b) Forward premium (in terms of the U.S. dollar)

		Germany	Japan
1992	Jan	5.6836	-1.0891
	Feb	5.7879	-1.0996

Mar	5.8600	-1.1037
Apr	5.9089	-1.1057
May	5.9417	-1.1066
Jun	5.9632	-1.1070

c) Eurocurrency deposit rates

		U.S.	Germany	Japan
1992	Jan	4.3071	9.9907	3.218
	Feb	4.1597	9.9476	3.060
	Mar	4.0886	9.9486	2.985
	Apr	4.0545	9.9634	2.949
	May	4.0383	9.9800	2.932
	Jun	4.0311	9.9943	2.924

B. In-sample forecasts for the Eurocurrency rates (actual data provided for input variables--domestic deposit rates and forward premia):

		U.S.	Germany	Japan
1992	Jan	4.2210	9.6573	5.1677
	Feb	4.2279	9.6329	5.2616
	Mar	4.3918	9.6563	4.9978
	Apr	4.1613	9.6967	4.7789
	May	3.9846	9.7217	4.7999
	Jun	4.0143	9.7218	4.6705

Base spot exchange rate (last actual value of sample period):

Dec. 1991: D-mark/dollar = 1.563
Yen/dollar = 128.04

References

1. Abraham, Bovas and Johannes Ledolter. Statistical Methods for Forecasting. John Wiley, 1983.

2. Adams, Charles and Bankim Chadha, "Structural Models of the Dollar," IMF Staff Papers, vol. 38, no. 3, pp. 525-59.

3. Baillie, Richard T. and Patrick C. McMahon. The Foreign Exchange Market: Theory and Econometric Evidence. Cambridge University Press, 1989.

4. Clendenning, E. Wayne. The Euro-Dollar Market. Oxford University Press, 1970.

5. Cumby, Robert E. and Maurice Obstfeld, "A Note on Exchange-Rate Expectations and Nominal Interest Differentials: A Test of the Fisher Hypothesis," The Journal of Finance, June 1981, pp. 697-703.

6. de Grauwe, Paul. International Money. Oxford University Press, 1989.

7. Dufey, Gunter and Ian H. Giddy. The International Money Market. Prentice-Hall, Inc., 1978.

8. Frankel, Jeffrey A., "Measuring International Capital Mobility: A Review," Papers and Proceedings of the American Economic Association, January 1992, pp. 197-202.

9. Gibson, Heather D. The Eurocurrency Markets, Domestic Financial Policy and International Instability. St. Martin's Press, 1989.

10. Giddy, Ian H., "An Integrated Theory of Exchange Rate Equilibrium," Journal of Financial and Quantitative Analysis, Dec. 1976, pp. 883-92.

11. Goodrich, Robert L. Applied Statistical Forecasting. Business Forecast Systems, Inc., 1989.

12. Goodrich, Robert L. and Eric A. Stellwagen. Forecast Master Plus. Documentation Version 1.01, Business Forecast Systems, Inc., 1983.

13. Granger, C.W.J. and Paul Newbold. Forecasting Economic Time Series, 2nd ed. Academic Press, Inc., 1986.

14. Herring, Richard J., "The Interbank Market," Eurodollars and International Banking, Paolo Savona and George Sutija, eds. MacMillan, 1985.

15. Herring, Richard J. and Richard C. Marston, "The Forward Market and Interest Rates in the Eurocurrency and National Money Markets," Eurocurrencies and the International Monetary System, Carl H. Stem, John M. Makin and Dennis E. Logue, eds., American Enterprise Institute, 1976.

16. Herring, Richard J. and Richard R. Marston. National Monetary Policies and International Financial Markets. North-Holland Publishing Company, 1977.

17. International Monetary Fund. International Financial Statistics. Electronic database.

18. Kubarych, Roger M. Foreign Exchange Markets in the United States. Federal Reserve Bank of New York, August 1978.

19. Levi, Maurice D. International Finance, 2nd ed. McGraw-Hill Publishing Company, 1990.

20. Levich, Richard M. The International Money Market: An Assessment of Forecasting Techniques and Market Efficiency. Contemporary Studies in Economic and Financial Analysis, vol. 22, 1979.

21. Llewellyn, David T. International Financial Integration. John Wiley, 1980.

22. Marston, Richard C., "Interest Arbitrage in the Euro-Currency Markets," European Economic Review, January 1976, pp. 1-13.

23. McKinnon, Ronald I. Money in International Exchange: The Convertible Currency System. Oxford University Press, 1979.

24. Sohmen, Egon, "The Theory of Forward Exchange," Princeton Studies in International Finance No. 17, 1966.

25. Stein, J., "The Forward Rate and the Interest Parity," Review of Economic Studies, March 1965, pp. 113-26.

26. Stock, James and Mark W. Watson, "Variable Trends in Economic Time Series," Journal of Economic Perspectives, vol. 2 (Summer, 1988), pp. 147-74.

27. Wolff, Christian C.P., "Forward Foreign Exchange Rates, Expected Spot Rates, and Premia: A Signal-Extraction Approach," The Journal of Finance, vol. XLII, No. 2, June 1987, pp. 395-406.

Quantitative Policy Objectives: A Case for Moving Targets

Andreas Muller, Department of Health Services Administration, University of Arkansas at Little Rock

Introduction

Planners in government and large corporations are often called upon to propose quantitative objectives which are typically expressed in fixed rates pertaining to a future year. The corporate average fuel efficiency (CAFE) standard proposed to be 40 mpg by the year 2001 is one example[1]; another one is the objective to reduce the U.S. infant mortality rate to 7/1,000 live births by the year 2000.[2]

Stating policy objectives in quantitative terms has several advantages. By definition, quantitative targets allow numerical comparisons; that is, the difference between the base line period and a future period can be expressed as absolute or relative change. Quantitative objectives also help monitor target achievement and potentially make government agencies, or corporations more accountable. Quantitative targets also foster clearer communication between the policy designers and those who have to implement the policy. It is not surprising that in results-oriented circles of government and the corporate world there is growing support for expressing policy objectives in quantitative terms.

The Problem

Comparing quantitative targets with actual performance is a deceptively simple task. For example, a five-year objective is set 20% below the baseline level. A program is implemented which is expected to achieve the 20% reduction. Progress toward the objective is monitored and measurement in the last year of the planning period indicates a 25% reduction compared with the baseline period. A strong temptation exists to declare the program a success and to consider the objective "accomplished". On the other hand, if only a 2% reduction, or an unexpected 5% increase occurred, then the temptation exists to declare the program a failure. Indeed, none of the inferences is justified without assuming that policy relevant factors other than the program intervention did not change, or that those changes were mutually offsetting.

External factors, that is, factors over which government, or the corporation has little control, typically affect policy objectives and need to be taken into account. Without explicit consideration of major policy relevant factors, quantitative targets are likely to be misspecified and incorrect inferences will be drawn about target achievement.

The problem will be illustrated and an adjustment method suggested using the *National Health Promotion and Disease Prevention Objectives*, health status objective 9.3 which states:

> *"Reduce deaths caused by motor vehicle crashes to no more than 1.9/100 million vehicle miles traveled and 16.8 per 100,000 people. (Baseline: 2.4 per 100 million vehicle miles traveled (VMT) and 18.8 per 100,000 people (age adjusted)".*[2] (p.274)

The objective implies that the mileage-based death rate and the age adjusted death rate shall be reduced by 21% and 11%, respectively. The larger reduction target in mileage-based motor vehicle death rates reflects trend adjustment for the period 1980-87. The writers of the policy document acknowledge that the achievement of the objectives will depend upon factors such as the volume of mass transit and various transportation safety initiatives.[2]

---

[1] Insurance Institute for Highway Safety. 1992. *Status Report.* Arlington, VA; Insurance Institute for Highway Safety. 27(6): 5.

[2] U.S Department of Health and Human Services, Public Health Service. 1992. *Health People 2000: National Health Promotion and Disease Prevention Objectives.* Boston, MA; Jones and Bartlett Publishers.

Time Series Analysis of Motor Vehicle Death Rates

Figure 1 shows the population (left axis) and mileage- based motor vehicle death rates (right axis) from 1950-1990.[3] The horizontal lines from 1987 to 2000 indicate the respective year 2000 policy objectives. The population-based death rate are not age-adjusted, thus somewhat overstating the gap between actual and policy target.

Since 1950 the population based motor vehicle death rate stayed in the 20 to 30/100,000 range, with death rates approaching their highest levels in the mid-1960s. Motor vehicle death rates based on vehicle miles traveled show an overall downward trend; otherwise resembling the population based death rates. The vertical grid line indicates the onset of the 1973 oil crisis.

Figure 2 shows U.S. per capita alcohol consumption and the unemployment rate for the years 1950-1990. Per capita alcohol consumption represents the amount of pure drinking alcohol (ethanol) consumed by the U.S. population 14 years of age and older.[4] The unemployment rate is measured by the per cent of the U.S. civilian labor force 16 years of age and older out of work and looking for work in the last two weeks.[5] Prior research found alcohol consumption positively and unemployment negatively related to motor vehicle fatality rates, respectively.[6,7,8,9,10] The 1973-74 oil crisis reduced motor vehicle deaths[6]; its impact needs to be incorporated into the statistical model. A transfer function model[11] can be stated as follows:

$$\Delta Y_t = C + \omega_1 \Delta UR_t + \omega_2 \Delta PCA_t + \omega_3 OILP73_t + \omega_4 OILP74_t + \theta(B)/\phi(B) * a_t; \quad (1)$$

Y_t	= annual motor vehicle death rates;
C	= constant(deterministic trend);
UR_t	= annual unemployment rates;
PCA_t	= per capita alcohol consumption;
$OILP73_t$	= dummy variable pulse coded "1" for 1973, otherwise "0";
$OILP74_t$	= dummy variable pulse coded "1" for 1974, otherwise "0";
$\theta(B)$	= nonseasonal MA operator of order $(1-\theta_1 B-\theta_2 B^2 -\ldots-\theta_q B^q)$;
$\phi(B)$	= nonseasonal AR operator of order $(1-\phi_1 B-\phi_2 B^2 -\ldots-\phi_p B^p)$;
a_t	= residuals $(NI(0)\sigma)$;

[3] U.S. Department of Transportation, Federal Highway Administration. 1991. *Highway Statistics* (various years). Washington, DC; U.S. Government Printing Office.

[4] National Institute on Alcohol Abuse and Alcoholism, Division of Biometry and Epidemiology, Alcohol Epidemiologic Data System. 1991. *Apparent Per Capita Alcohol Consumption: National, State, and Regional Trends: 1977-1988.* Surveillance Report # 16. Washington, DC; CSR,l Inc. Table 1.

[5] Department of Labor, Bureau of Labor Statistics. 1992. *Employment and Earnings* 39(1):162.

[6] Partyka, S.C. 1984. Simple models of fatality trends using employment and population data. *Accident Analysis and Prevention* 16:211-222.

[7] Wagenaar, A.C., Streff F.M. 1989. Macroeconomic conditions and alcohol-impaired driving. *J. of Studies on Alcohol* 50(3):217-225.

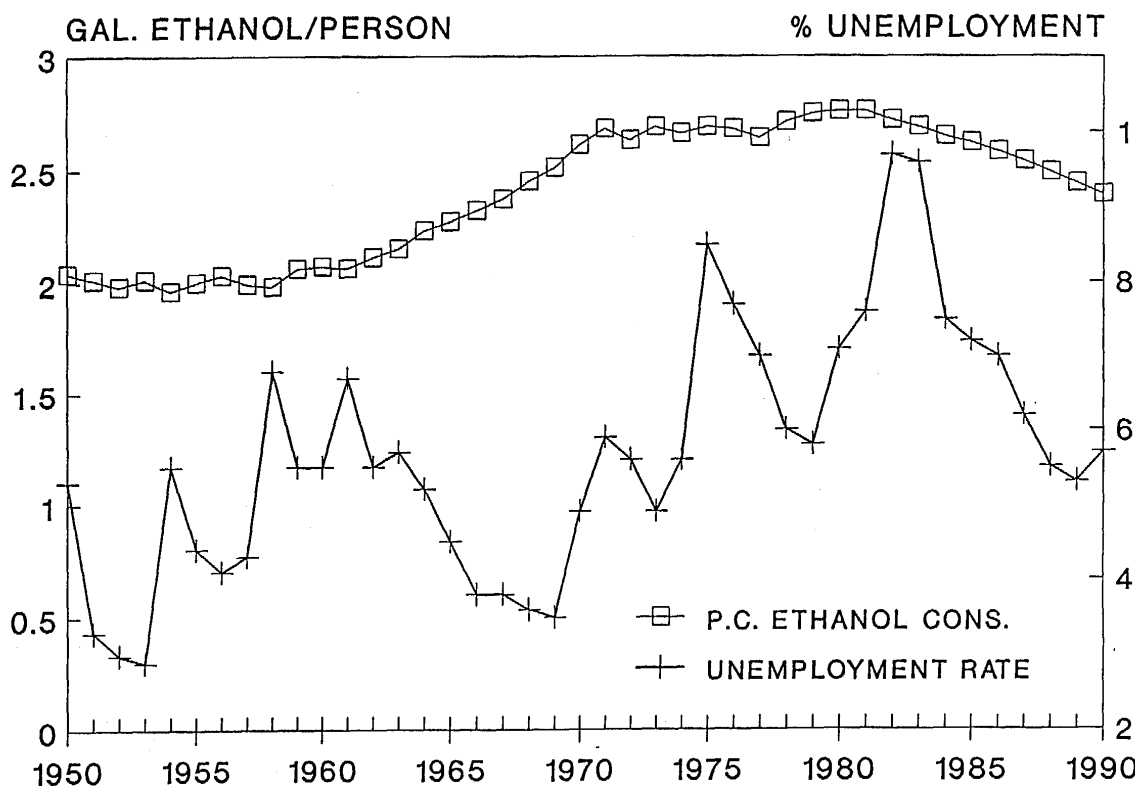
[8] Evans, W.N., Neville, D., Graham, J.D. 1991. General deterrence of drunk driving: evaluation of recent american policies. *Risk Analysis* 11(2):279-289.

[9] Muller, A. 1989. Business recession, alcohol consumption, drinking and driving laws: impact on Oklahoma motor vehicle fatalities and fatal crashes. *American J. of Public Health* 79(10):1366-1370.

[10] Leigh J.P., Waldon, H.M. 1991. Unemployment and highway fatalities. *J. of Health Politics, Policy and Law* 16(1):135-156.

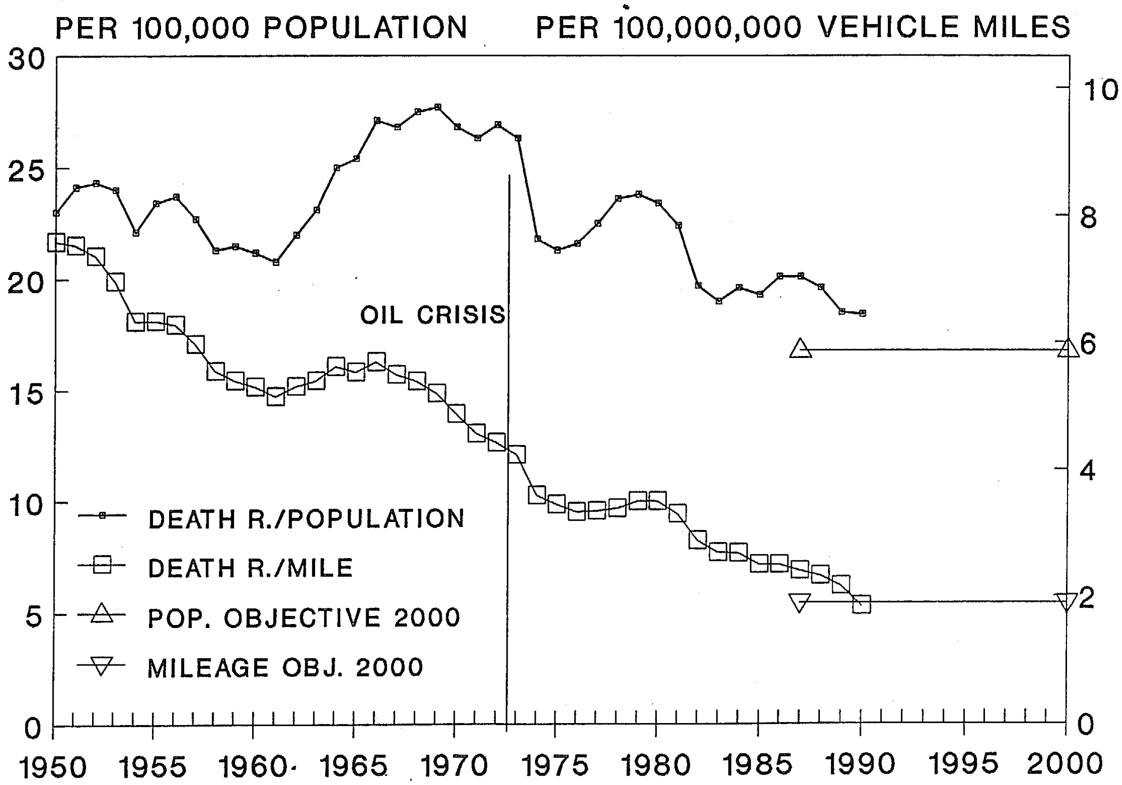
[11] Pindyck, R.S., Rubinfeld, D.L. 1991. *Econometric Models and Economic Forecasts* (3rd ed.). New York, NY; Macmillan.

U.S. PER CAPITA ALCOHOL CONSUMPTION AND UNEMPLOYMENT RATES, 1950-1990



NIAAA, DOL

U.S. MOTOR VEHICLE DEATH RATES (1950-90) AND RELATED SAFETY OBJECTIVES



HIGHWAY SAFETY STATISTICS

To stabilize the variance more, mileage-based death rates were ln-transformed. Since the autoregressive parameters of the output series approached the bounds of stationarity, all time series with the exception of the pulse inputs were once differenced (cf. Pankratz[12]). The model only stipulates contemporaneous unemployment and alcohol consumption effects. An examination of the assumption showed that neither lagged relationships, nor feedbacks were present.

The effect of the oil crises was measured by two dummy variable pulses for 1973 and 1974, respectively. In the final model specification a first-order decay effect associated with the 1974 pulse was found to be statistically significant (see Table 1). The nonseasonal *autoregressive-integrated-moving average* model $\{\theta(B)/\phi(B)a_t\}$ captures the lagged time dependencies of the output series. The model was estimated with RATS 4.0 software[13] using maximum likelihood estimation.

Table 1 shows that a 1% increase in the unemployment rate is estimated to reduce population based death rates by -.67 deaths/100,000, or 2.9%, and the mileage based death rate by approximately 2.0%. The coefficient pertaining to per capita ethanol consumption indicates that a 1 gallon increase in the consumption of alcohol increases the population based motor vehicle death rate by 6.95 deaths/100,000, or about 30% and the mileage based death rate by about 26%. These estimates are comparable with other studies explaining U.S. motor vehicle death rates.[8-10]

The impact of the 1973-74 oil crisis is estimated to have reduced population based death rates by 23% and decayed somewhat thereafter settling at a 19% below levels prevailing before the onset of the crisis. The impact of the oil crisis was less pronounced on mileage based death rates. There was 10.5% reduction in 1974 decaying to approximately 7% reduction over the next few years.

The significant constant term in the mileage-based model indicates a 3.5% annual reduction in the rates independent of other model inputs. It is noteworthy that neither the autoregressive term, nor the constant was statistically significant in the population based model indicating no longterm deterministic trend.

The results of the transfer function analysis illustrate that motor vehicle death rates are significantly affected by external factors such as the oil 1973-74 crisis, unemployment and alcohol consumption.[14] The findings further imply that any departure of policy related factors from their baseline values (1987) will effect the measure used to gauge policy achievement. That is, if policy related external factors depart significantly from baseline values, then the baseline, or the policy objective needs to be adjusted.

Adjusting Policy Baseline, or Objectives

Ideally, unemployment rates and per capita alcohol consumption would be forecasted with little error to the year 2000. Then, the forecasted values could be inserted in equation (1) and adjusted baseline values calculated for the year 2000. However, longterm forecasts of times series which show highly stochastic behavior, i.e. unemployment, may not be very useful since they tend to yield forecasts with large errors.[15]

Performing a sensitivity analysis is an alternate approach to dealing with uncertainty about the future. From past data reasonable ranges can be established for the input variables. For the period 1950-1990, unemployment varied between 2.9 and 9.7%, while per capita alcohol consumption ranged between 1.96 and 2.76 U.S. gallons ethanol. In 1987, the baseline year, unemployment rate was 6.2% and per capital alcohol consumption was 2.54 gallons. Likely future levels of unemployment and ethanol consumption can be expressed as differences from their 1987 baseline levels and multiplied by coefficients in Table 1, i.e. -.67 for unemployment and 6.95 for per capita alcohol consumption. The results are displayed in Table 2.

Table 2 estimates changes in motor vehicle death rates as a result of changes in unemployment and per capita alcohol consumption from their respective baseline levels. The italicized bold entries indicate conditions under

[12] Pankratz, A (1991) *Forecasting with Dynamic Regression Models.* New York, NY; John Wiley and Sons. 177-184.

[13] Doan, T.A. 1992. *Rats 4.0.* Evanston, IL; Estima.

[14] Alcohol consumption may be considered a partially external factor, since it is responsive to government interventions such as taxation, tougher drunk driving laws, and prohibition.

[15] McCleary, R., Hay, R.A. 1980. *Applied Time Series Analysis for the Social Sciences.* Beverly Hills, CA; Sage Publ. chpt. 4.

Table 1

TRANSFER FUNCTION MODELS OF U.S. MOTOR VEHICLE DEATH RATES*

INPUT	MV DEATHS/POPULATION ω_i	t	MV DEATHS/MILE(LN) ω_i	t
D-UNEMPLOYMENT RATE	-.67	-7.28	-.020	-4.61
D-P.C. ETHANOL CONS.	6.95	3.40	.256	2.01
OIL CRISIS PULSE '73	-1.48	2.39	-.040	-1.24
OIL CRISIS PULSE '74	-3.82	-6.28	-.105	-3.54
OIL CRISIS: δ_1	-.32	-2.24	-.493	-2.84
NOISE MOD. $\phi(1)$	-	-	.346	1.72
CONSTANT	-	-	-.035	-4.55
s.e.e.	.61		.032	
R^2 adj.	.948		.993	
DURBIN-WATSON	1.82		1.79	
Q(12)	8.31;	P=.60	Q(10) 7.28	p=.61
d.f.	36		34	
Mean	22.87		1.45	

*All time series are in first differences except for oil crisis which is measured by two pulses and a first-order decay effect.

Table 2

ESTIMATED CHANGE IN MOTOR VEHICLE DEATH RATES (100,000 POP.) DUE TO CHANGE IN UNEMPLOYMENT AND P.C. ETHANOL CONSUMPTION FROM 1987 BASELINE

UNEMPLOYMENT RATE	PER CAPITA ETHANOL CONSUMPTION (gal.) 2.00 (-.54)	2.25 (-.29)	2.50 (-.04)	2.75 (.21)
3% (-3.2)	-1.60	.13	1.87	3.60
4% (-2.2)	**-2.28**	-.54	1.21	2.93
5% (-1.2)	**-2.95**	-1.21	.53	2.26
6% (-.2)	**-3.62**	-1.88	-.13	1.59
7% (.8)	**-4.29**	**-2.55**	-.80	.92
8% (1.8)	**-4.96**	**-3.22**	-1.47	.25
9% (2.8)	**-5.63**	**-3.89**	**-2.14**	-.42
10% (3.8)	**-6.30**	**-4.56**	**-2.81**	-1.09

Numbers in parentheses indicate difference from 1987 baseline levels: per capita ethanol consumption=2.54 gal.; unemployment rate=6.2%.

which the year 2000 objective, a reduction in the population based death rate by 2/1000, is "achieved". For instance, if alcohol consumption were 2.5 gal. in the year 2000 and the unemployment rate 3%, the baseline, or the year 2000 objective would need to be adjusted by 1.87/100,000. In contrast, if the unemployment rate were 9% in 2000, then the baseline, or the policy objective would need to be adjusted downward by 2.14/100,000.

Table 3 shows the estimated per cent change in mileage based motor vehicle death rates by the year 2000 under varying assumptions of unemployment and per capita alcohol consumption. The per cent estimates incorporate an adjustment for the significant longterm deterministic trend[16], i.e. an annual reduction of 3.5% in mileage based death rates between 1987 and 2000.

Since all percentage estimates in Table 3 are less than the national objective (-21%), the year 2000 target is likely to be attained. This observation is not surprising, because the long run trend in mileage based death rates is expected to be about 37% lower, or 1.51/100,000,000 vehicle miles in the year 2000.[17] However, possible variation in per capita alcohol consumption and unemployment rates could result in mileage based death rates ranging from .99 (2.4*(1-.586)) to 1.8 (2.4*(1-.252)) per 100,000,000 vehicle miles by the year 2000.

Discussion

Planners interested in setting realistic objectives know that objectives need to be calibrated with regard to program, or policy effectiveness. What is less well known, or sometimes ignored, is that policies are carried out in a dynamic context in which factors not under government, or corporate control act independently upon policy outcomes. Under such circumstances, policy baselines are not well-described by fixed rates, or simple trends, but may be more like "moving targets".

Consequently, either baseline, or policy objective needs to be adjusted to compensate for departures from baseline values. *Barring such adjustments*, erroneous conclusion will be derived about policy targets, or their attainment. For instance, during boom times, public health and traffic safety programs may appear largely ineffective while during recessions they may appear spectacularly effective, when, indeed, neither conclusion is justified.

The statistical adjustment method proposed in this presentation can be extended to more input variables, although the resulting scenarios will get more complex. It is recommended that only the most important (best predicting) factors are selected for adjustment.

The adjustment weights (coefficients) can be derived from multivariate methods other than transfer function analysis (i.e. pooled, cross-sectional time series analysis, vector autoregressive models, state space models). A meta-analysis of the pertinent literature is recommended when many sources exist. Particular attention ought to be paid to trends when longterm objectives are of interest. Deterministic trends need to be carefully examined and, if present, incorporated in the adjustment.

A limitation of the simple sensitivity analysis proposed in this paper is that it is not helpful in answering questions such as: What is the probability of a reduction of 2 motor vehicle deaths per 100,000 population with or without government action by the year 2000? To answer such questions, more complex multivariate simulation methods ought to be employed.

Conclusion

Quantitative policy objectives ought to be adjusted for changes in major policy relevant factors. The common practice to assume static conditions, or to adjust only for "recent trends," appears inadequate and may result in misleading inferences about the assessment and attainment of policy objectives.

[16] Expressed in per cent, the trend adjustment for the period 1987-2000 was calculated by the formula $((1-.035)^{13}-1)*100=-37.1$. To simplify the presentation, forecast errors are ignored.

[17] The standard error associated with the year 2000 forecast is approximately $\pm .23$ in ln units; or ranging from 1.2 to 1.9/100,000,000 VMT.

Table 3

ESTIMATED PER CENT CHANGE IN MOTOR VEHICLE DEATH RATES (PER 100,000,000 VEHICLE MILES) DUE TO CHANGE IN UNEMPLOYMENT RATES, PER CAPITA ETHANOL CONSUMPTION AND LONGTERM TREND, 1987-2000

UNEMPLOYMENT RATE	PER CAPITA ETHANOL CONSUMPTION (gal.)			
	2.00 (-.54)	2.25 (-.29)	2.50 (-.04)	2.75 (.21)
3% (-3.2)	-44.4	-38.0	-31.6	-25.2
4% (-2.2)	-46.5	-40.1	-33.7	-27.3
5% (-1.2)	-48.5	-42.1	-35.7	-29.3
6% (-.2)	-50.5	-44.1	-37.7	-31.3
7% (.8)	-52.5	-46.1	-39.7	-33.4
8% (1.8)	-54.6	-48.2	-41.8	-35.4
9% (2.8)	-56.6	-50.2	-43.8	-37.4
10% (3.8)	-58.6	-52.2	-45.8	-39.4

Numbers in parentheses indicate difference from 1987 baseline levels: per capita ethanol consumption=2.54 gal.; unemployment rate=6.2%.

References

Insurance Institute for Highway Safety. 1992. *Status Report.* Arlington, VA; Insurance Institute for Highway Safety. 27(6):5.

U.S. Department of Health and Human Services, Public Health Service. 1992. *Health People 2000: National Health Promotion and Disease Prevention Objectives.* Boston, MA; Jones and Bartlett Publishers.

U.S. Department of Transportation, Federal Highway Administration. 1991. *Highway Statistics* (various years). Washington, DC; U.S. Government Printing Office. Table FI-1.

National Institute on Alcohol Abuse and Alcoholism, Division of Biometry and Epidemiology, Alcohol Epidemiologic Data System. 1991. *Apparent Per Capita Alcohol Consumption: National, State, and Regional Trends: 1977-1988.* Surveillance Report #16. Washington, DC; CSR, Inc. Table 1.

Department of Labor, Bureau of Labor Statistics. 1992. *Employment and Earnings* 39 (1) : 162.

Partyka, S.C. 1984. Simple models of fatality trends using employment and population data. *Accident Analysis and Prevention* 16:211-222.

Wagenaar, A.C., Streff F.M. 1989. Macroeconomic conditions and alcohol-impaired driving. *J. of Studies on Alcohol* 50(3):217-225.

Evans, W.N., Neville, D., Graham, J.D. 1991. General deterrence of drunk driving: evaluation of recent american policies. *Risk Analysis* 11(2):279-289.

Muller, A. 1989. Business recession, alcohol consumption, drinking and driving laws: impact on Oklahoma motor vehicle fatalities and fatal crashes. *American Journal of Public Health* 79(10):1366-1370.

Leigh J.P., Waldon, H.M. 1991. Unemployment and highway fatalities. *J. of Health Politics, Policy and Law* 16(1):135-156.

Pinkyck, R.S., Rubinfeld, D.L. 1991. *Econometric Models and Economic Forecasts* (3rd ed.). New York, NY; Macmillan.

Pankratz, A. (1991) *Forecasting with Dynamic Regression Models.* New York, NY; John Wiley and Sons. 177-184.

Doan, T.A. 1992. *Rats 4.0.* Evanston, IL; Estima.

McCleary, R., Hay, R.A. 1980. *Applied Time Series Analysis for the Social Sciences.* Beverly Hills, CA; Sage Publ. chpt. 4.

Latin Hypercube Sampling for Sensitivity and Uncertainty Analysis of Energy Forecasting Model

Ramesh Dandekar and Nancy Kirkendall, Energy Information Administration

Abstract

The Energy Information Administration (EIA) of the Department of Energy uses a wide collection of models for a mid-term forecast of U.S. energy production, supply, distribution, and consumption. The models are mostly deterministic and provide point forecasts for several times in the future. This paper addresses model forecast accuracy as a function of the accuracy of the model inputs. The use of Latin Hypercube sampling (LHS) technique for sensitivity and uncertainty analysis is demonstrated on a coal supply/demand model. The LHS technique provides the information needed for evaluating the impact of all model inputs and their interactions on the forecasts with a minimum number of computer runs. The results are compared with the commonly used "one-variable-at-a-time" sensitivity analysis.

Introduction

The Energy Information Administration (EIA) of the Department of Energy uses deterministic models for a mid-term energy forecast. These models are used to determine the "expected" response of the national (and international) energy system to "most likely" values of input parameters. As a result, the major focus of EIA's mid-term forecasting analysis is within a few percentiles of the middle of the distribution of input parameters. The accuracies of the point forecasts from these models depend on the accuracy of the model inputs. This paper addresses model forecast accuracy as a function of the accuracy of the model inputs. Only selected model inputs were varied. The results were examined, and observations were made concerning unusual features of the model. Some of these features may be well known to modelers, and easily explained. Others may illustrate potential problems with the model results.

The study also serves to demonstrate the applicability of Latin Hypercube sampling (LHS) technique to conduct sensitivity and uncertainty analysis of EIA's modeling system. LHS was originally proposed by McKay et. al. (1979). Correlations among the input variables was incorporated by using a method described in Iman and Conover (1982). For this first application the input variables were specified as being uncorrelated.

Assumptions

The LHS technique requires specifying the statistical distribution functions of all input variables. To implement a range around the "most likely" value for input variables, we chose to use a uniform distribution to characterizes all input variables. This should be conservative as it includes more observations in the tails of distribution.

In deciding on the range of the uniform distribution, our objective was to have a wide enough range to minimize the possibility of excluding a realistic combination of input variables from the analysis. A range of plus or minus 25 percent around the base values for all input variables was judged to be adequate for this purpose. The study assumes that the input variables are uncorrelated.

Candidate Model

This study used two of EIA's coal models. The first model, the Resource Allocation and Mine Costing (RAMC) model, generates coal supply curves using input geological and financial parameters. The second model, the Coal Supply and Transportation Model (CSTM), uses the coal supply curves from the RAMC model and coal demand forecasts from the Intermediate Future Forecasting System (IFFS) to produce coal production and price forecasts. The supply/demand linkage is done through a surface transportation network solution algorithm imbedded in the CSTM model.

The RAMC model consists of two sub-modules: the geological submodule and the financial submodule. In the geological submodule, the available coal reserves by coal type and coal supply region are used to create hypothetical mining operations with their own operating characteristics. In the financial submodule, the costing of these mines is done to determine their minimum acceptable coal prices. These two modules operate independently within the RAMC model. As a result, the decision was made to analyze the two modules separately using two independent LHS experiments. Due to the space limitations, the discussion in this paper is limited to the analysis of the financial sub-module.

The CSTM model connects demand from 47 coal demand regions (43 domestic regions and 4 export regions) from the IFFS modeling system with 32 coal supply regions in the RAMC model. The supply-demand linkage is done through a transportation network consisting of over 200 nodes and 700 links. These nodes and linkages include both rail and water transport modes. To satisfy input coal demands by region, the least cost coal supply quantities and prices are determined by use of a least cost optimization algorithm imbedded in the CSTM model.

As a part of this demonstration project, the coal demand coming from the IFFS modeling system was kept constant at its base level. Similarly, the transportation network from the base forecast was used without any changes throughout the project. Such a configuration allowed us to study the effects of changes in the coal supply side input variables on the changes in the regional coal outputs.

The effects of nine financial input parameters from the RAMC coal supply model were evaluated for the forecast year 2000.

Input Financial Variables

The financial input variables to the RAMC model belong to two broad categories. The variables belonging to the first category are related to coal extraction costs. These costs are strictly based on engineering considerations of a mining operation. Variables belonging to this category includes: capital costs (initial and deferred), labor costs, annual operation and maintenance (O&M) costs, utility costs (such as water, electricity and other fuels), and costs of other miscellaneous supplies. In addition, the engineering cost components include estimates for escalation rates, utility discount rates and rate of return on investment. In the RAMC model these costs are estimated either on a judgmental basis or are derived using a detailed engineering analysis of a hypothetical mining operation. This study focuses on the sensitivities of this category of financial variables.

The second category of financial input variables consists of social and welfare costs imposed by local and Federal governments. These costs are usually imposed as taxes, royalties, fees and insurance payments. In EIA forecasting models, including the RAMC model, these costs have been implemented at their current levels. This sensitivity analysis ignores the effects of changes in this category of financial variables.

Overview of Latin Hypercube Design

As a part of the experiment, fifty combined RAMC/CSTM computer model runs were made to estimate the effects of changes in the nine coal supply side financial parameters. The choice of making fifty computer runs was strictly arbitrary. All the nine financial parameters were changed simultaneously over the range +-25% of their base values in these fifty computer runs. LHS was used to identify the sample points in the multivariate input space. These financial parameters were assumed to be independent. A method developed by Iman and Conover was used to induce a desired rank correlation of zero among the 9 financial variables.

Output From the Experiment

The RAMC uses 32 coal supply regions. These 32 basic coal supply regions are further aggregated into eleven major coal supply regions. The demand component of the model consists of 43 domestic demand regions and four overseas demand regions. Coal production and price forecasts from the CSTM model provides information on the coal production for each coal supply region by coal mining method (surface, deep and total).

Each of the fifty model runs gave different values for the regional forecasts. The values from these fifty computer runs were used to calculate percent change from the corresponding base forecast values. The mean and standard deviation of the percent change values were computed for each variable in each region. Additionally, the empirical CDF was calculated along with 5, 50, and 95 percentile values. Although all input variables followed uniform distributions, the output histograms for all output variables generally showed a mound shaped distribution with varying degrees of skewness. This indicates that the output variables are not dominated by changes in only one input variable. The summary statistics from the experiment for total of surface and deep mining activity is presented in Table 1. Similar tabular outputs, in the same format as table 1, were generated separately for surface and deep mining methods, but are not included here due to space limitations.

Summary Statistics

The first column in Table 1 shows the base forecast values in millions of tons. The remaining five columns provide the summary statistics obtained from the fifty computer runs. These statistics are measured as a percent deviation from the base value in the first column. Included are the mean, the standard deviation, the 5 percentile value, the median, and the 95 percentile value.

The mean and the median indicates the extent of possible shifts from the base case forecast. Positive values indicate that the forecast values tend to be larger than the base case, whereas negative values indicate that forecasts tend to be smaller.

The standard deviation (third column) is a measure of the dispersion of forecasts. The relatively large standard deviations in these tables indicate relatively high uncertainty in the forecasts.

The 5 and 95 percentile values provide an empirical 90 percent confidence interval for the forecast values. These percentiles can also be used to evaluate the symmetry of the distribution of forecasts.

The output from the fifty computer runs were also used to compute correlations between each of the model output variables. The variables showing very high correlation coefficients, both positive and negative (exceeding 0.80 in absolute value) are identified in Table 2.

A positive correlation indicates that the production in the two regions moves together, either up or down, showing that they react to changes in the input variables in the same way. A negative correlation indicates that production in the two regions moves in opposite directions in reaction to changes in the input variables. This could be the result of competitive markets for coal as implemented in the model.

Changes in Production Pattern

Confidence Intervals

As a result of changes in the financial variables, 90 percent confidence intervals for forecasts of surface production at the U.S. level were -4.75 percent to 3.59 percent; for deep production were -4.52 percent to 6.14 percent; and for total production were -.33 percent to .22 percent. The small impact on the U. S. Total production is expected as total demand was not changed. For finer levels of geographic detail, the 90 percent confidence intervals are much wider.

No Impact of Input Variables

There were four regions which had non-zero total production in the base case scenario, but which showed no response to changes in the financial variables. These were Shreveport, LA (deep mines), and Corsicana, TX, Winslow AZ and Healy AK (surface mines). It therefore appears that in the RAMC/CSTM model, coal produced from these four supply regions does not compete with coal from other regions. Based on the explanation by EIA coal industry specialist, the production for these four supply regions is dominantly lignite. Lignite is seldom transported more than a few miles from the mine.

High Variation

Regions highlighted below as having "high variation" are those which had base case production in excess of 10 million tons, and a standard deviation of greater than 10 percent. Those that had high variation in response to changes in financial variables include Cambridge OH, Clinton TN, Cordova AL and Carbondale CO. Based on the explanation by the coal industry specialist, the potential instabilities of coal production in Cambridge OH, Cordova AL and Carbondale CO are well known to coal supply modelers. They stem from model reliance on the assumption that consumer coal choices are dominated by minimization of short term delivered coal price.

TABLE 1:

SUMMARY STATISTICS FROM FIFTY LHS MODEL RUNS

SUPPLY REGION	(MM TONS) BASE VALUE	<- PERCENT DEVIATION FROM BASE VALUE -> MEAN	STD DEV	5 PERC	MEDIAN	95 PERC
=============	==========	======	=======	=======	=======	=======
PA JOHNSTOWN PA	61.27	0.51	6.24	-7.89	-0.55	11.18
OH CAMBRIDGE OH	20.45	0.23	16.47	-23.29	1.72	28.85
MD LONACONING MD	8.63	-0.36	8.11	-14.34	-0.60	12.85
NV CLARKSBURG WV	53.86	-1.21	2.92	-4.80	-0.68	1.06
SV CHARLESTON WV	149.82	-0.09	1.56	-2.20	-0.33	3.37
VA APPALACHIA VA	55.07	0.57	2.12	-2.20	0.30	4.26
EK HAZARD KY	143.41	-0.61	3.85	-7.44	-0.54	5.26
TN CLINTON TN	17.50	0.23	10.30	-18.05	1.70	16.65
AL CORDOVA AL	28.07	-1.63	16.11	-31.33	0.62	24.33
WK CENTRAL CITY KY	39.61	1.80	3.49	-4.26	1.20	8.35
IL CENTRALIA IL	56.90	-2.84	3.56	-9.17	-1.65	1.38
IN HUNTINGBURG IN	27.73	0.49	3.81	-4.83	0.05	5.94
IA OTTUMWA IA	3.50	0.27	1.53	-0.03	0.00	1.45
MO CLINTON MO	0.47	1.79	23.42	-28.06	-1.58	36.29
KS PITTSBURG KS	0.19	3.30	76.66	-100.00	9.79	106.70
AR RUSSELVILLE AR	0.84	6.83	17.29	-18.73	7.98	41.65
LA SHREVEPORT LA	4.11	0.00	0.00	0.00	0.00	0.00
OK TULSA OK	4.68	-6.29	17.32	-40.72	-3.41	18.64
TX CORSICANA TX	73.18	0.00	0.00	0.00	0.00	0.00
ND WILTON ND	30.05	-0.05	0.14	-0.44	0.00	0.00
SD LEMMON SD	0.00	0.02	0.04	0.00	0.00	0.13
EM SIDNEY MT	0.00	0.00	0.00	0.00	0.00	0.00
WM BILLINGS MT	51.42	0.22	1.41	-1.35	-0.30	2.71
NW GILLETTE WY	187.59	-0.29	0.30	-0.84	-0.23	0.10
SW ROCK SPRINGS WY	27.05	2.07	2.36	-1.26	1.98	6.25
CN HAYDEN CO	0.00	0.00	0.00	0.00	0.00	0.00
CS CARBONDALE CO	18.45	14.87	22.21	-15.60	4.03	52.54
UT SUNNYSIDE UT	30.12	-0.36	6.51	-9.49	-1.61	10.89
AZ WINSLOW AZ	10.65	0.00	0.00	0.00	0.00	0.00
NM GALLUP NM	23.23	-1.71	3.51	-7.12	-0.24	1.94
WA CENTRALIA WA	1.81	-2.70	9.82	-20.22	0.25	11.67
AK HEALY AK	2.01	0.00	0.00	0.00	0.00	0.00
U. S. Total	1131.65	-0.01	0.15	-0.33	0.00	0.22
NORTHERN APPALACHIA	144.21	-0.22	1.10	-1.65	-0.11	1.38
CENTRAL APPALACHIA	365.80	-0.18	2.34	-3.98	0.06	4.15
SOUTHERN APPALACHIA	28.07	-1.63	16.11	-31.33	0.62	24.33
MIDWEST	124.24	-0.62	1.54	-3.45	-0.36	1.52
WEST INTERIOR	86.98	-0.25	0.83	-1.70	-0.17	0.95
NORTH GREAT PLAINS	296.10	0.04	0.19	-0.15	-0.01	0.46
ROCKIES & SOUTHWEST	82.44	2.71	6.13	-6.33	0.67	13.31
NORTHWEST & ALASKA	3.82	-1.28	4.65	-9.57	0.10	5.51
APPALACHIA	538.07	-0.27	0.71	-1.58	-0.17	0.76
INTERIOR	211.21	-0.46	0.70	-1.72	-0.33	0.64
WEST	382.37	0.61	1.20	-1.06	0.14	2.85
EAST OF MISS. RIVER	662.31	-0.33	0.38	-1.06	-0.14	0.08
WEST OF MISS. RIVER	469.34	0.45	0.85	-0.68	0.08	2.01

TABLE 2

CORRELATION COEFFICIENTS AMONG SUPPLY REGIONS

SUPPLY REGION and MINE TYPE	Cor Coef	SUPPLY REGION and MINE TYPE
NORTHERN APPALACHIA SURFACE	-0.961	NORTHERN APPALACHIA DEEP
CENTRAL APPALACHIA DEEP	-0.930	NORTHERN APPALACHIA SURFACE
SOUTHERN APPALACHIA SURFACE	-0.871	NORTHERN APPALACHIA DEEP
SOUTHERN APPALACHIA SURFACE	-0.917	CENTRAL APPALACHIA DEEP
SOUTHERN APPALACHIA TOTAL	-0.943	CENTRAL APPALACHIA TOTAL
MIDWEST SURFACE	-0.800	NORTHERN APPALACHIA DEEP
WEST INTERIOR DEEP	-0.869	NORTHERN APPALACHIA SURFACE
WEST INTERIOR SURFACE	-0.892	NORTHERN APPALACHIA DEEP
WEST INTERIOR SURFACE	-0.887	CENTRAL APPALACHIA DEEP
WEST INTERIOR SURFACE	-0.832	WEST INTERIOR DEEP
WEST INTERIOR TOTAL	-0.811	SOUTHERN APPALACHIA TOTAL
NORTH GREAT PLAINS DEEP	-0.918	NORTHERN APPALACHIA SURFACE
NORTH GREAT PLAINS DEEP	-0.822	SOUTHERN APPALACHIA SURFACE
NORTH GREAT PLAINS DEEP	-0.846	WEST INTERIOR SURFACE
ROCKIES & SOUTHWEST TOTAL	-0.927	CENTRAL APPALACHIA TOTAL
ROCKIES & SOUTHWEST TOTAL	-0.887	WEST INTERIOR TOTAL
NORTHWEST & ALASKA DEEP	-0.864	SOUTHERN APPALACHIA DEEP
NORTHWEST & ALASKA DEEP	-0.890	ROCKIES & SOUTHWEST DEEP
NORTHWEST & ALASKA TOTAL	-0.902	SOUTHERN APPALACHIA TOTAL
NORTHWEST & ALASKA TOTAL	-0.929	ROCKIES & SOUTHWEST TOTAL
APPALACHIA SURFACE	-0.966	APPALACHIA DEEP
INTERIOR SURFACE	-0.853	APPALACHIA DEEP
INTERIOR SURFACE	-0.815	INTERIOR DEEP
WEST TOTAL	-0.959	APPALACHIA TOTAL
EAST OF MISS. RIVER SURFACE	-0.986	EAST OF MISS. RIVER DEEP
WEST OF MISS. RIVER TOTAL	-0.957	EAST OF MISS. RIVER TOTAL
WEST INTERIOR DEEP	0.816	CENTRAL APPALACHIA DEEP
WEST INTERIOR DEEP	0.818	NORTHERN APPALACHIA DEEP
ROCKIES & SOUTHWEST TOTAL	0.824	SOUTHERN APPALACHIA TOTAL
SOUTHERN APPALACHIA SURFACE	0.850	NORTHERN APPALACHIA SURFACE
WEST INTERIOR SURFACE	0.851	MIDWEST SURFACE
INTERIOR SURFACE	0.854	APPALACHIA SURFACE
NORTH GREAT PLAINS DEEP	0.864	WEST INTERIOR DEEP
WEST INTERIOR TOTAL	0.871	CENTRAL APPALACHIA TOTAL
WEST INTERIOR SURFACE	0.874	SOUTHERN APPALACHIA SURFACE
NORTH GREAT PLAINS DEEP	0.900	CENTRAL APPALACHIA DEEP
CENTRAL APPALACHIA DEEP	0.922	NORTHERN APPALACHIA DEEP
NORTH GREAT PLAINS DEE	0.925	NORTHERN APPALACHIA DEEP
WEST INTERIOR SURFACE	0.932	NORTHERN APPALACHIA SURFACE
NORTHWEST & ALASKA TOTAL	0.933	CENTRAL APPALACHIA TOTAL
NORTHWEST & ALASKA DEEP	0.935	WEST INTERIOR DEEP
NORTHWEST & ALASKA TOTAL	0.936	WEST INTERIOR TOTAL
ROCKIES & SOUTHWEST DEEP	0.948	SOUTHERN APPALACHIA DEEP

All but Clinton, TN also showed high variation in response to financial variables for deep mines. Cordova, AL showed high variation in response to financial variables for surface mines. Clarksberg WV, and the region Northern Appalachia also had high variation for surface mines in response to the financial variables.

The Southern Appalachia region is solely made up of Cordova, AL. As a result it too shows excessive variation in response to financial parameters. The Southern Appalachian region has a smaller base value and larger standard deviation than the seven other coal supply regions in the RAMC. Thus, the rationale for identifying Southern Appalachia as a separate supply region needs to be re-evaluated.

Change in Mean Forecast of Production

A simple statistical test of the mean percent change yields some indication of those regions where the symmetrical changes in the input variables tend to favor coal production in a region (a positive mean indicates increased production). The test indicates that the mean is significantly different from zero (at the 95 percent level) if it is larger in absolute value than the population standard deviation divided by 3.5, (roughly 2 divided by the square root of 50).

Table 1 indicates that the changes in the financial variables cause slight, but significant, increases in production West of the Mississippi, and corresponding decreases East of the Mississippi. Table 2 shows that the correlation between production values in these regions is -.96 for the financial variables.

The increase in the West occurs primarily in the Rockies and Southwest (Table 1). This comes from deep mines. The largest single contributor is Carbondale CO, which shows a large increase in total production of 15 percent and in deep production of 20 percent in response to changes in the financial variables.

Part of the decrease in production East of the Mississippi occurs in the Midwest and is due to a decrease in deep production. This occurs in response to changes in financial variables. Also in the Midwest there is a smaller, but still significant, increase in surface production due to changes in financial variables.

In response to financial variables there is a decrease in production from surface mines in Northern Appalachia and an increase in production from deep mines. The correlation between Northern Appalachian deep and surface mines was -.96 for the financial experiment.

Observations Based on Correlations

Table 2 shows correlations from the financial experiment. It shows that Southern Appalachian total coal production and production from the Rockies and Southwest (which are positively correlated with each other) are strongly negatively correlated with production in three different regions: Central Appalachia, West Interior, and Northwest and Alaska.

Based on the coal industry specialist, the inverse relationships with Central Appalachia and West Interior are to be expected, since these are the nearest competitive regions providing low sulfur and high sulfur steam coals respectively. Central Appalachia also competes with Southern Appalachia (Alabama) as a source of metallurgical coals in both the domestic and export markets.

The high negative correlation between Southern Appalachia and Northwest & Alaska is more difficult to explain, as these regions' products do not compete in any market. Production in these three regions also shows pairwise positive correlation. These correlation patterns could illustrate the effects of competition. The true potential for such shifts in the national coal market has been of interest to coal producers and consumers since mid-1970's. Competition between these sources has been sharpened by regulatory incentives for the use of low sulfur bituminous coal.

Southern Appalachia deep coal, which is solely made up of Alabama deep coal, shows strong positive correlation between deep production in the Rockies and Southwest. Both regions have strong negative correlation with deep production in the Northwest and Alaska. This indicates that deep coal produced from the Rockies and the Southwest in combination with Southern Appalachia deep coal could compete with deep coal produced from Northwest and Alaska. At the same time deep production in the Northwest and Alaska is positively correlated with deep production in West Interior, Central and Northern Appalachia and the Northern Great Planes.

Surface production in Southern and Northern Appalachia, West Interior and the Midwest are pairwise positively correlated. They are negatively correlated with deep production from West Interior, Central and Northern Appalachia.

Traditional Sensitivity Analysis Using Data From LHS

The analysis so far has been restricted to evaluation of summary statistics based on the model output. The next step was to quantify the average contribution of each of the input variables to the average model output uncertainty. Such an evaluation provides valuable insight into the relative importance of each input variable in determining model outcome.

To accomplish this objective, a linear regression model of the form

$$\frac{(Y-Y_b)}{Y_b} = \sum_{j=1}^{m} \alpha_j \frac{(X_j-X_{bj})}{X_{bj}}$$

was estimated. Where, Y = Model output variable, X_j = j^{th} Model input variable, Y_b = Base Forecast of the model output variable, X_{bj} = Base value of the j^{th} model input variable and m = the total number of model input variables selected for the analysis. In this regression model all the variable values are expressed as a relative change from their base values. This regression model assumes that the cumulative effect of changes in all the input variables on the model output is linear. The regression coefficients and related R^2 values from these analyses are presented in Table 3. In the table the "*" next to the regression coefficient indicates that the related t statistic is less than 2 and therefore the regression coefficient is not significantly different from zero. Similar tables were generated for each coal mining methods and are used in the analysis that follows.

In table 3, a positive value for the regression coefficient indicates that the relative change in the output variable is in the same direction as the relative change in the input variable. The larger the value of the regression coefficient, the more influential is the input variable in determining the model outcome. The input variables with larger regression coefficients should be carefully evaluated. The R^2 value is located in the first column of the table. It is a measure of fraction of the output variation which is explained by the variation in the input variables. The R^2 value shown here is the "adjusted R^2" as this problem is stated as a regression through the origin.

Regression Fit

Tables 3 shows the results of regression analysis applied to total mines by supply region using data from the financial experiment. Similar tables were also generated for surface and deep mining methods but are not included in this paper.

For the U. S. total production (Table 3) all financial variables were statistically significant except for utility discount rate. However, only labor rate, labor productivity and escalation factors appear to be important in determining the relative changes in total coal production. For deep and surface mines at the National level the three important variables are labor rate, labor productivity and annual supplies & parts cost.

The regression coefficients for deep and surface mines at the National level are generally of opposite sign. Thus for example, an increase in labor rates tends to result in a decrease in deep production and an increase in surface production. Increases in labor productivity and annual supplies & parts cost, however, result in an increase in deep production and a decrease in surface production.

For most regions, it appears that there is little impact on production of changes in deferred capital costs, water and power costs, utility discount rates and industrial ROR. For these variables the estimated coefficients tend not to be statistically significant. Initial capital costs tend to be significant for only surface coal production in the West and West of the Mississippi river. The sign in those regions is negative, indicating that an increase in initial capital costs tends to result in a decrease in surface production.

The R^2 statistics shown in table 15 vary widely from one supply region to other. This shows that the linear model is good in some regions and poor in others.

Table 3 - REGRESSION COEFFICIENTS - Latin Hypercube Design - Total Mines

SUPPLY REGION	RSQUARE	INI CAP	DEF CAP	LABR RATE	LABR PROD	W&S COST	OTHR SUP	ESCA FACT	UT DIS RT	IND. ROR
PA JOHNSTOWN PA	0.912	0.030*	0.041	0.176	-0.259	0.014*	0.158	0.224	0.033*	0.039
OH CAMBRIDGE OH	0.904	-0.198	-0.040*	-0.501	0.631	-0.117	-0.381	-0.598	-0.065*	-0.136
MD LONACONING MD	0.978	-0.144	-0.011*	-0.271	0.338	-0.040	-0.204	-0.252	-0.006*	-0.068
NV CLARKSBURG WV	0.139	0.064	-0.004*	0.069	-0.047*	-0.050*	0.044*	0.061	-0.006*	-0.006*
SV CHARLESTON WV	0.000	0.014*	0.009*	0.003*	0.006*	0.014*	-0.019*	0.001*	-0.027*	-0.003*
VA APPALACHIA VA	0.816	-0.029	-0.014*	-0.093	0.080	0.012*	-0.028	-0.057	0.005*	-0.006*
EK HAZARD KY	0.940	-0.034	-0.019*	-0.153	0.179	-0.018*	-0.029	-0.114	0.006*	-0.015*
TN CLINTON TN	0.941	-0.013*	-0.069	-0.372	0.485	-0.017*	-0.192	-0.269	-0.052	-0.016*
AL CORDOVA AL	0.925	0.064*	0.042*	0.715	-0.712	0.117	-0.052*	0.413	0.036*	0.034*
WK CENTRAL CITY KY	0.000	0.011*	-0.007*	0.080	-0.114	-0.010*	0.053*	0.011*	0.024*	-0.028*
IL CENTRALIA IL	0.000	0.077	-0.005*	0.089	-0.093	0.009*	0.126	0.100	0.013*	0.016*
IN HUNTINGBURG IN	0.882	-0.088	-0.005*	-0.093	0.105	0.013*	-0.154	-0.117	0.002*	-0.024*
IA OTTUMWA IA	0.043	-0.023*	-0.013*	0.021*	0.001*	-0.008*	-0.033	-0.015*	-0.006*	-0.009*
MO CLINTON MO	0.915	-0.298	-0.041*	-0.364	0.466	-0.198	-1.253	-0.650	-0.019*	-0.204
KS PITTSBURG KS	0.936	0.995	0.157*	1.118	-1.406	0.522	4.170	2.229	0.025*	0.694
AR RUSSELVILLE AR	0.677	-0.224	0.184*	0.454	-0.660	0.076*	-0.759	0.219	-0.095*	0.089*
LA SHREVEPORT LA	0.000	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*
OK TULSA OK	0.704	-0.228	-0.079*	-0.540	0.792	-0.117*	-0.107*	-0.551	0.075*	-0.123*
TX CORSICANA TX	0.000	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*
ND WILTON ND	0.193	0.001*	-0.001*	0.005	-0.002*	0.000*	0.003	0.002*	-0.001*	0.001*
SD LEMMON SD	0.193	0.000*	0.000*	-0.002	0.001*	0.000*	-0.001	-0.001*	0.000*	0.000*
EM SIDNEY MT	0.000	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*
WM BILLINGS MT	0.648	-0.021	-0.008*	-0.031	0.036	0.000*	-0.039	-0.054	-0.006*	-0.007*
NW GILLETTE WY	0.000	0.007*	0.000*	0.003*	-0.003*	0.001*	0.004*	0.008	0.001*	-0.001*
SW ROCK SPRINGS WY	0.000	-0.054*	0.042*	0.037*	-0.036*	-0.010*	-0.052*	-0.014*	-0.015*	0.005*
CN HAYDEN CO	0.000	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*
CS CARBONDALE CO	0.339	0.161*	0.151*	0.754	-0.842	0.109*	0.445	0.786	0.092*	0.209*
UT SUNNYSIDE UT	0.864	0.079	-0.031*	0.094	-0.300	0.031*	0.180	0.192	0.012*	0.095
AZ WINSLOW AZ	0.000	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*
NM GALLUP NM	0.529	-0.002*	0.004*	-0.087	0.162	-0.001*	-0.060	-0.113	-0.024*	-0.022*
WA CENTRALIA WA	0.850	-0.124	-0.067*	-0.323	0.481	-0.094	-0.055*	-0.282	-0.015*	-0.094
AK HEALY AK	0.000	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*
U. S. Total	0.956	0.002	0.001	0.006	-0.006	0.001	0.002	0.006	0.000*	0.001
NORTHERN APPALACHI	0.221	0.000*	0.009*	0.013*	-0.018*	-0.032	0.017*	0.018*	0.002*	-0.009*
CENTRAL APPALACHIA	0.938	-0.013	-0.009*	-0.090	0.108	0.000*	-0.032	-0.066	-0.011*	-0.009*
SOUTHERN APPALACHI	0.925	0.064*	0.042*	0.715	-0.712	0.117	-0.052*	0.413	0.036*	0.034*
MIDWEST	0.449	0.019*	-0.006*	0.045	-0.056	0.004*	0.040	0.023	0.014*	-0.007*
WEST INTERIOR	0.750	-0.015	-0.003*	-0.023	0.036	-0.006*	-0.012	-0.027	0.003*	-0.006*
NORTH GREAT PLAINS	0.589	-0.004	0.002*	0.000*	0.001*	0.000*	-0.009	-0.005	-0.002*	-0.001*
ROCKIES & SOUTHWES	0.694	0.065*	0.024*	0.178	-0.252	0.036*	0.149	0.214	0.018*	0.075
NORTHWEST & ALASKA	0.849	-0.059	-0.032*	-0.153	0.228	-0.045	-0.026*	-0.134	-0.007*	-0.045
APPALACHIA	0.719	-0.005*	-0.001*	-0.021	0.031	-0.003*	-0.020	-0.018	-0.005*	-0.007*
INTERIOR	0.000	0.005*	-0.005*	0.017	-0.018	0.000*	0.019	0.003*	0.009*	-0.006*
WEST	0.608	0.010*	0.007*	0.037	-0.051	0.007*	0.025	0.041	0.002*	0.015*
EAST OF MISS. RIVE	0.000	-0.001*	-0.002*	-0.008*	0.015	-0.001*	-0.009	-0.010	-0.001*	-0.007*
WEST OF MISS. RIVE	0.543	0.006*	0.005*	0.026	-0.035	0.005*	0.018	0.028	0.003*	0.011*

Comparison of Results With One-at-a-Time Sample

To compare the results from LHS with those from the method of changing only one input variable at a time, an additional set of model runs were made. The set varied the nine financial input variables over the same range, -25 percent to +25 percent, one-at-a-time. As 50 observations of each variable in a one-at-a-time experiment would have required 900 additional computer runs, a reduced number of observations was used for each variable. A total of 10 observations were made for each variable at 5 percentage point intervals.

The same linear regression models used in the LHS experiments were also used with the one-at-a-time experiment. Results of the regression analysis applied to production by region are given in Table 4.

The findings from Table 4 for the one-at-a-time design are similar to those from Table 3 for the LHS design. In particular, conclusions concerning the important explanatory variables and the signs of the estimated coefficients tend to be the same. The actual values of the estimated coefficients are sometimes similar, and sometimes quite different. This may be due in part to interactions among the variables, and it may be due in part to the difference in sample sizes. The LHS design used a total of 50 runs, with 50 different observed values of each input variable. The one-at-a-time design used a total of 90 runs, with 10 different observed values of each input variable.

Conclusions

- The forecasts for coal supply regions of Alabama, Ohio and South Colorado appear to be highly variable in response to changes in both financial and geological input variables.

- Of the nine financial variables selected for the analysis; labor rates, labor productivity and escalation factors appear to be the three most important variables in determining the level of total coal production at the national level.

- Based on the financial input uncertainties, the national level 90 percent confidence interval for delivered coal prices is approximately plus and minus 5 percent. The demand regions east of the Mississippi have relatively wider 90 percent confidence intervals for price than the western demand regions.

- After labor rates and labor productivity, annual supplies & parts cost is the most important financial input variable to explain changes in surface and deep mine production for all supply regions.

- For the both financial and geological experiments, the LHS design provides results comparable to one-at-a-time sample design. By comparable, we mean that the two procedures identify the same input variables as being important and the signs of the estimated coefficients are the same.

- For both the financial and geological experiments, the values of the coefficients estimated by regression methods applied to the LHS data and to the one-at-a-time sample may be quite different. This is most likely due to interactions and other nonlinear effects, although for the financial variables the difference in the sample size may be a contributing factor.

Table 4 - REGRESSION COEFFICIENTS - Orthogonal Design - Total Mines

SUPPLY REGION	RSQUARE	INI CAP	DEF CAP	LABR RATE	LABR PROD	W&S COST	OTHR SUP	ESCA FACT	UT DIS RT	IND. ROR
PA JOHNSTOWN PA	0.774	0.069	0.017*	0.175	-0.255	0.020*	0.124	0.226	0.012*	0.029*
OH CAMBRIDGE OH	0.927	-0.195	-0.045*	-0.551	0.741	-0.096	-0.258	-0.630	-0.030*	-0.046*
MD LONACONING MD	0.975	-0.184	-0.020	-0.291	0.337	-0.027	-0.175	-0.248	-0.009*	-0.075
NV CLARKSBURG WV	0.000	0.036	0.002*	0.022*	-0.009*	0.005*	0.007*	0.030	0.004*	0.011*
SV CHARLESTON WV	0.000	0.014*	-0.014*	0.019*	0.004*	0.002*	-0.002*	-0.019*	-0.009*	-0.011*
VA APPALACHIA VA	0.794	-0.029	-0.020	-0.062	0.083	-0.006*	-0.049	-0.037	-0.005*	-0.013
EK HAZARD KY	0.951	-0.050	-0.011*	-0.153	0.171	-0.013	-0.024	-0.109	-0.006*	-0.027
TN CLINTON TN	0.935	-0.127	-0.123	-0.469	0.551	-0.060	-0.150	-0.313	-0.051	-0.081
AL CORDOVA AL	0.936	0.029*	0.098	0.769	-0.877	0.045*	-0.037*	0.494	0.037*	0.016*
WK CENTRAL CITY KY	0.000	-0.053*	0.016*	-0.015*	-0.046*	-0.005*	0.013*	-0.041*	0.000*	-0.018*
IL CENTRALIA IL	0.136	0.086	0.017*	0.120	-0.097	-0.006*	0.099	0.098	0.011*	0.038*
IN HUNTINGBURG IN	0.728	-0.069	-0.006*	-0.088	0.090	-0.006*	-0.202	-0.112	-0.008*	-0.031*
IA OTTUMWA IA	0.149	0.000*	0.000*	0.000*	0.000*	0.000*	-0.038	0.000*	0.000*	0.000*
MO CLINTON MO	0.883	-0.301	-0.079*	-0.366	0.474	-0.033*	-1.344	-0.899	-0.038*	-0.135
KS PITTSBURG KS	0.941	1.157	0.219*	1.264	-1.535	0.104*	4.431	2.916	0.190*	0.561
AR RUSSELVILLE AR	0.803	-0.339	-0.014*	0.291	-0.436	-0.002*	-0.991	0.005*	-0.009*	-0.171
LA SHREVEPORT LA	0.000	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*
OK TULSA OK	0.558	-0.210	0.001*	-0.339	0.623	-0.037*	-0.161	-0.166	-0.021*	-0.108*
TX CORSICANA TX	0.000	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*
ND WILTON ND	0.150	0.000*	0.000*	0.002	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*
SD LEMMON SD	0.150	0.000*	0.000*	-0.001	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*
EM SIDNEY MT	0.000	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*
WM BILLINGS MT	0.588	-0.004*	0.000*	-0.057	0.066	0.000*	-0.051	-0.080	-0.002*	-0.001*
NW GILLETTE WY	0.000	0.000*	0.000*	0.009	-0.009*	0.000*	0.009	0.012	0.001*	0.000*
SW ROCK SPRINGS WY	0.026	-0.031*	0.024*	0.032*	-0.006*	0.000*	-0.074	-0.004*	0.000*	-0.015*
CN HAYDEN CO	0.000	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*
CS CARBONDALE CO	0.375	0.594	0.330*	0.915	-1.001	0.236*	0.765	0.990	0.140*	0.407
UT SUNNYSIDE UT	0.723	0.034*	-0.096	0.091	-0.305	-0.006*	0.100	0.169	-0.012*	0.024*
AZ WINSLOW AZ	0.000	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*
NM GALLUP NM	0.760	-0.054	-0.001*	-0.164	0.184	-0.015*	-0.093	-0.177	-0.007*	-0.027*
WA CENTRALIA WA	0.831	-0.130	-0.011*	-0.284	0.456	-0.031*	-0.038*	-0.214	-0.013*	-0.064
AK HEALY AK	0.000	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*
	0.964	0.001	0.001	0.006	-0.006	0.000	0.002	0.005	0.000*	0.001
NORTHERN APPALACHI	0.000	0.004*	0.000*	-0.013*	0.013*	-0.005*	0.008*	0.003*	0.002*	0.005*
CENTRAL APPALACHIA	0.921	-0.025	-0.019	-0.084	0.107	-0.008*	-0.025	-0.071	-0.009*	-0.021
SOUTHERN APPALACHI	0.936	0.029*	0.098	0.769	-0.877	0.045*	-0.037*	0.494	0.037*	0.016*
MIDWEST	0.040	0.007*	0.012*	0.030	-0.039	-0.006*	0.005*	0.007*	0.003*	0.005*
WEST INTERIOR	0.596	-0.014	0.000*	-0.015	0.028	-0.002*	-0.017	-0.007	-0.001*	-0.007
NORTH GREAT PLAINS	0.852	-0.004	0.002	-0.002	0.005	0.000*	-0.010	-0.007	0.000*	-0.002
ROCKIES & SOUTHWES	0.605	0.130	0.039*	0.192	-0.283	0.046*	0.182	0.233	0.025*	0.092
NORTHWEST & ALASKA	0.830	-0.062	-0.005*	-0.134	0.216	-0.014*	-0.018*	-0.101	-0.007*	-0.030
APPALACHIA	0.560	-0.014	-0.008	-0.020	0.031	-0.004*	-0.017	-0.022	-0.004*	-0.012
INTERIOR	0.000	-0.001*	0.007*	0.012*	-0.011*	-0.004*	-0.004*	0.001*	0.002*	0.000*
WEST	0.532	0.025	0.010*	0.039	-0.055	0.010*	0.031	0.044	0.005*	0.018
EAST OF MISS. RIVE	0.117	-0.010	-0.004*	-0.011	0.018	-0.005*	-0.013	-0.016	-0.002*	-0.009
WEST OF MISS. RIVE	0.506	0.018	0.008*	0.029	-0.039	0.008*	0.022	0.034	0.004*	0.014

References

Doctor P.G. (1989) "Sensitivity and Uncertainty Analyses for Performance Assessment Modeling," Engineering Geology, 26(1989): pp. 411-429.

Iman R.L. and Conover W. J. (1982), "A Distribution-Free Approach to Inducing Rank Correlation Among Input Variables", Commun. Stat., B11(3): pp. 311-334.

Iman R. L. and Helton J. C. (1985), "A Comparison of Uncertainty and Sensitivity Analysis Techniques for Computer Models", Sandia National Laboratory.

Iman R. L. and Helton J. C. (1987), "An investigation of Uncertainty and Sensitivity Analysis Techniques for Computer Models," Risk Analysis, 8(1), pp. 71-90.

Iman R. L. and Helton J. C. (1991), "The Repeatability of Uncertainty and Sensitivity Analysis for Complex Probabilistic Risk Assessments," Risk Analysis, 11(4), pp. 591-606.

McKay M.D., Conover W. J. and Beckman, R. J. (1979), "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code", Technometrics 21(2): pp. 239-245.

Morgan M. G., and Henrion M (1991), Uncertainty, A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis, Cambridge University Press.

Short Run Forecasts and Long Run Dynamics: A Case Study Using Nitrogen Fertilizer Prices

Harry Vroomen, Economic Research Service and Frederick Joutz, Energy Information Administration and the George Washington University

The purpose of this paper is to compare alternative model specifications used in short run forecasts when serial correlation is present. The "textbook" correction for first-order serial correlation recommends the Cochrane-Orcutt procedure. However, this approach imposes restrictions on the long-run dynamics which can lead to a misspecified model and suboptimal forecasts. In addition, the textbook model produces biased confidence intervals. We demonstrate how to incorporate the long run dynamics into the short run model.

U.S. fertilizer prices have been highly variable since the mid-1970's and exhibited increased volatility since 1980. The prices of nitrogen materials, in particular, have varied significantly over this period. For example, the retail price of anhydrous ammonia fell 38 percent from May 1984 to October 1986, but rose nearly 29 percent by April 1989. Prices then changed direction again, falling 20 percent by October 1990, before climbing 18 percent by April 1993. Prices of other nitrogen fertilizer materials have followed a similar pattern (Figure 1).

This variability complicates the planning process for fertilizer suppliers and users. The USDA provides fertilizer price forecasts through the Situation and Outlook Program on agricultural inputs. Input manufacturers need to forecast fertilizer prices to plan production levels and decide on contract terms for future delivery. Similarly, accurate forecasts can assist farmers in making informed decisions with respect to crop mix and the timing of fertilizer purchases. Accurate fertilizer price forecasts can help reduce costs, maximize profits, and foster the efficient operation of the market.

This study compares five short-run price forecasting models for the retail prices of anhydrous ammonia (AA), urea (UREA), nitrogen solutions (NS), ammonium nitrate (AN), and ammonium sulfate (AS). These materials account for 97.3 percent of the single-nutrient nitrogen fertilizer used in the United States and 77.4 percent of the total nitrogen applied (Tennessee Valley Authority). Price forecasts are generated for the spring, the peak demand season for fertilizer.

The paper is organized into five sections. First, the model is explained. Then, issues relating to serial correlation and common factor restrictions are discussed. Third, we present the procedure for developing the forecasts. This is followed by the forecast comparisons. The conclusion is the fifth section.

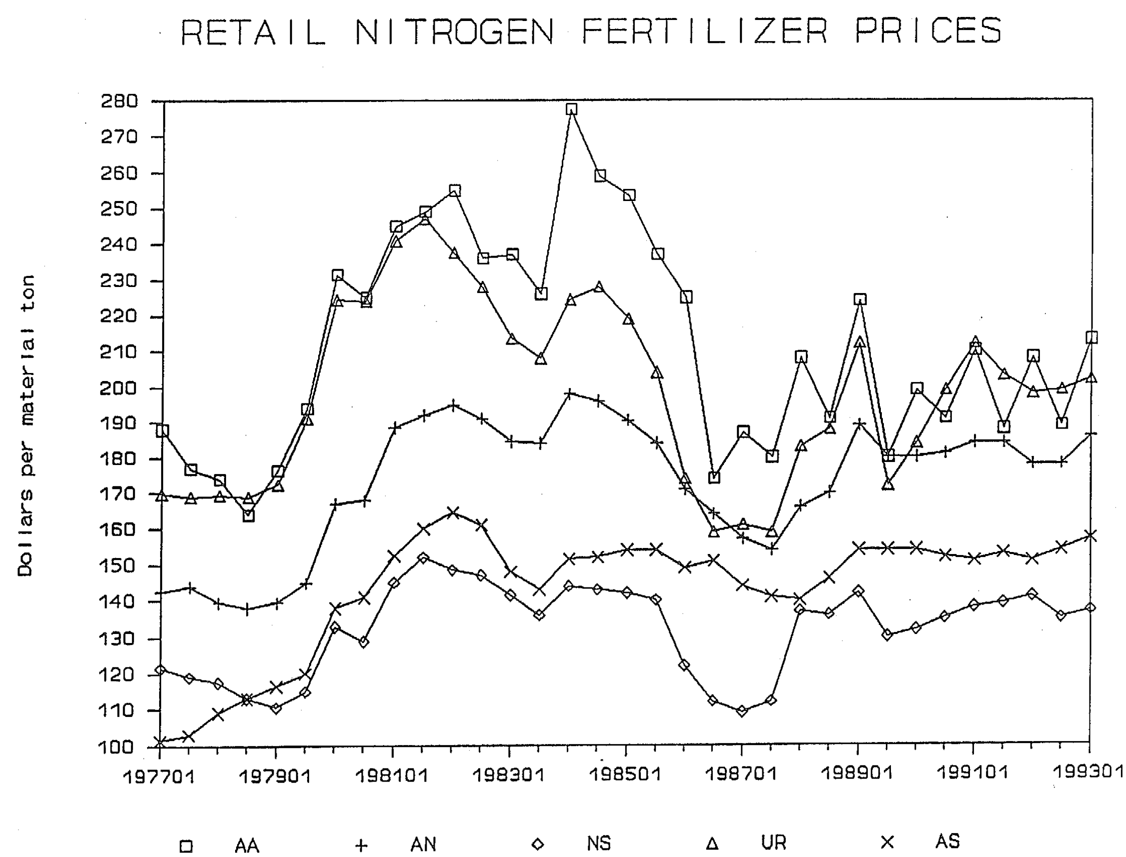
The Model

We begin by assuming that retail prices are approximated by a linear relationship with the wholesale price for AA and transportation costs. AA is the source of nearly all nitrogen fertilizer used in the United States (Andrilenas and Vroomen, 1990). It may be applied directly to the soil or converted into other nitrogen fertilizers, such as UREA, NS, AN, and AS. The retail prices of these nitrogen materials would reflect price changes at the wholesale level if retailers followed a markup pricing scheme. However, retail prices should also be affected by changes in marketing costs not reflected in f.o.b. prices. Transportation costs represent a significant share of the final price a farmer pays for fertilizer, and rail is the predominant method of transporting fertilizer in the United States. Marketing costs are thus represented by the cost of rail transportation. Consequently, the retail price model for each nitrogen material is specified as:

$$RPN_{i,t} = \beta_0 + \beta_1 WHOLAA_t + \beta_2 RAIL_t + e_t \qquad (1)$$

where:

RPN =	retail price of nitrogen material i
WHOLAA =	wholesale price of AA
RAIL =	total rail freight rate index (Dec. 1984=100)
i =	AA, UREA, NS, AN, and AS, respectively
$e_{i,t}$ =	a stochastic disturbance term.



RETAIL NITROGEN FERTILIZER PRICES
Dollars per material ton
280
270
260
250
240
230
220
210
200
190
180
170
160
150
140
130
120
110
100
197701
197901
198101
198301
198501
198701
198901
199101
199301
AA
AN
NS
UR
AS

Figure 1

The data was collected from various issues of the U.S. Department of Agriculture publication, Agricultural Prices. From 1977 to 1985, retail fertilizer prices were reported for March, May, October, and December. Since 1986, however, retail prices have only been available for April and October. To form a continuous data set, March and May retail prices were averaged to construct an April price for years preceding 1986, while reported prices for April were used for subsequent years. Consequently, the retail price equations in (1) were estimated with biannual data (April and October) for 1977-92.

Table 1 contains the summary statistics from the OLS regressions for the retail price equations over the sample period 1977 through 1987.

All of the structural parameters have the hypothesized signs and all but two are significant at the 95-percent level. Four of the five autoregressive parameters are also statistically significant at 5-percent, the fifth is significant at 6-percent. In addition, R^2's indicate that the explanatory variables capture most of the variation in the retail prices of the nitrogen fertilizer materials. Preliminary results indicated that the Durbin-Watson statistic was significant in all the retail price regressions.

Liu (1987) argues that ignoring serial correlation of the residuals can greatly reduce the forecasting performance of a multi-variable model. All equations in (1) were thus estimated with a maximum-likelihood autoregressive technique. Autoregressive techniques use the time series part of a model as well as the systematic part in generating predicted values and so are useful forecasting tools.

Common Factor Restrictions And Serial Correlation Correction

This section discusses the implication of correcting for serial correlation in estimating models and making forecasts. We start by describing three maintained models: 1) a static model, 2) a model with AR(1) errors, and 3) a first order autoregressive distributed lag model, ADL(1,1). The appendix discusses the case where there are lagged endogenous variables in the original model and the AR and ADL models have longer lags. Hendry and Mizon (1978) addressed the issues related to autocorrelation and dynamic specification with a case study on the demand for money in the U.K.

The first two maintained models typify the conventional approach to econometric modeling. Gilbert (1990) refers to this as the average economic regression (AER) view of econometrics. The researcher uses a specification known to "be correct", our equation (1), because it is derived from theory. In the single equation case the static model is specified as:

$$y_t = \beta_0 + x_t\beta + \epsilon_t \tag{2}$$

where x_t is a row vector with k elements and ϵ_t is a random disturbance. The objective is to obtain estimates $b_0,\ldots,b_k$ of the coefficient vector $\beta_0,\ldots,\beta_k$. Several issues, among them, serial correlation, heteroscedasticity, multicollinearity, stationarity vs. differencing, simultaneity, and seasonality confront the applied researcher. These issues present themselves through low Durbin-Watson statistics, insignificant coefficients, wrong signs, extremely large t-values on the lagged dependent variable, etc. The researcher attempts to resolve these problems by respecifying the model and collecting more data.

In the process, variables are added and deleted from the model until the correct signs appear, variables are significant, first order autocorrelation disappears, and the explanatory power measured by $\bar{R}^2$ is relatively high. In the presence of a low Durbin-Watson statistic, the typical correction or respecification is:

$$\begin{aligned} y_t &= \beta_0 + x_t\beta + \epsilon_t \\ \epsilon_t &= \rho_1\epsilon_{t-1} + u_t \quad \textit{where } u_t \sim (0,\sigma_u^2) \textit{ and } |\rho_1|<1 \end{aligned} \tag{3}$$

The Cochrane-Orcutt procedure or maximum likelihood technique with grid search is often used to estimate the parameters. The autocorrelation coefficient is often referred to as a nuisance parameter necessary for efficient estimation of the $\beta's$. It is used to patch up the original theoretical model, because of "bad" test statistics.

There are three issues to consider in this case. First, the error terms may be autocorrelated of a higher order, AR(p). Further tests can be conducted using the Lagrange multiplier approach to detect the number of lags. Second, the model is not the true one. We shall assume that the relevant variables have not been omitted to avoid

Table 1
OLS Results for Retail Price Equations, 1977-1987

Retail Price	constant	WHOLAA	RAIL	RBARSQ	SSR	DW-statistic
AA	51.34 (16.47)	0.806 (0.098)	0.766 (0.171)	0.84	3486.0	1.54 (0.059)
UREA	82.4 (19.4)	0.794 (0.114)	0.213 (0.202)	0.72	4829.0	1.34 (0.019)
NS	70.14 (10.34)	0.342 (0.061)	0.198 (0.107)	0.66	1374.5	1.07 (0.002)
AN	60.53 (10.58)	0.371 (0.062)	0.742 (0.11)	0.84	1438.4	1.16 (0.005)
AS	43.83 (9.08)	0.139 (0.054)	0.925 (0.094)	0.85	1059.8	0.81 (0.1e-3)

Standard errors are given in parentheses. The exact Durbin-Watson statistic and probability were computed using SHAZAM 7.0.

this problem. Finally, the dynamic specification of the true model is incorrect. Certain variables are omitted which are themselves autocorrelated. This case, which is the focus here, suggests that the true model contains richer dynamics than that originally specified.

Model specification is sometimes referred to as an art form as much as a science. While it is unlikely that the researcher will finally discover the true dynamic specification, in cases where the order of autocorrelation is low, he can detect misspecification of the model. If the AR(1) model in (3) is maintained, it implies that:

$$y_t = \beta_0*(1-\rho_1) + x_t\beta - x_{t-1}\beta\rho_1 + \rho_1 y_{t-1} + \epsilon_t \tag{4}$$

or rewriting this using backshift notation:

$$(1-\rho_1 L)y_t = \beta_0(1-\rho_1 L) + (1-\rho_1 L)x_t\beta + \epsilon_t \tag{5}$$

A more general model is the autoregressive distributed lag ADL(1,1) where the first number in the parentheses is the number of lags of the dependent variable and the second is the maximum lag on the explanatory variables. This model is written as:

$$y_t = \pi_0 + x_t\beta + x_{t-1}\gamma + \rho y_{t-1} + u_t \tag{6}$$

The difference between the ADL(1,1) and the AR(1) models is that the former do not place (nonlinear) restrictions on the coefficients. The AR(1) expression in backshift notation reveals the common factor restrictions.

If the common factors are correct, then there is a direct relationship between the autocorrelation coefficient, the coefficient on the lagged dependent variable, and the coefficients on the lagged explanatory variables. The relationship imposes the restriction(s) $\gamma = -\beta*\rho$. This can be tested for using the procedure suggested by Sargan (1964) which is a Wald test. Unfortunately, there is no exact test like the t-test as in the linear restriction case. In finite samples, LaFontaine and White (1986) find that identical algebraic expressions for the (nonlinear) restrictions do not lead to unique statistical conclusions.

A general approach to testing for the common factor restrictions, or COMFAC analysis, is the asymptotic F-test or the Likelihood Ratio test. In finite samples, with lagged dependent variables and or nonlinearity in the parameters, the test is not exact. The statistic for testing the AR(1) model against the ADL(1,1) model is given by

$$\frac{(SSR_r - SSR_u)/r}{SSR_u/(n-k-r-2)} \sim F(r,n-k-r-2)$$

where the sum of squared residuals from the AR(1) model are SSR_r, and the sum of squared residuals from the ADL(1,1) are SSR_u. The degrees of freedom in the denominator are equal to the total number of observations, n, the number of variables in the static model, k, the number of additional unique parameters in the ADL model, r, one lagged dependent variable and the intercept term. The degrees of freedom for the test, r, refers to the number of identifiable parameters in the ADL(1,1) model. The number of restrictions, r, is not always equal to k. The Likelihood ratio test is

$$LR = \left[\frac{RSS_r}{RSS_u}\right]^{n/r}; \ where \ -2\ln LR \sim \chi^2_{(r)} .$$

If the null hypothesis is not rejected from either the F-test or Likelihood ratio test, the imposition of the AR(1), common factors restriction, is a convenient restriction and not a nuisance. The model can be estimated more parsimoniously without loss of consistency.

In Table 2 we present the results for the tests of common factors using the LR test and asymptotic F-test. In all five retail price equations, the common factor restrictions from correcting for the first order serial correlation in (2) using the Cochrane-Orcutt procedure is rejected. The alternative model is an ADL(1,1).

We simplified the alternative model by deleting insignificant variables. The results are given in Table 3. The lagged dependent variable provides valuable explanatory power in every equation. The transportation cost variable, RAIL, appears to have been masking the dynamics of the lagged dependent variable. It is only significant in the UREA price equation. The lagged wholesale price, LWHOLAA, only appears in the AA equation. First order

Table 2 Test for Common Factors Restriction, 1977-87				
Retail Price	AR(1) RSS	ADL(1,1) RSS	F-statistic	L.R. statistic
AA	3247.0	1401.3	18.54 (0.68e-4)	17.647 (0.15e-3)
UREA	2225.0	654.97	27.18 (0.72e-5)	25.681 (0.26e-5)
NS	805.58	286.48	22.50 (0.22e-4)	21.712 (0.19e-4)
AN	845.86	296.72	22.81 (0.21e-4)	21.999 (0.17e-4)
AS	583.34	366.33	12.74 (0.49e-3)	9.7701 (0.76e-2)

The p-value is given in parentheses.

Table 3 Autoregressive Distributed Lag Models for 1977-1987									
Retail Prices	constant	WHOLAA	LWHOLAA	RAIL	LRAIL	LRP	RBARSQ	RSS	Durbin's h
AA[1]	34.24 (13.70)	0.575 (0.079)	0.393 (0.087)			0.287 (0.066)	0.93	1240.2	-1.40
UREA	37.84 (11.49)	0.476 (0.067)		-0.276 (0.107)		0.633 (0.079)	0.95	855.9	0.21
NS	20.36 (8.92)	0.236 (0.035)				0.615 (0.078)	0.91	346.02	-0.98
AN	17.78 (7.76)	0.262 (0.0345)				0.707 (0.051)	0.96	349.56	-1.51
AS	17.05 (8.242)	0.119 (0.038)				0.785 (0.061)	0.92	472.17	0.34

1. The lagged retail price for AA enters with a "2-period" lag - one year instead of six months.

serial correlation is not present according to the Durbin's h statistic.

Since we will be constructing a set of rolling forecasts, we reestimated the model for each year. The models continued to reject the common factor restrictions based on an AR(1) specification for the error term. The ADL regression results for the sample period 1977-92 are presented in Table 4 as a simple test for model stability. The coefficients do not change much from the model estimated through 1977, however the model fit appears to decline slightly as all the $\bar{R}^2$'s are lower.

Developing Model Forecasts

Retail price forecasts for the selected nitrogen materials are generated for the Spring, the peak demand season for fertilizer. Forecasts are made for each April from 1988-93. Out-of-sample forecasts are generated 6 months ahead at a time. For example, retail price forecasts for April 1988 were computed from model coefficients estimated with data through October 1987. This rolling horizon procedure was followed six times for each set of equations as the time period for each was sequentially updated until the final set of price forecasts for April 1993 was developed from equations estimated through October 1992. Sequentially updated forecasting incorporates new information in parameter estimates and is the efficient way to use this model.

In order to develop forecasts from the estimated retail price equations, we first need values for the explanatory variables appearing in these equations. The OLS, AR(1), and ADL models all require forecasts of $WPAA_t$ and $RAIL_t$. Forecasts of these variables are developed from autoregressive-integrated-moving-average (ARIMA) models.

The U.S. wholesale price of AA is determined at New Orleans, near most of the domestic ammonia capacity; thus, prices at this market are used in the study (Tennessee Valley Authority, July, 1989). F.o.b. prices, U.S. Gulf, are available for AA starting in February 1977 from Green Markets (McGraw-Hill, Inc., 1990). Following the iterative technique of identification, estimation, and diagnostic checking popularized by Box and Jenkins (1970), the series WHAA was identified as an SARIMA(2,1,0)(1,0,0):

$$(1-\phi_1 L-\phi_2 L^2)(1-\phi_{12} L^{12})(1-L)WHAA_t = e_t$$

where L is the backshift operator, ϕ are autoregressive parameters, and $\{e_t\}$ is a sequence of white noise. Maximum-likelihood estimation of this model over the period February 1977-October 1987 yields:

$$\phi_1 = 0.686(0.088) \qquad \phi_2 = -0.179(0.088) \qquad \phi_{12} = 0.240(0.090),$$

where the quantities shown in parentheses are standard errors. Parameter estimates are all statistically significant at the 95-percent level and lie within the bounds of invertibility. Diagnostic checks do not reveal any serious model inadequacies and the residuals of the model do not differ from white noise as indicated by the Q-statistics of 3.76 and 17.78, which are not significant when compared to a χ^2 statistic with 9 and 21 degrees of freedom, respectively (Box and Pierce, 1970). No intercept was used because it did not differ significantly from zero.

Similarly, using biannual data (April and October), the ARIMA model for the RAIL series was identified as an ARIMA(1,1,0):

$$(1-\phi_1 L)(1-L)RAIL_t = \delta + \epsilon_t$$

Data for the total rail freight rate index is available from the U.S. Department of Labor. Maximum-likelihood estimation of this model over the period April 1977-October 1987 yields:

$$\delta = 1.894(1.109) \qquad \phi_1 = 0.668(0.167)$$

Diagnostic checks reveal no serious model inadequacies. The Q-statistic of 5.43, which is not significant when compared to a χ^2 statistic with 11 degrees of freedom, indicates that the residuals of the model do not differ from

Table 4
Autoregressive Distributed Lag Models for 1977-1992

Retail Prices	constant	WHOLAA	LWHOLAA	RAIL	LRAIL	LRP	RBARSQ	RSS	Durbin's h
AA_1	53.73 (15.02)	0.513 (0.089)	0.306 (0.096)			0.299 (0.081)	0.86	3454.6	-0.22
UREA	45.07 (16.72)	0.456 (0.087)		0.101 (0.123)		0.462 (0.103)	0.82	3218.7	1.35
NS	28.71 (12.73)	0.158 (0.043)				0.646 (0.102)	0.72	1283.9	1.75
AN	20.46 (10.90)	0.178 (0.041)				0.768 (0.066)	0.87	1174.0	1.61
AS	16.72 (7.65)	0.08 (0.029)				0.83 (0.051)	0.90	653.14	1.24

1. The lagged retail price for AA enters with a "2-period" lag - one year instead of six months.

white noise.

The ARIMA models for WHAA and RAIL, estimated with data through October, are used to forecast these variables for the following April. Retail price forecasts from the OLS models are then developed by incorporating the forecasts for $WPAA_t$ and $RAIL_t$ into the respective equations for AA, UREA, NS, AN, and AS. Additional information is required to develop price forecasts from the AR(1) and ADL models. In addition to forecasts of $WPAA_t$ and $RAIL_t$, the AR(1) models require a lagged residual, e_{t-1}, which is available at time t. Similarly, the ADL models require lagged values of WPAA, RAIL and the dependent variable. These values are also available at time t.

Evaluation of Forecasting Models

Tables 5 and 6 provide summary statistics for the three models: ordinary least squares (OLS), Cochrane-Orcutt (AR1), and Autoregressive Distributed Lag (ADL). We report the mean percent errors (MPE), the root mean square error (RMSE), and the mean absolute percent error (MAPE). The statistics are for the six one step ahead forecast errors. The unconditional forecast results are given in Table 5 and the conditional forecast results are in Table 6.

The results are mixed as to which model produces the best forecast. However the ADL model appears to perform the worst of the three models. It produces the highest MPE in three of the five price equations using the conditional and unconditional technique. Similarly, the RMSE from the ADL model is also the highest in four of the five equations. The ADL model, while statistically superior, does not produce numerically better forecasts.

Conclusion

We test models used in predicting five nitrogen fertilizer product prices. The original model specified was a simple static model. It was found to have errors which exhibited first order autocorrelation. Then, we compared the "textbook" approach of correcting for this problem using the Cochrane-Orcutt procedure against the more general ADL model. The later was found to reject the common factor restrictions implied by the former. When the three models one step ahead forecasting power was compared there is no clear winner. We expected the ADL model to perform best, but it appeared to perform the worst. This may be due to several reasons: 1) the forecast horizon was restricted to one period, 2) the sample is too small, meaning the results are not significantly different, 3) this is a case where a well fitting model is not a good forecasting model.

Table 5
Comparison of Unconditional Forecasts 1988-1993

	Mean Percent Error			RMSE			MAPE		
Retail Prices	OLS	AR(1)	ADL	OLS	AR(1)	ADL	OLS	AR(1)	ADL
AA	1.62	1.73	8.51	1.94	2.52	4.23	2.01	2.33	8.51
UREA	8.56	1.49	6.50	4.62	6.53	7.72	5.86	5.08	8.18
NS	5.73	4.08	7.97	2.66	1.52	5.14	5.73	4.08	7.97
AN	1.50	0.01	5.77	3.89	3.60	3.55	3.44	2.59	5.81
AS	-2.34	-4.07	2.03	4.32	3.23	1.94	3.05	4.07	2.03

Table 6
Comparison of Conditional Forecasts 1988-1993

	Mean Percent Error			RMSE			MAPE		
Retail Prices	OLS	AR(1)	ADL	OLS	AR(1)	ADL	OLS	AR(1)	ADL
AA	-1.25	-1.30	6.61	5.45	5.78	8.12	4.44	4.70	7.78
UREA	3.39	0.15	0.50	5.41	3.78	6.59	4.97	2.76	5.33
NS	4.45	3.11	7.21	5.06	2.09	6.61	4.62	3.11	7.21
AN	0.38	0.53	5.17	4.35	2.67	5.31	3.09	2.07	6.11
AS	-2.84	-3.91	1.71	3.29	1.80	2.66	3.21	3.91	2.32

References

P. Andrilenas and H. Vroomen, "Fertilizer," in Seven Farm Input Industries, Ag. Econ. Report No. 635, US Dept. Agr., Econ. Res. Serv., September 1990, pp 14-29.

G.E.P. Box and G.M. Jenkins, "Time Series Analysis, Forecasting and Control," Holden-Day, Inc., San Francisco, CA, 1970.

G.E.P. Box and D.A. Pierce, "Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models", Journal of the American Statistical Association, 65, 1509-1526(1970).

Gilbert, Christopher L.; 1990, "Professor Hendry's Econometric Methodology," Modeling Economic Series: Advanced Texts in Economic Series, ed. C.W.J. Granger, Oxford Press, chapter 13.

Hendry, David F. and Grayham E. Mizon; 1978, "Serial Correlation as a Convenient Simplification, not a Nuisance: A Comment on a Study of the Demand for Money by the Bank of England," The Economic Journal, 88, 549-563.

Lafontaine, F. and Ken White; 1986, "Obtaining any Wald Statistic You Want," Economics Letters, vol. 21, 35-40.

Lon-MU Liu, "Sales Forecasting Using Multi-equation Transfer Function Models," Journal of Forecasting, 6, 223-238(1987).

McGraw-Hill, Inc., Green Markets: Fertilizer Market Intelligence Weekly, New York, October 29, 1990, and earlier issues.

Sargan, J.D.; 1964, "Wages and Prices in the United Kingdom: A Study in Econometric Methodology," in Harts, Mills, and Whitaker (eds). Econometric Analysis for National Economic Planning. London: Butterworth, pp. 22-54.

Tennessee Valley Authority, Economics and Marketing Staff, "North American Fertilizer Capacity Data," Muscle Shoals, AL, July 1989.

Tennessee Valley Authority, Economics and Marketing Staff, "Commercial Fertilizers," Muscle Shoals, AL, December, 1989 and earlier issues.

U.S. Department of Labor, Bureau of Labor Statistics, Producer Price Indexes, monthly reports, 1977-92.

U.S. Department of Agriculture, National Agricultural Statistics Service, "Agricultural Prices," Washington, DC, various annual summaries, 1977-92, and April, 1993.

White, Ken; 1993, SHAZAM User's Reference Manual Version 7.0, McGraw-Hill.

Modeling and Forecasting U.S. Patent Application Filings

Kay Adams and Gus Mastrogianis, U.S. Patent and Trademark Office

Introduction

This paper summarizes the results from a research project to develop a methodology for predicting patent application filings at the U.S. Patent and Trademark Office (USPTO). Like many government agencies worldwide, the USPTO has switched to a fee-based funding environment; thus the need for reliable forecasts of revenue and resource allocation. Also, the USPTO is engaged in a long-term international cooperative effort with the European and the Japanese Patent Offices (EPO and JPO respectively) to track and project patent trends.

The research focus to date has been on aggregate patent application filings. We developed and compared naive, ARIMA, and econometric forecasting models at the one and two-year ahead horizon using annual and quarterly time series data. The Mean Error (ME), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE) of the models from within-sample forecasting comparisons, suggest that the annual and quarterly econometric models are more reliable than the other two types of models. Both the annual and quarterly econometric models consistently outperform the naive and ARIMA models in predicting the actual patent application filing levels whether it is one-year ahead or two-year ahead within-sample forecasting.

The paper is organized into six sections. First, we present a review of the literature on patent application filings and related economic activities (e.g., the relationship between R&D and patenting). This motivates the construction of the econometric models. Description of the data used in the research is presented in the second section. The annual and quarterly empirical models are presented in sections three and four, respectively. Then, the within-sample unconditional forecast at the one and two-year horizons are compared. The final section concludes and discusses additional likely future research activities at the USPTO.

I. Patent and Macroeconomic Activity

Griliches (1990) stated that a patent represents an output of the inventive process having passed a two-part test of the economic threshold. First, the invention has been deemed worthy of an initial investment of resources by the inventor or the business organization. And once developed, an expectation of some level of utility or marketability is also tied to the invention. Thus, patent application filings represent an important measure of successful inventions meeting an expected economic value of the patent right.

In acknowledging the distinction between the process and output of innovation, Acs and Audretsch (Germany, 1989) provided empirical evidence which suggests that patents do provide a good, albeit imperfect, measure of innovative activity. Tempering this view somewhat, Griliches (1989) argued that patents (grants) are not a "constant yardstick indicator of either inventive input or output" (p. 291). Given that all inventions do not result in patents, that not all inventions are patentable, and that patents vary considerably in their economic impact, the problems associated with using patents as the sole measure of inventive output are easily recognizable. However, Griliches does contend that the aggregate count of patents may serve as a measure of shifts in technology.

Patents are intrinsically linked to the larger economic picture -- to the process of innovation and to technological and scientific change. Based on a review of domestic patent application filings from the Civil War to World War II, Schmookler (1954) concluded that industrial invention is economically caused. In his view, the level of inventive activity is represented by such factors as the number of technical trained workers, industrial (variable and not fixed) inputs, and the Gross National Product (p. 188). Schmookler also observed that "changes in the level of inventive activity cause changes in the number of patent applications filed" (p. 186).

Scherer (1983) tested the relationship between the R&D and patenting activity. He believed that the propensity to patent can be measured by examining the number of patents industrial corporations obtain per million dollars of company-financed research and development expenditures (p. 108). When expressed as a ratio, the measure becomes the number of patents to R&D expenditures.

Scherer found evidence which suggests that, within industries, patents rise in connection to R&D. Taken from 1976 and 1977, the survey covered 443 corporations and 15,112 U.S. patents originating from U.S. resident inventors and issued to the same corporations. This research supports the general view that R&D is a proxy measure of the inputs to the inventive system, while patents represent a qualified proxy measure of the output from the inventive process (Acs and Audretsch, 1989).

Based on the theory that R&D expenditures and patenting activities are positively related, Scherer's findings also showed:

1. patenting is closely associated with formally organized R&D activity; and
2. of the industries studied, approximately 75 percent had R&D expenditures with either constant or increasing rates of returns as demonstrated by patent grants (p. 111-117).

Pavitt (1982) supports the proposition that a strongly positive relationship exists between R&D and patenting activities. Scherer found evidence to support a lag of approximately nine months from the time of invention to the filing of a patent application (p. 108). In a related study, Bosworth and Westaway (U.K., 1984) recognized the existence of a lag between expenditures on R&D and the filing of a patent application. The model constructed for their study showed the time lag to be approximately one year. From this research, the authors proposed that a relatively high percentage of R&D programs yields fairly immediate results in the form of patents, typically on the order of one year (p. 144).

A study by Hall, Griliches, and Hausman (1986) focused on the analysis of the relationship between patenting and R&D activity at the firm level during the 1970s in the U.S. manufacturing sector. Based on previous research, recognition is given to the theory that successful research and development efforts generally lead to both patents and to the further commitment to invest in development. The authors saw annual R&D expenditures as investments which add to a firm's stock of knowledge. As they describe it, the stock of knowledge depreciates over time, rendering the contribution of older R&D less valuable. Using this assumption, they attempted to use patent application filings as a possible indicator of the value added to the knowledge base in any given year. Additionally, their study concluded there is a lag structure from R&D to patents, however, the timing of the lag remained uncertain.

Several other economic variables also have been shown to demonstrate evidence of relatedness to patents. For example, Griliches (1989) compared domestic patent applications with economic variables, such as real Gross National Product (GNP) and Gross Private Domestic Investment (GPDI). He concluded:

1. Fluctuations in R&D do affect the number of patent applications applied for, but less than proportionately, and to the extent testable, the causality runs from R&D to patents (p. 309-311).

2. An estimated 10 percent increase in defense expenditures causes a 5 percent decline in domestic patenting, implying that defense spending takes away resources from patentable research (p. 310).

3. The vast majority of the systematic short-run variability in aggregated domestic patenting is explained by the R&D and national defense R&D variables, and virtually none of the other economic variables (included in this study) make any significant contributions on their own (p. 311).

II. U.S. Patent and Macroeconomic Data

The USPTO patent applications filing data base was constructed by the Office of Long-Range Planning and Evaluation. The research goal was to forecast patent application filings by fiscal year. Most of the patent application filing data were readily available. However, some of the data formats varied. Patent application filing data differed by reporting periods or increments (e.g., weekly, biweekly, calendar, and fiscal year) and patent type/category (i.e., utility, plant, reissue, and design). In some cases, patent application filing data, most notably quarterly filings, were estimated using several past data reporting structures to derive the fiscal year utility, plant, and reissue (UPR) data format.

1. Annual USPTO Patent Application Filings

Annual patent application filings were available for total, domestic origin, and foreign origin. The amount and format of the patent application filing data varied among the three annual data sets. Total annual UPR patent application filings were available from fiscal year 1949 to fiscal year 1991 from the Commissioner of Patents and Trademarks Annual Reports. See Figure 1. Total annual patent application filings were consistently maintained as utility, plant, and reissue data since fiscal year 1949.

However, one inconsistency found in total annual UPR application filings was the switch to a new basis for a fiscal year. With the adoption of a new fiscal year in 1976, the year shifted forward by one quarter. Since quarterly data were not available prior to 1976, the switch was inconsequential for the quarterly data set. But the lack of quarterly data prior to that time made it impossible to adjust annual total UPR patent application filings to this change. Therefore, the years 1949 to 1976 are under the prior fiscal year basis (July to June) and 1977 forward are under the present fiscal year basis (October through September). The major reason for this concession was to have as many annual total UPR application filing observations as possible for the development of forecasting models.

In fiscal year 1949, patent application filings totaled nearly 66,000; by 1991 annual filings were approximately 168,000. Over the full sample period, the average annual growth rate was 2.06 percent. There is a clear upward trend to the data with a noticeable increase in the 1980s. During the 1980s, the average growth rate was in excess of 4 percent or about 2.5 percent higher than for the 1950s, 1960s, and 1970s. In fiscal year 1983, there was a decrease in filings of about 20,000 from 1982, because of an announced sharp increase in the patent application filing fees. The actual sample for the identification, estimation, and forecasting periods is 1956 through 1991. This was done because of data limitations for the other variables in the econometric models. Dropping the first seven observations did not significantly affect the ARIMA model's performance.

2. Quarterly USPTO Patent Application Filings

A quarterly data set for total UPR patent application filings was constructed from weekly, biweekly, and monthly total patent application filing data. A data set for the first quarter of 1976 through the third quarter of 1992 was transcribed from dozens of paper reports, patent application filing data varied by patent type and reporting periods. A process was developed at the USPTO to estimate total quarterly UPR patent application filings from the paper reports.

The first step of the process was to identify which quarters should be constructed since there are currently two time periods that can be used to construct quarterly UPR patent application filing data at the USPTO. One is based on the first and last day of the patent production process, and the second is based on estimating the first and last day of each fiscal quarter using monthly patent application filing data. There is a considerable degree of variability in the number of days in each quarter based on the patent production process. For instance, the cutoff dates for the end of each quarter in the patent production process may occur as many as fourteen days before the last day in a fiscal quarter or may start as early as fourteen days before the first day of a fiscal quarter.

Since the research includes quarterly economic data, which are based on the first and last day of a calendar quarter and easily converted to a fiscal quarter, total quarterly UPR application filing data were constructed based on estimated calendar monthly patent application filings. Total quarterly patent application filing data based on monthly application filings are consistent with calendar quarters. Total UPR, UPR&D (aggregate of UPR and design patents), and utility weekly and biweekly patent application filings were used to estimate monthly UPR application filings. Total monthly UPR application filings were then used to estimate the total quarterly UPR patent application filings.

Total quarterly UPR application filings are shown in Figure 2. There are several noteworthy characteristics of the data. First, the number of filings prior to 1982 was fairly constant. In 1982Q4, there was a big increase due to the dumping or "surge" of filings prior to the first and largest patent application filing fee increase since 1966. Starting in 1984, there is an upward trend to the data that appears to slow down in 1991-1992. The increase is from about 26,000 filings per quarter to 42,000 filings per quarter by 1992. Another notable feature of the data series is the saw-toothed pattern between consecutive observations around the average or trend.

3. USPTO Patent Application Filing Fees

In addition to patent application filing data, patent application filing fees prior to fiscal year 1965 were also needed for this study. Basic utility and reissue patent application filing fees were used to create a fee time series from fiscal years 1932 to 1965 and the first quarter of fiscal year 1976 to the third quarter of 1992. After fiscal year 1982, the filing fee data set reflects fees charged to large entities for utility and reissue patent application filings. Actual dates of filing fee changes (i.e., the dates when new filing fees became effective) were identified in U.S. Public Law documents and the Commissioner of Patents and Trademarks Annual Reports.

Patent application filing fee data were constructed for only utility and reissue application filings, which accounted for 93.95 percent of total UPR&D application filings in fiscal year 1991. From fiscal years 1949 to 1991, utility and reissue patent application filings have averaged 93.34 percent of total UPR&D application filings. It should be noted that filing fees for plant patents were excluded because plant fee data were not identified for fiscal years prior to 1965, and the application filing fee for plant patents is lower than fees for utility and reissue patents. Also, plant patent application filings only accounted for 0.23 percent of total USPTO patent application filings in fiscal year 1991 and averaged 0.16 percent of total filings since fiscal year 1965.

From fiscal year 1983 to the present, the USPTO has offered large and small entity filing fees. Only large entity filing fees were used to develop forecasting models. There were several reasons why only the large entity fee was selected. Large entities accounted for 66.61 percent of total patent application filings on average from fiscal year 1983 to fiscal year 1991, and both the large and the small entity fees have changed simultaneously during that period.

Although annual utility and reissue patent application filing fee data were available back to fiscal year 1932, it was only necessary to use the fee data from fiscal years 1949 to 1991 since these are the only fiscal years for which total annual UPR patent application filings are available. For quarterly utility and reissue fee data, a database for only the first quarter of fiscal year 1976 to the third quarter of fiscal year 1992 was created because of the similar constraint imposed by total quarterly UPR patent application filings. Once the fees were compiled, they were adjusted for inflation using the 1987 Gross Domestic Product implicit price deflator to derive real fee changes.

4. Economic and Demographic Data

Annual domestic economic and demographic data were available starting with calendar year 1949. Several of the economic indicators were only available in 1982 U.S. dollars, in nominal U.S. dollars, and in 1987 U.S. dollars. With the exception of Gross Private Domestic Investment (GPDI), the 1987 GDP implicit price deflator (GDP deflator) was selected for converting nominal economic data to real economic data. For GPDI, a separate 1987 GPDI price deflator was identified.

Quarterly domestic and foreign data were available for the majority of economic variables but not for any of the demographic variables (e.g., Scientists and Engineers engaged in R&D and graduates with a Bachelor's in Mathematics, Science, and Engineering). However, the data sets are more complete for the time period studied (1976-1992). One notable exception is the economic variable, R&D Expenditures (both domestic and foreign), that is not reported by quarter from the sources identified.

Table 1 lists all of the variables, their respective sample periods and coverage used in the study.

III. Annual Forecasting Models

1. Naive Model for Annual USPTO Patent Application Fillings

For the purposes of this project, the software package Forecast Pro was used in the expert system mode to select the best naive model. The Holt exponential smoothing model with a linear trend was found to minimize the root mean square error. This model's technical details are discussed below and followed by the empirical results.

In many real world applications, the data are not so simple in their patterns that simple exponential smoothing provides a reliable forecast. For example, the annual data on UPR patent application filings exhibit an upward trend over time. When there is a trend in the data, the simple smoothing forecasting model will have large errors that fluctuate from positive to negative or vice versa. A common method for adjusting the simple smoothing forecasting model for trend is Holt's method, named after its originator C. C. Holt (1957). The exposition below follows Wilson and Keating (1990). Three equations and two smoothing constants are used in the model.

$$F_{t+1} = \alpha X_t + (1-\alpha)(F_t + T_t) \quad (3.1)$$
$$T_t = \beta(F_t - F_{t-1}) + (1-\beta)T_{t-1} \quad (3.2)$$
$$H_{t+h} = F_{t+1} + hT_{t+1} \quad (3.3)$$

where:

F_{t+1} =	Smoothed value for period $t+1$
X_t =	Actual value in period t
F_t =	Smoothed value for period t
T_t =	Trend estimate
α =	Smoothing constant $(0 < \alpha < 1)$
β =	Smoothing constant for trend estimate $(0 < \beta < 1)$
h =	Number of periods ahead to be forecast
H_{t+h} =	Holt's forecast value for period $t+h$
$(F_t + T_t)$ =	Actual forecast of X_t made in period t

Equation (3.1), the one period ahead forecast or smoothed value, adjusts F_{t+1} for a fraction of this period's level, X_t, and for growth of the previous period, T_t, by adding T_t to the smoothed value of the previous period, F_t. The forecast can also be expressed as an adaptive or error learning model.

$$F_{t+1} = (F_t + T_t) + a(X_t - F_t - T_t) \quad (3.4)$$

A one period ahead forecast is equal to the previous periods forecast plus the difference between the current actual value and its forecast.

The trend estimate is calculated in equation (3.2), where the difference of the last two smoothed values is calculated. Because these two values have already been smoothed, the difference between them is assumed to be an estimate of trend in the data. The logic behind β in equation (3.2) is the same logic used in the simple smoothing model for the constant α. The most recent trend, $(F_{t+1} - F_t)$, is weighted by β and the last previous smoothed trend, T_t, is weighted by (1 - β). The sum of the weighted values is the new smoothed trend value T_{t+1}.

Equation (3.3) is then used to forecast h periods into the future by adding the product of the trend component, T_{t+1}, and the number of periods ahead to forecast, h, to the current value of the smoothed data F_{t+1}. This method accurately accounts for any linear trend in the data.

Table 2 contains the results from the Holt exponential smoothing model. The model suggests that a one-year ahead forecast should be about 80 percent of this year's patent application filings plus 20 percent of the forecast error for this year. The Durbin-Watson statistic suggests that consecutive errors are uncorrelated, however, the statistic is not appropriate, because there is no intercept in the model [See Durbin and Watson(1971)]. The Ljung-Box-Pierce test for the correlation of errors through lag 18 is also insignificant.

2. Box-Jenkins (ARIMA) Model for Total Annual USPTO Patent Application Filings

Since patent application filings have a clear upward trend, the series was transformed to natural logarithms to reduce the relative increase from one year to the next. Table 3 provides the mean, variance, autocorrelation function, and partial autocorrelation function for the filings data in levels and first differences of natural logarithms.

The autocorrelation of filings in levels at lag 1 is 0.85, and it decays slowly or exponentially to 0.69 at lag 2, 0.56 at lag 3, and so on. The partial autocorrelation at lag 1 is the same by definition as the autocorrelation at that lag. It is significantly different from zero when divided by the asymptotic standard error 0.17. The magnitude of the autocorrelation at lag 1 and the slow rate of decay are indicative of a nonstationary series with a unit root. The Box-Jenkins procedure recommends transforming the data into first differences of natural logarithms. This has a convenient interpretation, because the data is in annual growth rates.

The second half of the Table 3 presents the results in first differences of natural logarithms. The average growth rate is approximately 2.3 percent with a standard deviation of 5.1 percent. The autocorrelation at lag 1 is -0.19, but the standard error of the autocorrelations is 0.17 indicating that last year's growth rate is weakly correlated (or a marginal predictor) of this year's growth rate. The implied t-ratio is greater than 1.0, but not greater than 1.65 if a 10 percent significance level were chosen. Similarly, autocorrelations at longer lags are not individually significantly different from zero.

The modified Ljung-Box-Pierce statistics are also tabulated in Table 3. The statistics test whether autocorrelations through a given number of lags are significantly correlated with the current value of the series. The statistics are distributed as Chi-square with the number of degrees of freedom, df, equal to the number of lags. The p-value(s) are the probabilities of rejecting the null hypothesis (of no correlation) when the alternative is true. If the values were less than 0.1, there would be some evidence of correlation between that number of lags and the current value. None of the Ljung-Box-Pierce statistics are significant suggesting that a possible ARIMA model would be in first differences without a constant (or drift) term for the growth rate. This specification, ARIMA(0,1,0), is referred to as a random walk in the time series literature. The standard error (deviation) is given below the annual growth rate. The estimate for the growth is insignificant from zero since the t-ratio is less than one.

Table 4 presents results of a test of the specification ARIMA(1,1,0) with a constant. The AR(1) parameter is as expected from the examination of the autocorrelation function. The t-ratio is slightly greater than one, suggesting that it provides marginal additional explanatory power. The constant term is about 2.7 percent, one half a percent higher in the simple random walk with drift model, but within a one standard error bound. The autocorrelation function and the Ljung-Box-Pierce statistic suggest that there is no further autocorrelation in the residuals of the estimated model. R-square is low, 0.0368, but this is not unusual in ARIMA models. At the bottom of the table is a F-test comparing the two models using the residual sum of squares. The F-statistic is 5.26 which suggests the random walk with drift model is rejected at the 5 percent significance level.

Thus, a candidate ARIMA model is the ARIMA(1,1,0) shown below.

$$\begin{aligned}\Delta F_t &= \underset{(0.0095)}{0.02741} - \underset{(0.1684)}{0.1945}\cdot\Delta F_{t-1} + \varepsilon_t\end{aligned}$$

The constant term can be interpreted as a year to year growth rate after the last year's effect on the growth of patent application filings. The negative coefficient on the lagged growth rate suggests that positive above average growth in one year will be followed by below average growth the next.

3. Econometric Model for Annual USPTO Patent Application Filings

The econometric model for patent application filings is based on a simple theory of innovation and economic growth. Firms and inventors invest resources to improve existing technologies and develop new technologies. One observable output of the process is a patent filing. This represents the firms' and inventors' (initial) claim to an invention. A filing is made to both protect the property rights of the claim and to earn rents or royalties from the invention.

It is assumed that the economic theory underlying the patent application filing process can be generalized to the aggregate level. A simple model of UPR patent application filings would assume that filings are a function of past inventions (or filings), the cost or price of a filing, and the resources invested toward inventions. The first component can be captured by including lagged values of filings in the model. The price of a filing is measured by the filing fee paid upon application; lagged or previous fees can be used to derive estimates of both short run and long run price elasticities. This measure is only a proxy, because it omits patent attorney fees. The resources invested for inventions can include both economic and demographic components. Economic variables could be represented by expenditures on research and development (R&D) and public and private investment expenditures. The number of scientists and engineers engaged in R&D can serve as a proxy for the human resources devoted to the process of invention.

The approach for finding an econometric model begins with a general model which is reduced to a parsimonious one via empirical tests of model fit and parameter stability. Below, the model building and testing process is described in two parts. First, an initial set of models are estimated and found to be deficient in terms of parameter stability.

The initial or most general model specified was a single equation autoregressive distributed lag regression (ADL) of filings F_t on filings for the last four years, the current and past filing fees, FEE_{t-i}, over the last four years, and the expenditures on research and development in the United States, ARD_{t-i} over the last four years. The ARD series is constructed as a three year moving average. This series attempts to smooth expenditure and reflects a commitment to research leading to patenting. The length of the moving average was chosen arbitrarily recognizing a potential degrees of freedom problem.

$$F_t = \alpha + \sum_{i=1}^{4} \phi_i \cdot F_{t-i} + \sum_{i=1}^{4} \beta_i \cdot ARD_{t-i} + \sum_{i=4}^{4} \gamma_i \cdot FEE_{t-i} + \varepsilon_t \qquad (3.5)$$

All variables are in natural logarithms. Both the fee and expenditure data were adjusted to 1987 dollars by the GDP deflator. The current fee is used, because the prices were known in advance by patent applicants and current prices can affect filing behavior. The sample period used is from 1956 to 1991.

The fee series can be treated as exogenous. R&D data enter the model through a lag, because current R&D expenditures are unlikely to result in a filing immediately. Also, the filing data are based on fiscal years whereas the R&D expenditure data are only available on a calendar year basis. Thus, using the one-year lag of ARD is effectively a three quarter lag. R&D expenditures are implicitly assumed to be exogenous in the model presented here. This implies that there is no feedback relationship between R&D and the output of patents, and the parameters of interest are invariant to the relationship. There was insufficient current data on scientists and engineers to be useful in the model. The most recent estimates are from 1989.

A more parsimonious model was searched for by reducing the number of lags across all variables and by variables individually. The lag length for the ADL is based on minimizing the Akaike Information Criterion (AIC) and Bayesian Schwartz Criterion (BSC) as recommended in Granger and Newbold (1986) and Judge et. al. (1988). The results are presented in Table 5. The unrestricted model with 4 lags yields an AIC and BSC of -6.85 and -6.19, respectively. It is minimized when there is one lag on filings, one lag on research and development expenditures, and the current and one lag of fees; the AIC is -7.02 and the BSC is -6.80. Other variations of the initial ADL restricting filings to one lag and both filings and fees to one lag were tried and yielded the same result.

Results from the final model are presented in Table 6. The econometric model chosen for forecasting purposes used past filings, a three year moving average of research and development expenditures, and a new variable for the price, LP. The price series was transformed to take account for the impact of the announced fee structure changes in 1983. Prior to 1983, the fee, lagged one and two years, is used and after that period the current and lagged fees were used in the new price series. Tests for autocorrelation resulted in a Durbin's H statistic of 0.84, suggesting the residuals are not autocorrelated at lag one. The Lagrange Multiplier statistics and Ljung-Box-Pierce statistics imply there is no further autocorrelation of the residuals. The tests for heteroscedasticity suggests that there is none present. The Jarque-Bera asymptotic LM normality test implies that the null hypothesis of no skewness and excess kurtosis cannot be rejected.

Sequential Chow Tests revealed a structural break in the model beginning in 1982. The Chow test statistic was calculated as:

$$Chow = \frac{(sse - sse1 - sse2)/k}{see/(n1 + n2 - 2k)}$$

where

sse = sum of squared errors for full sample
$sse1$ = sum of squared errors first part of sample
$sse2$ = sum of squared errors second part of sample
k = number of estimated coefficients
$n1$ = number of observations in first part
$n2$ = number of observations in second part

The statistic is distributed as an F, thus if it is less than the critical value from an $F(k, n1 + n2 - 2k)$, there is no evidence of a structural break. In the tests, the F statistic has 5 degrees of freedom in the numerator and 25 degrees of freedom in the denominator. The critical values are 2.6 at five percent and 3.86 at one percent. Despite this problem, the model will be used for forecasting purposes. Alternative models were tried, but could not encompass this model and in many cases had more severe problems.

The model chosen only includes a single lag of past filings, last year's three year moving average of R&D, and current and last year's real fees.

$$\underset{}{F_t} = \underset{(0.4433)}{0.21665} + \underset{(0.0725)}{0.81199} \cdot F_{t-1} + \underset{(0.0557)}{0.17632} \cdot ARD_{t-1} - \underset{(0.0191)}{0.1226} \cdot FEE_t + \underset{(0.0195)}{0.13268} \cdot FEE_{t-1}$$

The dynamics of the model are interesting. The short run effect of the three year moving average of R&D expenditures suggests that a one percent increase leads to 0.17 percent increase in patent application filings. However, the long run impact of a one percent increase is 0.9378 percent or almost one. This is consistent with research from the literature review. The immediate effect of a one percent increase in filing fees is -0.1226, while the effect from the previous year is a positive 0.1329. The net effect is zero. Patent application filing fees are set by statute. Historically, patent application filing fees have not been set as an impediment to the innovation and invention process. From an economic perspective, patent application filing fees can be viewed as user costs. The negative-positive sign pattern captures both the surge effect of announced fee increases and the declining real fee(s) between fee increases.

A simple adjustment was tried using a dummy variable, SURGE, which takes on the value 1 in 1982, -1 in 1983, and 0 in all other years. This was designed to capture the dumping or surge of filings just before the fee increases. In both cases, the fee increase was announced and passed by Congress as P.L. 97-247. The SURGE was used as a replacement for the FEE series and came up with the same lag specification for filings and R&D. Unfortunately, the model fit was not as "good" as with the FEE series.

IV. Quarterly Forecasting Models

The models of quarterly patent application filings are based on data from the first quarter of 1976 (1976Q1) through the third quarter of 1992 (1992Q3). Quarterly filings data are collected on a fiscal year basis, October 1st to September 30th.

1. Naive Model for Quarterly USPTO Patent Application Filings

The software package Forecast Pro, when run in expert mode, did not select any of its "canned" naive extrapolative models. Rather, it recommended a Box-Jenkins type (or ARIMA) type model. We decided to construct several naive models, because of the simplicity, ease of use, and forecasting potential. The first and most obvious was a simple random walk with trend starting in 1984. This was ruled out, because it would imply extremely high growth rates of 8.8 percent per annum in filings or a doubling of filing applications every of 7.9 years.

Another model which was ultimately chosen and performs quite well is a transfer function type model. It combines features of an autoregressive model and deterministic components. The functional form of the model is shown below:

$$F_t = \alpha + \beta_1 \cdot F_{t-1} + \beta_2 \cdot F_{t-2} + \beta_3 \cdot F_{t-4} + \beta_4 \cdot SURGE + \varepsilon_t \qquad (4.1)$$

Filings from three quarters ago did not contribute to the model fit so they were omitted from the specification. The SURGE variable is zero except 1982Q4 where it equals one and 1983Q1 where it equals minus one. The results from the model are presented in Table 7. All variables are significant except the constant which is kept in for testing purposes. The fit is extremely high, adjusted R-square is 0.95. There is no evidence of first order autocorrelation of the residuals from the Durbin's H statistic of -0.73. The residuals appear to be normal from the Jarque-Bera test, and the variance is constant from tests for heteroscedasticity.

2. Box-Jenkins (ARIMA) Model for Quarterly USPTO Patent Application Filings

The graph and autocorrelogram of the patent application filings showed that the series is nonstationary. As with the annual data, the quarterly data are transformed to first differences of natural logarithms. If only the post 1983 data were used, the ARIMA model would behave poorly. Furthermore, the annual growth rate of 8.8 percent for that period is not expected to continue as a long term trend. The ARIMA model for quarterly patent application filings will need to be used with caution, because of the drawbacks caused by the clear structural changes.

Table 8 contains the autocorrelation functions of the series in natural logarithms of levels and first differences. The nonstationary properties of the series are clearly presented by persistence (or slow decay) of the autocorrelations in levels. There is no clear pattern to either the autocorrelations or partial autocorrelations of the series in first differences. This suggests that a mixed model including both autoregressive (AR) and moving average (MA) components would be the best candidate model.

Since the maximum lag length for a significant partial autocorrelation is three, that was the longest lag chosen for either the AR or MA components. Various combinations of models were tried ARIMA(1,1,1), ARIMA(2,1,1), ARIMA(3,1,1), ARIMA(3,1,2), ARIMA(1,1,2), and ARIMA(1,1,3) were tried along with pure integrated AR and MA models. The model which fit best is an ARIMA(1,1,1); the results are given in Table 9. The negative sign on the AR coefficient is caused by the saw-toothed pattern in the data noted earlier. The residuals appear to be uncorrelated individually or as a group from the Ljung-Box-Pierce statistics. However, there is strong evidence of excess kurtosis.

3. Econometric Model for Quarterly USPTO Patent Application Filings

The quarterly model of patent application filings is based on the same economic logic as the annual model. Filings are an output of the inventive process.

Lagged filings from the same quarter in previous years are used to capture the dynamics and feedback from past inventive activity. Research and development expenditures are not available on a quarterly basis, so they were proxied for by using investment and gross domestic product. Because filings are the result of long run investment and economic trends, the explanatory variables are used or transformed to reflect this property. The real Gross Domestic Product, real Gross Private Domestic Investment and real Gross Private Non-Residential Investment series were converted to four quarter moving averages, referred to as AGDPQ, AGPDI, and AGIN respectively. This smoothes the series to reflect the trend in economic activity over the past year. The economic activity variables are on a calendar year basis. Filing fees were adjusted by the 1987 GDP deflator. An expected real fee series, ERFEE, was created to control for the announcement effect on fees. The series uses next period's fee as the fee in the current period. The SURGE variable is used to capture the impact of a change in the fee structure in 1983Q1 as in the naive model.

The technique for estimating the quarterly econometric model follows the general to specific approach used in the annual model. An ADL model with a fixed lag length is estimated. Fewer lags are used in an attempt to find a parsimonious model. The initial model is specified as follows:

$$F_t = \alpha + \sum_{i=1}^{4} \phi_i \cdot F_{t-4\cdot i} + \sum_{i=1}^{4} \beta_i \cdot ARDPQ_{t-4\cdot i} + \sum_{i=0}^{4} \gamma_i \cdot ERFEE_{t-i} + SURGE_t + \varepsilon_t \qquad (4.2)$$

The two investment series were tried as replacements for AGDP, but did not appear to improve the predictive power. The AIC and BSC were used in selecting the appropriate lag length of the models. Tests for lag length were also performed using the two investment series, but they did not yield significantly better results. Also, lags two through four in the fee series provided no additional explanatory power, so the tests were conditioned upon the current quarter and one lagged quarter of the series. The results are summarized in Table 10. A model with filings lagged four and eight quarters, the four quarter moving average of real GDP lagged four and eight quarters, and the current and current expected real fee series and lagged one quarter appear to provide a parsimonious fit.

$$F_t = \alpha + \phi_1 \cdot F_{t-4} + \phi_2 \cdot F_{t-8} + \beta_1 \cdot AGDPQ_{t-4} + \beta_2 \cdot AGDPQ_{t-8} + \gamma_1 \cdot ERFEE_t + \gamma_2 \cdot ERFEE_{t-1} + SURGE_t + \varepsilon_t \qquad (4.3)$$

Results from the model are provided in Table 11. The model fits the data quite well; adjusted R squared is 0.961, and the standard error of the estimate is 1259 with a mean of the dependent variable of 31329. The standard error is about four percent of the mean value. The Durbin's H statistic is -1.33 indicating first order autocorrelation is not a serious problem. There is no further autocorrelation from the residual correlogram. The null hypothesis of no heteroscedasticity cannot be rejected. The Jarque-Bera test for normality of residuals indicates there is no skewness and excess kurtosis.

The dynamics of the model are interesting. The short run effect of a billion dollar increase in the four quarter moving average of GDP lagged one year is associated with an increase in 10 patent application filings. A lag of one year is actually three quarters, because of the fiscal year calendar year distinction. On the other hand, the long run impact is about 17 filings. The dynamics are calculated as follows:

$$\text{Effect of \$1 Billion increase in GDP} = \frac{\hat{\beta}_1 + \hat{\beta}_2}{1 - \hat{\phi}_1 - \hat{\phi}_2} = \frac{8.8869}{1 - 0.47237} = 16.84$$

This result is interesting, because it implies for every billion dollars of real GDP there are 17 patent application filings. If real GDP is $5 trillion, that means there will be approximately 85,000 filings which is about equal to the number of domestic utility filings in 1991. The latter are about 49 percent of total UPR filings.

The net effect of the expected fees is zero, but they do appear to marginally contribute to the model. As previously mentioned, patent application filing fees are set by statute. Historically, patent application filing fees have not been set as an impediment to the innovation and invention process. From an economic perspective, patent application filing fees can be viewed as user costs. The negative-positive sign pattern captures both the surge effect of announced fee increases and the declining real fee(s) between fee increases.

V. Evaluation of the Forecasting Models

Tables 12 and 13 contain the results from the within-sample unconditional forecast evaluation of the annual and quarterly models respectively. The two-year ahead predictions in the annual model are dynamic in that they are based upon the one-year ahead forecasted value of patent application filings. The unconditional forecasts use the actual values for fees and R&D expenditures in the predicted period(s). Filing fees were known in advance, except for the inflation impact on their real values. The R&D measure enters into the two-year ahead prediction, and there it is through the last, most recent year in a three year moving average term.

The quarterly models forecast out eight quarters; the first four are summed for the one-year ahead prediction and the fifth through eight quarters are summed for the two-year ahead prediction. It is dynamic as well, because predicted the values of (lagged) patent application filings are used as the forecast horizon lengthens.

A rolling set of one to two-year ahead forecasts are constructed using naive, ARIMA, and econometric models estimated through 1984. Then, the models are re-estimated using data through 1985 to produce one and two-year ahead forecasts. The procedure continues until 1990 when only a one-year ahead forecast is made. There are ultimately seven one-year ahead predictions and six two-year ahead forecasts.

In Tables 12 and 13, the actual filing level is given in the second column. Then the one and two-year ahead prediction errors from the naive, ARIMA, and econometric forecasts respectively are given in the next six columns. For example, when the model is estimated using data through 1984, the one-year ahead prediction error made in 1985 is 6,524 and the two-year ahead error is 10,629.

The annual econometric model outperforms the annual naive and ARIMA models. Its mean percent error is nearly half of the other two at 3.15 percent for the one-year ahead and 5.85 percent at two-year ahead. One note of caution, the forecasts from the annual econometric model appear to deteriorate starting in 1989. The RMSE from the econometric model is also the smallest among the three annual models. This may just be the result of using the exogenous variables, a criticism of forecast comparisons noted by Fair (1984). The naive model underpredicts in every case but the last one-year ahead forecast. The ARIMA model forecasts are less than actual levels in every case also. In the first half of the evaluation period, the econometric model overpredicts, thereafter it overpredicts.

In the within-sample forecast comparison, the quarterly forecasts using the econometric model generated smaller mean percent error and RMSE than the other two quarterly models for the one-year and two-year ahead forecast. The mean percent errors 2.18 at the one-year and 3.59 at the two-year ahead are smaller than the annual econometric model error. As with the annual model, the quarterly model overpredicts in the mid 1980s and then underpredicts in the late 1980s and early 1990s. The naive and ARIMA models also underpredict until 1990.

VI. Conclusions

A comparison of the one and two-year ahead *ex post* forecast suggests that the quarterly models consistently outperform the annual models in predicting the future levels of patent application filings. When the mean errors, the root mean squared errors, and the mean absolute errors from the *ex post* forecasts of the quarterly models are compared to the annual models, it is clear that the quarterly naive, ARIMA, and econometric models generate forecast errors much smaller than their respective annual models.

The error statistics also suggest that the quarterly econometric model outperforms the other two quarterly models. Judging from the *ex post* forecast errors, the quarterly ARIMA model appears to slightly outperform the quarterly naive model.

Among the annual forecasting models, it is the econometric model that outperforms the other two annual models in both one and two-year ahead *ex post* forecast comparison. However, unlike the case of quarterly models, the annual naive model appears to be much more reliable than the annual ARIMA model when the errors from *ex post* forecasts are considered. While the mean errors for the one-year ahead forecast from the annual econometric and naive models were 3,288 and 3,531 respectively, it was 5,565 for the ARIMA model. The annual ARIMA model performs particularly poorly relative to the other two models in the two-year ahead forecasts as indicated by its large mean error, RMSE, and mean absolute error.

In short, the quarterly models perform better than the annual models. This certainly is the case when their respective Mean Percent Errors, Root Mean Squared Errors, and Mean Absolute Percent Errors from the within-sample forecast tests are compared to those of the annual models. Based on the error statistics computed from the one and two-year ahead within-sample forecasting, the top three models in ranking appear to be the quarterly econometric, the quarterly naive and the annual econometric models. The USPTO will continue to examine the performance of the different models on a periodic basis with the most recent observations available.

FIGURE 1

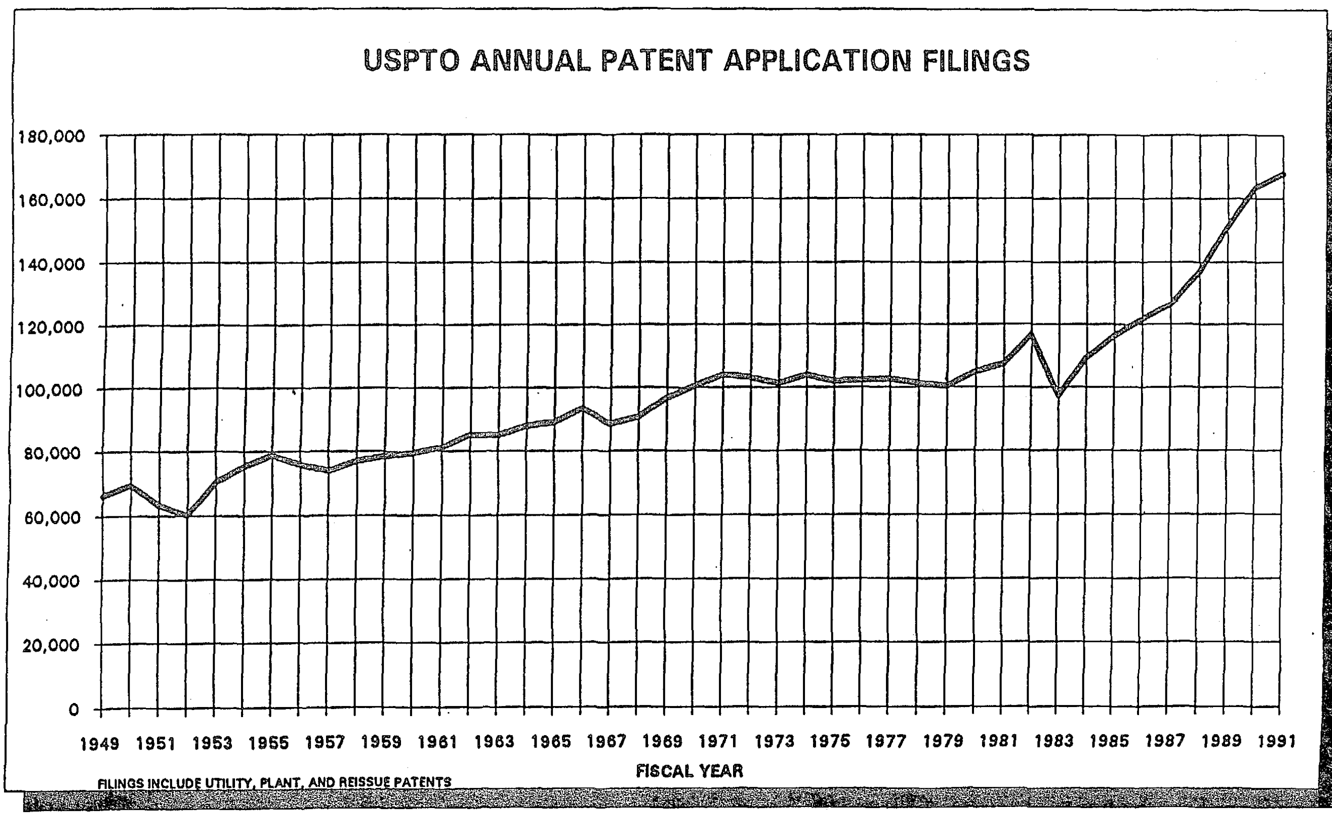

USPTO ANNUAL PATENT APPLICATION FILINGS
180,000
160,000
140,000
120,000
100,000
80,000
60,000
40,000
20,000
0
1949 1951 1953 1955 1957 1959 1961 1963 1965 1967 1969 1971 1973 1975 1977 1979 1981 1983 1985 1987 1989 1991
FISCAL YEAR
FILINGS INCLUDE UTILITY, PLANT, AND REISSUE PATENTS

FIGURE 2

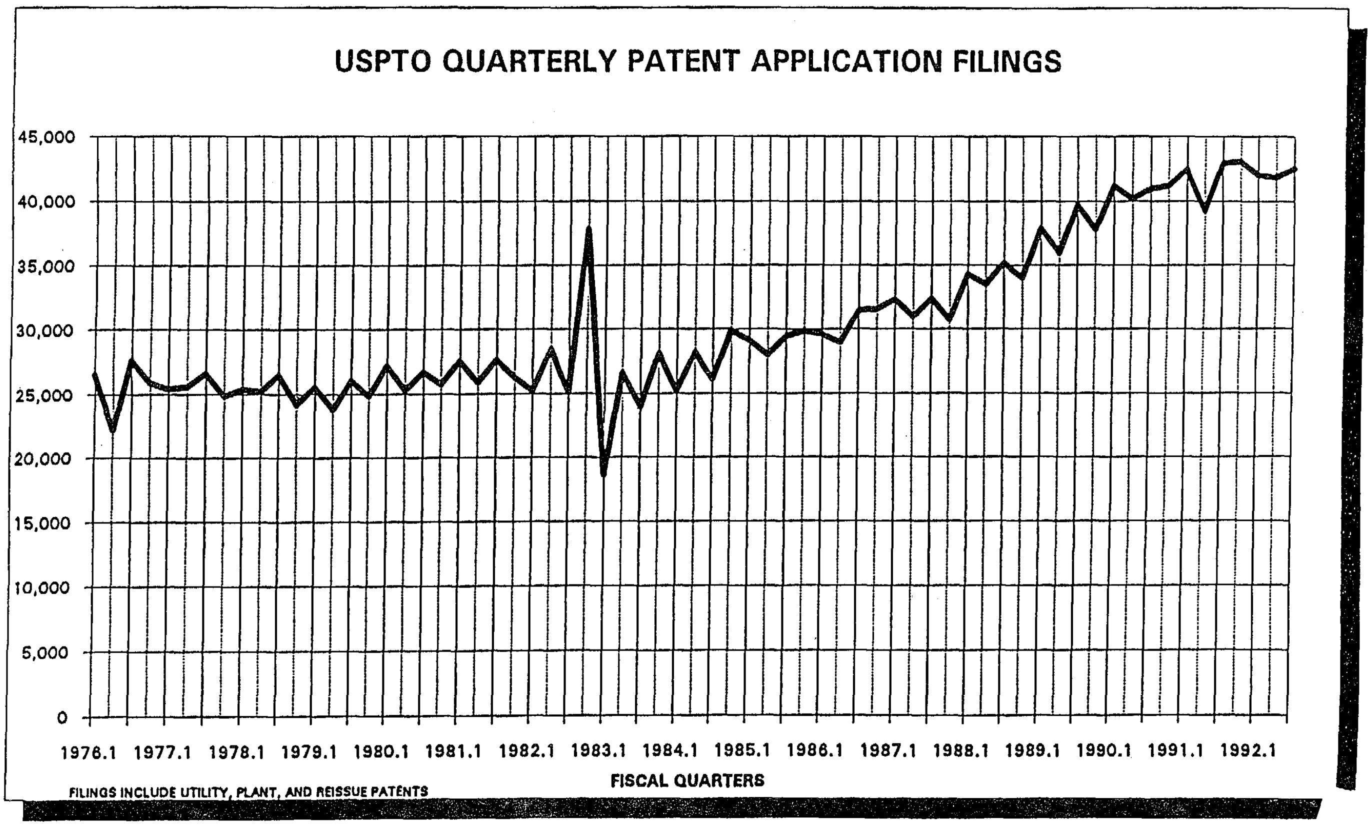

USPTO QUARTERLY PATENT APPLICATION FILINGS
45,000
40,000
35,000
30,000
25,000
20,000
15,000
10,000
5,000
0
1976.1
1977.1
1978.1
1979.1
1980.1
1981.1
1982.1
1983.1
1984.1
1985.1
1986.1
1987.1
1988.1
1989.1
1990.1
1991.1
1992.1
FISCAL QUARTERS
FILINGS INCLUDE UTILITY, PLANT, AND REISSUE PATENTS

FIGURE 3

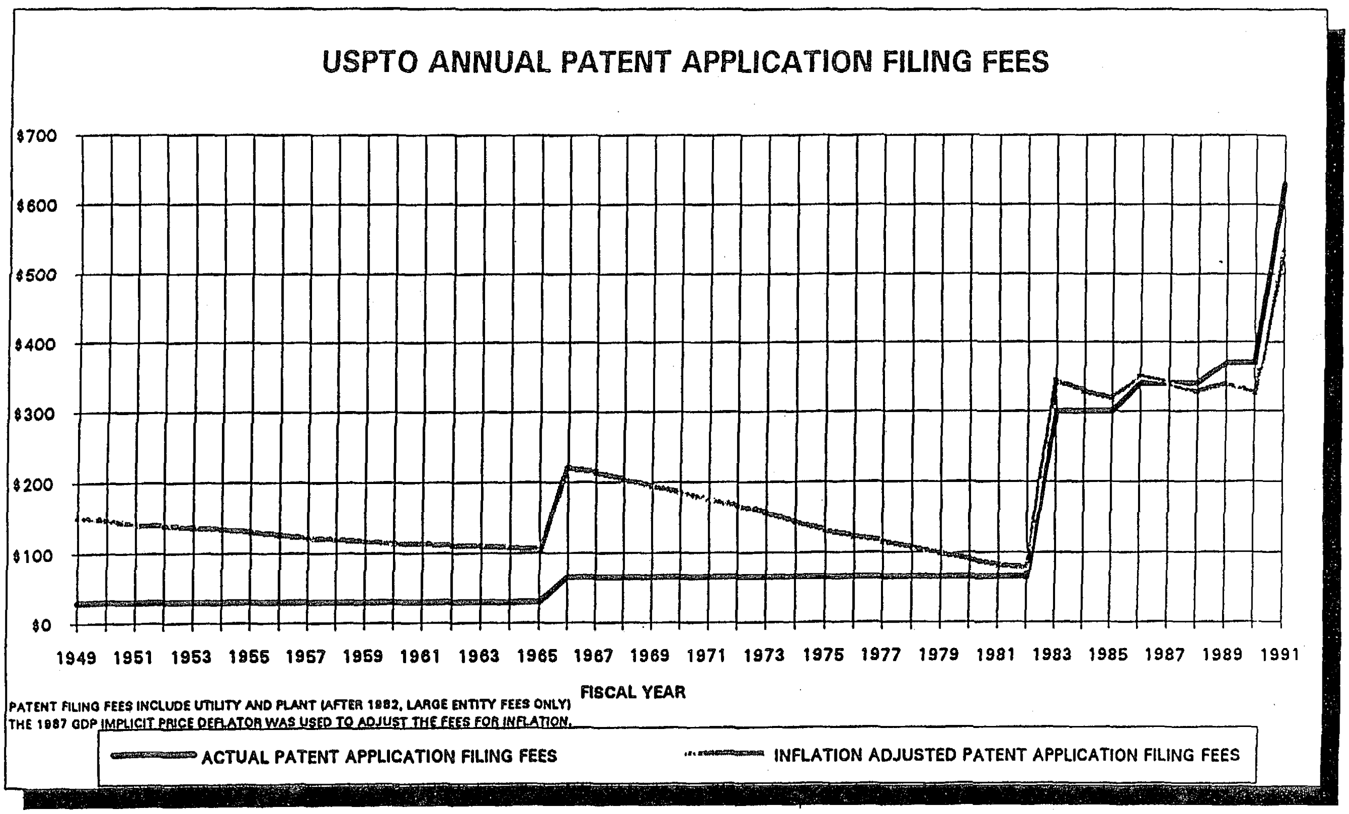

USPTO ANNUAL PATENT APPLICATION FILING FEES
$700
$600
$500
$400
$300
$200
$100
$0
1949 1951 1953 1955 1957 1959 1961 1963 1965 1967 1969 1971 1973 1975 1977 1979 1981 1983 1985 1987 1989 1991
FISCAL YEAR
PATENT FILING FEES INCLUDE UTILITY AND PLANT (AFTER 1982, LARGE ENTITY FEES ONLY)
THE 1987 GDP IMPLICIT PRICE DEFLATOR WAS USED TO ADJUST THE FEES FOR INFLATION.
ACTUAL PATENT APPLICATION FILING FEES
INFLATION ADJUSTED PATENT APPLICATION FILING FEES

FIGURE 4

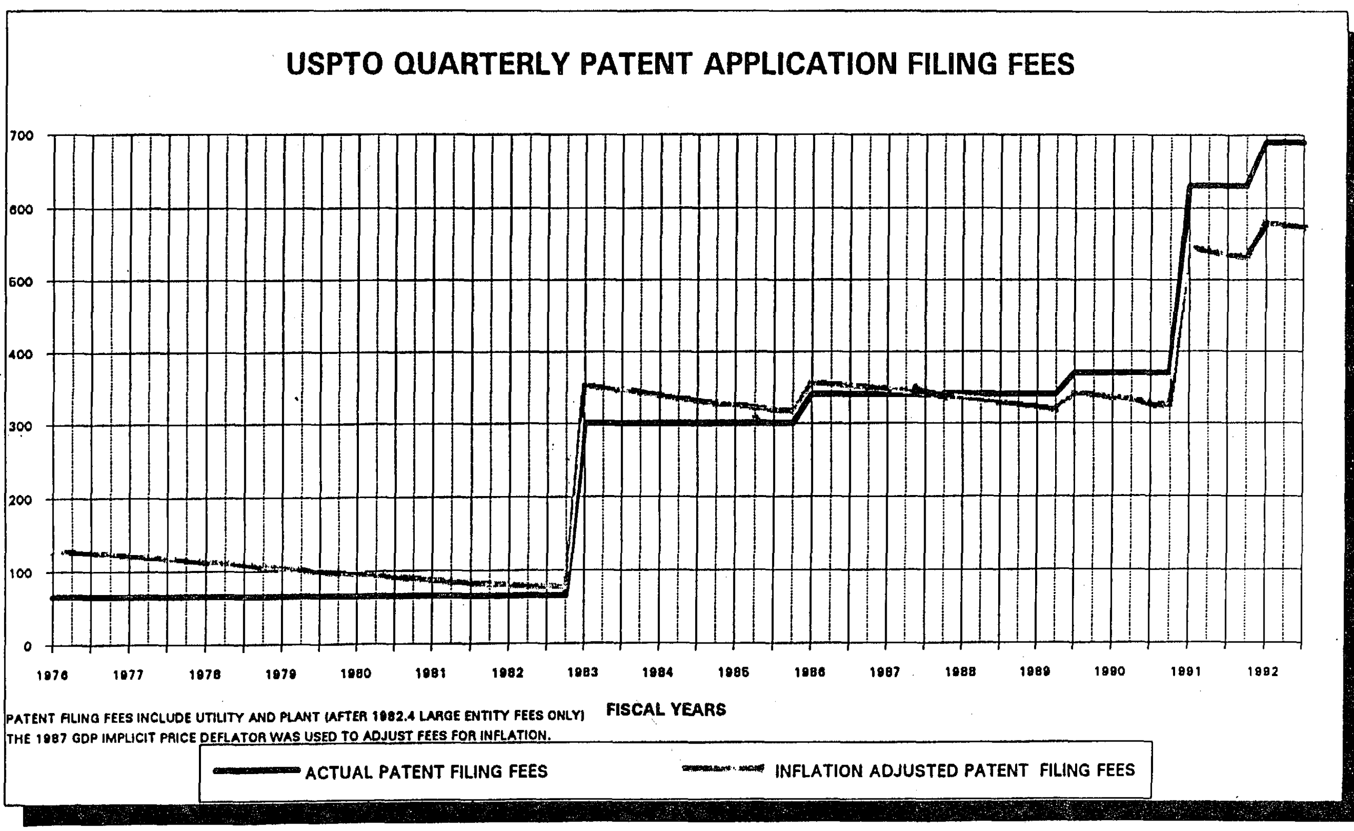

USPTO QUARTERLY PATENT APPLICATION FILING FEES
0
100
200
300
400
500
600
700
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
FISCAL YEARS
PATENT FILING FEES INCLUDE UTILITY AND PLANT (AFTER 1982.4 LARGE ENTITY FEES ONLY)
THE 1987 GDP IMPLICIT PRICE DEFLATOR WAS USED TO ADJUST FEES FOR INFLATION.
ACTUAL PATENT FILING FEES
INFLATION ADJUSTED PATENT FILING FEES

FIGURE 5

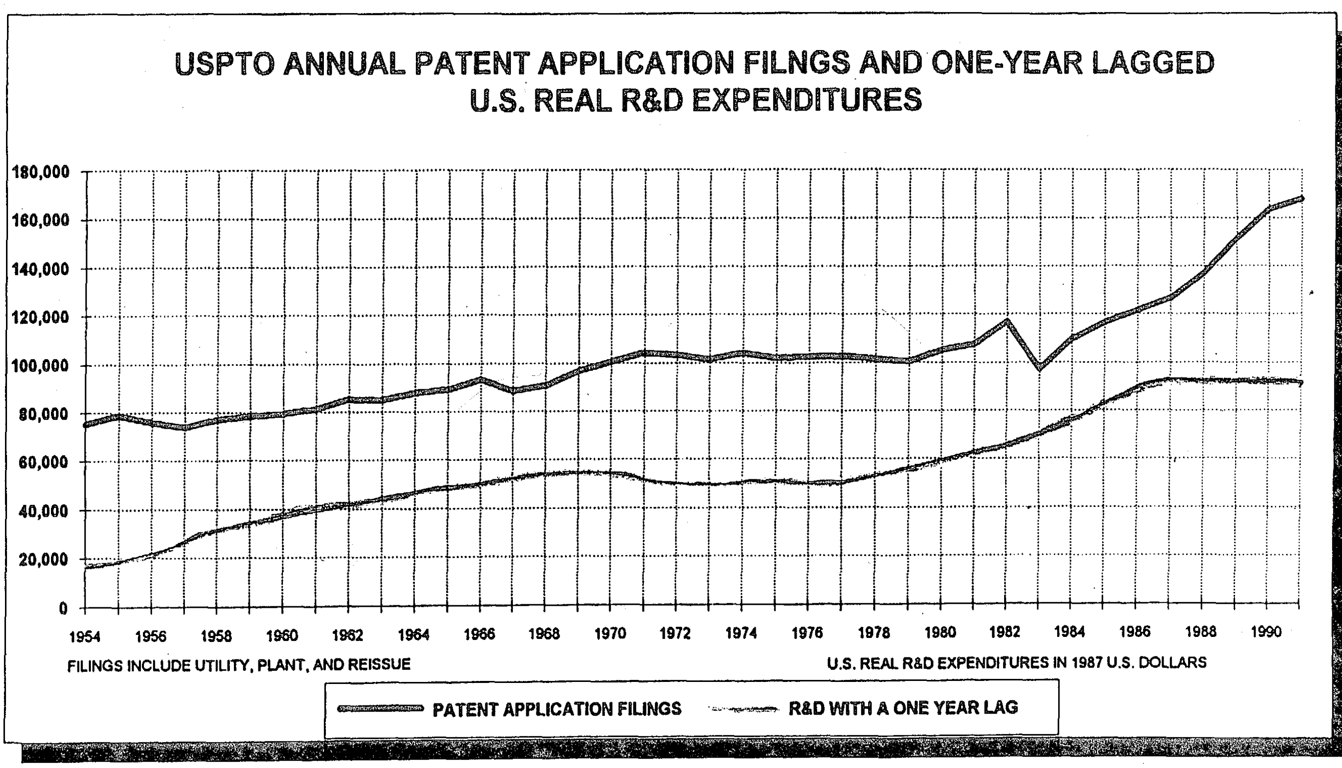

USPTO ANNUAL PATENT APPLICATION FILNGS AND ONE-YEAR LAGGED U.S. REAL R&D EXPENDITURES
180,000
160,000
140,000
120,000
100,000
80,000
60,000
40,000
20,000
0
1954
1956
1958
1960
1962
1964
1966
1968
1970
1972
1974
1976
1978
1980
1982
1984
1986
1988
1990
FILINGS INCLUDE UTILITY, PLANT, AND REISSUE
U.S. REAL R&D EXPENDITURES IN 1987 U.S. DOLLARS
PATENT APPLICATION FILINGS
R&D WITH A ONE YEAR LAG

FIGURE 6

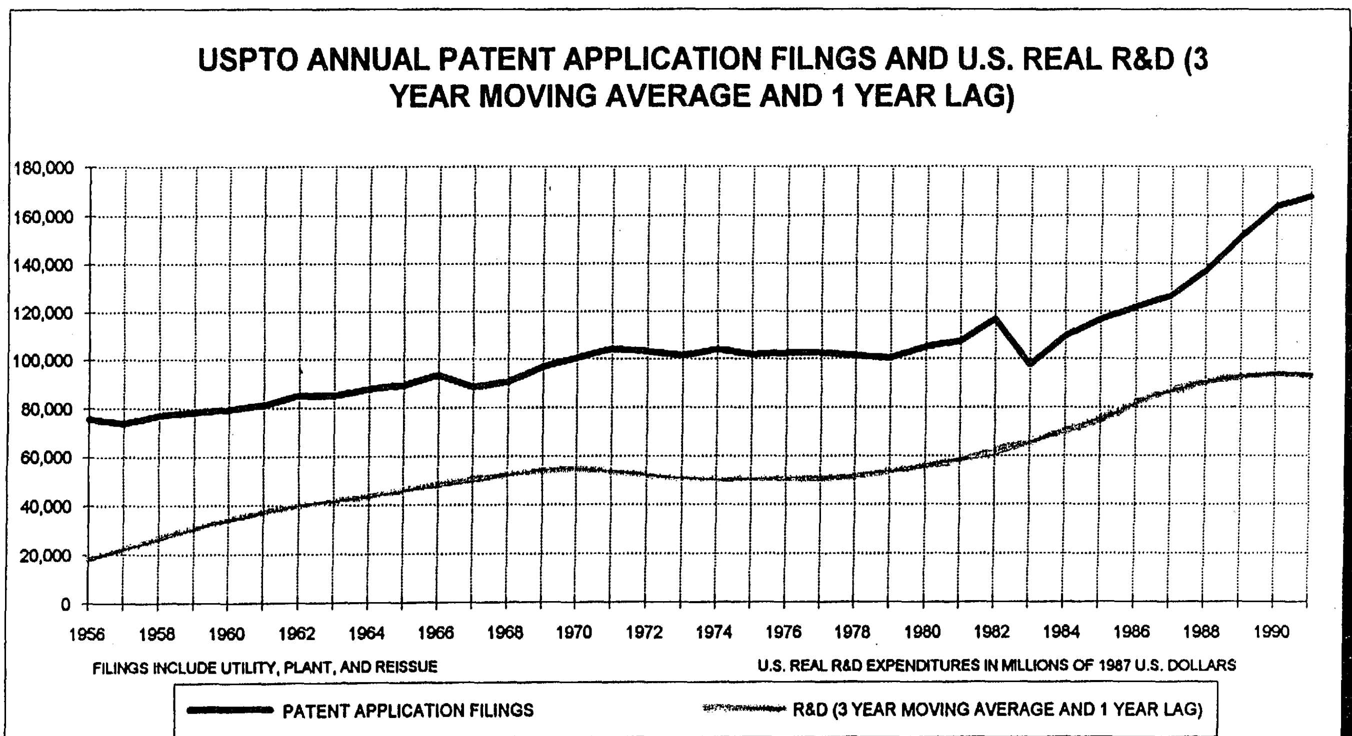

USPTO ANNUAL PATENT APPLICATION FILNGS AND U.S. REAL R&D (3 YEAR MOVING AVERAGE AND 1 YEAR LAG)
180,000
160,000
140,000
120,000
100,000
80,000
60,000
40,000
20,000
0
1956
1958
1960
1962
1964
1966
1968
1970
1972
1974
1976
1978
1980
1982
1984
1986
1988
1990
FILINGS INCLUDE UTILITY, PLANT, AND REISSUE
U.S. REAL R&D EXPENDITURES IN MILLIONS OF 1987 U.S. DOLLARS
PATENT APPLICATION FILINGS
R&D (3 YEAR MOVING AVERAGE AND 1 YEAR LAG)

FIGURE 7

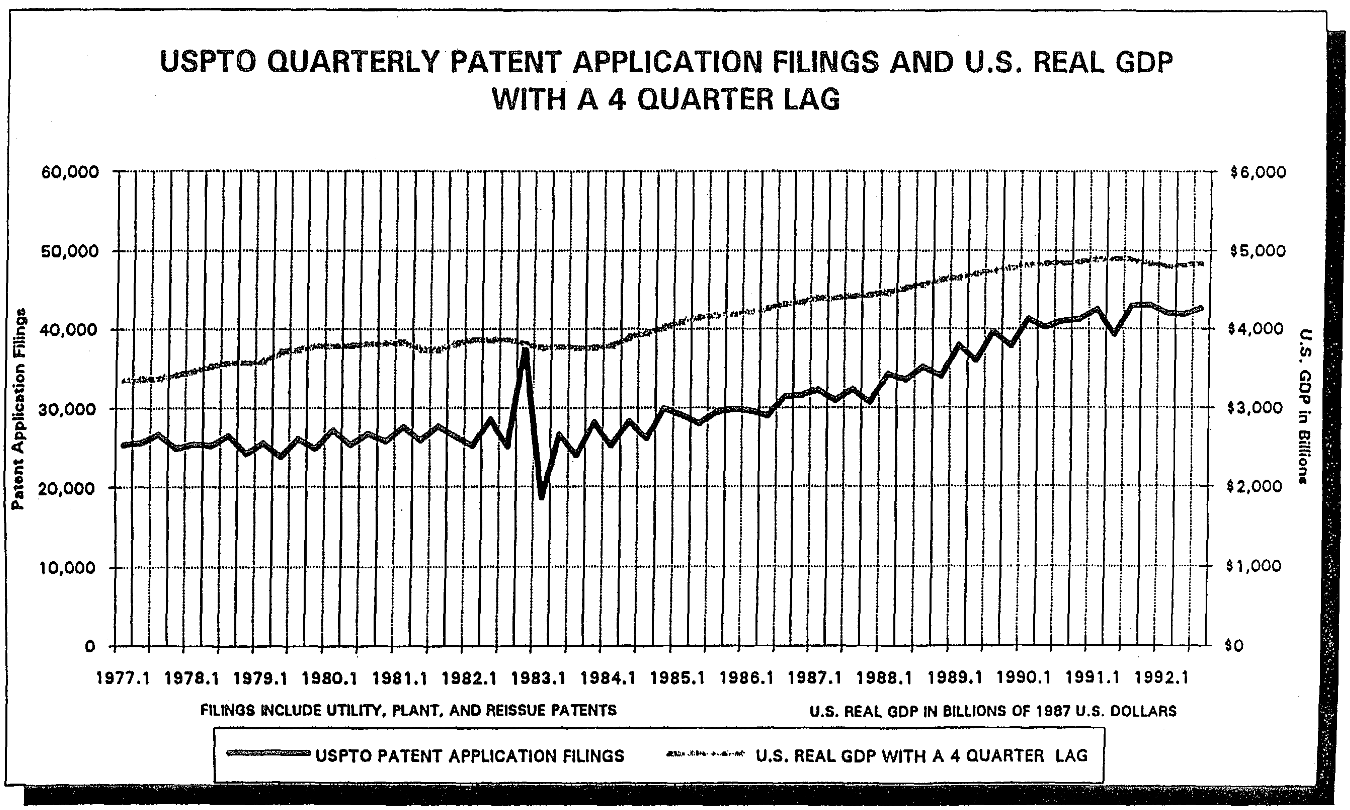

USPTO QUARTERLY PATENT APPLICATION FILINGS AND U.S. REAL GDP WITH A 4 QUARTER LAG
Patent Application Filings
60,000
50,000
40,000
30,000
20,000
10,000
0
U.S. GDP in Billions
$6,000
$5,000
$4,000
$3,000
$2,000
$1,000
$0
1977.1 1978.1 1979.1 1980.1 1981.1 1982.1 1983.1 1984.1 1985.1 1986.1 1987.1 1988.1 1989.1 1990.1 1991.1 1992.1
FILINGS INCLUDE UTILITY, PLANT, AND REISSUE PATENTS
U.S. REAL GDP IN BILLIONS OF 1987 U.S. DOLLARS
USPTO PATENT APPLICATION FILINGS
U.S. REAL GDP WITH A 4 QUARTER LAG

FIGURE 8

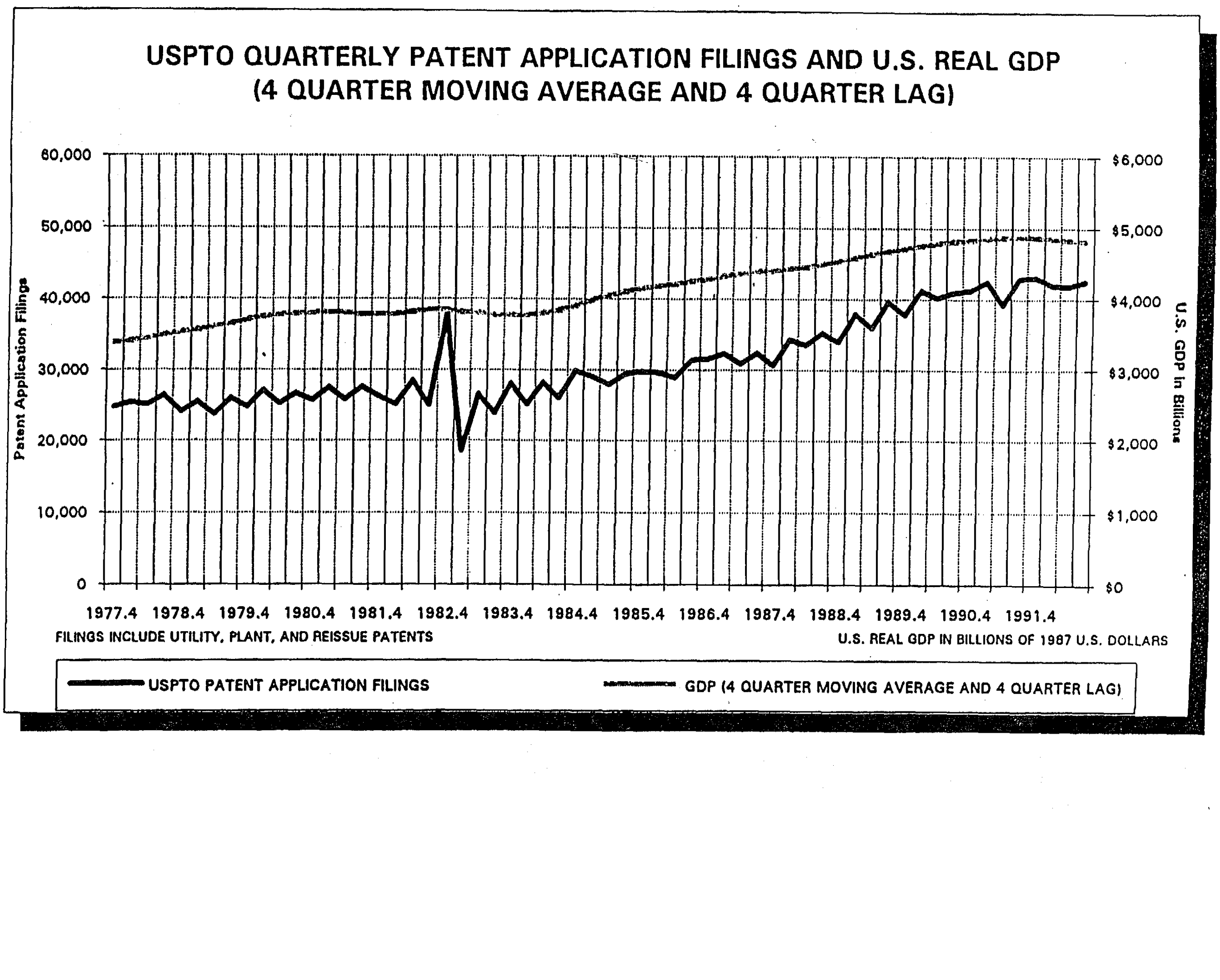

USPTO QUARTERLY PATENT APPLICATION FILINGS AND U.S. REAL GDP
(4 QUARTER MOVING AVERAGE AND 4 QUARTER LAG)
60,000
50,000
40,000
30,000
20,000
10,000
0
Patent Application Filings
$6,000
$5,000
$4,000
$3,000
$2,000
$1,000
$0
U.S. GDP in Billions
1977.4 1978.4 1979.4 1980.4 1981.4 1982.4 1983.4 1984.4 1985.4 1986.4 1987.4 1988.4 1989.4 1990.4 1991.4
FILINGS INCLUDE UTILITY, PLANT, AND REISSUE PATENTS
U.S. REAL GDP IN BILLIONS OF 1987 U.S. DOLLARS
USPTO PATENT APPLICATION FILINGS
GDP (4 QUARTER MOVING AVERAGE AND 4 QUARTER LAG)

TABLE 1

U.S. ANNUAL AND QUARTERLY ECONOMIC AND DEMOGRAPHIC DATA

ANNUAL ECONOMIC AND DEMOGRAPHIC INDICATORS	CALENDAR YEARS	BASE YEAR/CURRENCY	SOURCE
Real Gross Domestic Product in Billions	1959 - 1992	1987 U.S. Dollars	Citibase Database Services
Research and Development Expenditures in Billions	1953 - 1991	1987 U.S. Dollars	National Science Foundation
Scientists and Engineers Engaged in R&D in Thousands	1965 - 1988	-	National Science Foundation
Graduates in Science, Mathematics, and Engineering	1969 - 1985	-	U.S. Department of Education
Real Gross Private Domestic Investment in Billions	1959 - 1992	1987 U.S. Dollars	Citibase Database Services
10 Year Government Bond Interest Rates	1949 - 1992	-	Citibase Database Services
Business Formations	1963 - 1989	-	Dun and Bradstreet Corporation
1987 GDP Implicit Price Deflator	1959 - 1992	1987 U.S. Dollars	Citibase Database Services

QUARTERLY ECONOMIC AND DEMOGRAPHIC INDICATORS	QUARTERS	BASE YEAR/CURRENCY	SOURCE
Real Gross Domestic Product in Billions	1976.1-1992.3	1987 U.S. Dollars	Citibase Database Services
Research and Development Expenditures in Billions	NA	-	NA
Scientists and Engineers Engaged in R&D in Thousands	NA	-	NA
Graduates in Science, Mathematics, and Engineering	NA	-	NA
Real Gross Private Domestic Investment in Billions	1976.1 - 1992.3	1987 U.S. Dollars	Citibase Database Services
10 Year Government Bond Interest Rates	1976.1 - 1992.3	-	Citibase Database Services
Business Formations	NA	-	NA
GDP Implicit Price Deflator	1976.1 - 1992.3	1987 U.S. Dollars	Citibase Database Services

TABLE 2

NAIVE MODEL FOR ANNUAL PATENT APPLICATION FILINGS
HOLT EXPONENTIAL SMOOTHING: LINEAR TREND,
NO SEASONALITY
FORECAST PRO FOR WINDOWS VERSION 1.00B
SAMPLE PERIOD 1949 - 1992

Component	Smoothing Weight	Final Value
Level	0.837	1.7359e+005
Trend	0.193	6858.2

Standard Diagnostics

Sample size 44	Number of parameters 2
Mean 9.957e+004	Standard deviation 2.682e+004
R-square 0.9566	Adjusted R-square 0.9555
Durbin-Watson 1.974	Ljung-Box(18)=12.69 P=0.1902
Forecast error 5657	BIC 6023 (Best so far)
MAPE 0.0413	RMSE 5527
MAD 4007	

TABLE 3

ANNUAL PATENT APPLICATION FILINGS AUTOCORRELOGRAM
SAMPLE PERIOD 1956-1991

FILINGS IN NATURAL LOGARITHMS

MEAN = 11.526
VARIANCE = 0.41263E-01
STANDARD DEV. = 0.20313

LAGS	AUTOCORRELATIONS												STD ERR
1 -12	0.85	0.69	0.56	0.46	0.36	0.28	0.22	0.17	0.17	0.10	0.07	0.02	0.17
13 -24	0.00	-.01	-.02	-.02	-.04	-.06	-.08	-.11	-.16	-.20	-.24	-.26	0.38

LAGS	PARTIAL AUTOCORRELATIONS												STD ERR
1 -12	0.85	-.08	-.02	0.02	-.03	-.03	0.02	0.03	0.09	-.21	0.08	-.07	0.17

FILINGS IN FIRST DIFFERENCES OF NATURAL LOGARITHMS

MEAN = 0.22716E-01
VARIANCE = 0.26228E-02
STANDARD DEV. = 0.51214E-01

LAGS	AUTOCORRELATIONS												STD ERR
1 -12	-.19	0.09	-.02	0.18	0.02	-.06	-.11	0.13	-.14	0.05	-.01	-.11	0.17
13 -24	-.11	-.14	-.04	0.19	-.19	0.05	0.01	0.12	-.09	-.02	-.02	0.03	0.19

LAGS	PARTIAL AUTOCORRELATIONS												STD ERR
1 -12	-.19	0.05	0.01	0.19	0.09	-.07	-.16	0.05	-.11	0.03	0.09	-.14	0.17

MODIFIED BOX-PIERCE (LJUNG-BOX-PIERCE) STATISTICS (CHI-SQUARE)

LAG	Q	DF	P-VALUE	LAG	Q	DF	P-VALUE
1	1.36	1	.243	13	7.11	13	.896
2	1.65	2	.437	14	8.40	14	.867
3	1.67	3	.644	15	8.52	15	.901
4	3.09	4	.543	16	10.92	16	.814
5	3.11	5	.684	17	13.57	17	.697
6	3.25	6	.777	18	13.78	18	.743
7	3.86	7	.795	19	13.78	19	.796
8	4.62	8	.798	20	15.11	20	.770
9	5.55	9	.784	21	15.88	21	.776
10	5.69	10	.840	22	15.91	22	.821
11	5.70	11	.893	23	15.94	23	.858
12	6.43	12	.893	24	16.02	24	.887

TABLE 4

ANNUAL PATENT APPLICATION FILINGS ARIMA (1,1,0) MODEL

R-SQUARE = 0.0368 R-SQUARE ADJUSTED = 0.0077
VARIANCE OF THE ESTIMATE-SIGMA**2 = 0.26001E-02
STANDARD ERROR OF THE ESTIMATE-SIGMA = 0.50991E-01
AKAIKE INFORMATION CRITERIA -AIC(K) = -5.8379
SCHWARZ CRITERIA- SC(K) = -5.7490

	PARAMETER ESTIMATES	STD ERROR	T-STAT
AR(1)	-0.19454	0.1684	-1.155
CONSTANT	0.27412E-01	0.9482E-02	2.891

RESIDUALS

LAGS	AUTOCORRELATIONS												STD ERR
1 -12	0.02	0.05	0.04	0.20	0.05	-.08	-.11	0.08	-.11	0.03	-.02	-.15	0.17

MODIFIED BOX-PIERCE (LJUNG-BOX-PIERCE) STATISTICS (CHI-SQUARE)

LAG	Q	DF	P-VALUE
2	0.11	1	.741
3	0.16	2	.923
4	1.82	3	.611
5	1.91	4	.751
6	2.21	5	.819
7	2.77	6	.837
8	3.11	7	.874
9	3.71	8	.882
10	3.75	9	.927
11	3.77	10	.957
12	4.97	11	.933

ANALYSIS OF RESIDUALS
VALUES RANGE FROM -0.1920 TO 0.0873
SAMPLE MOMENTS OF RESIDUALS (USING THE DIVISOR 35) :
MEAN = -0.2351045E-03
VARIANCE = 0.2451460E-02
SKEWNESS = -1.371450
KURTOSIS = 7.257485
STUDENTIZED RANGE = 5.640804

RSS(ARIMA(1,1,0)) = 0.08589 and RSS(ARIMA(0,1,0)) = 0.09879

$$F\text{-}test = \left[\frac{(0.09878-0.08589)}{0.08589}\right]*\frac{35}{1} = 5.25$$

critical values:
$F(0.05,1,30)=4.17$
$F(0.01,1,30)=7.56$

TABLE 5

SPECIFICATION TESTS FOR LAG LENGTH IN FINAL AUTOREGRESSIVE DISTRIBUTED LAG MODEL

Sample 1959-1991

FIL	RD	FEE	AIC	BSC	$\bar{R}^2$
UNRESTRICTED MODELS STARTING WITH FOUR LAGS					
1-4	1-4	0-4	-6.85	-6.19	0.9750
1-3	1-3	0-3	-6.90	-6.38	0.9759
1-2	1-2	0-2	-6.91	-6.54	0.9758
1-1	1-1	0-1	-7.02	-6.80	0.9778
RESTRICTING LAGS ON FILINGS TO ONE					
1-1	1-4	0-4	-6.93	-6.42	0.9764
1-1	1-3	0-3	-6.92	-6.49	0.9754
1-1	1-2	0-2	-6.64	-6.97	0.9767
RESTRICTING LAGS ON FILINGS AND FEES TO ONE					
1-1	1-4	0-1	-7.01	-6.63	0.9768
1-1	1-3	0-1	-6.96	-6.63	0.9756
1-1	1-2	0-1	-7.00	-6.72	0.9769

TABLE 6.
ANNUAL PATENT APPLICATION FILINGS ECONOMETRIC MODEL
SAMPLE 1959 - 1991

```
R-SQUARE =   0.9807      R-SQUARE ADJUSTED =   0.9778
VARIANCE OF THE ESTIMATE-SIGMA**2 =  0.77181E-03
STANDARD ERROR OF THE ESTIMATE-SIGMA =  0.27781E-01
SUM OF SQUARED ERRORS-SSE=  0.20839E-01
MEAN OF DEPENDENT VARIABLE =   11.561
LOG OF THE LIKELIHOOD FUNCTION(IF DEPVAR LOG) = -297.984
 AKAIKE (1973) INFORMATION CRITERION- LOG AIC =  -7.0242
 SCHWARZ(1978) CRITERION-LOG SC =  -6.7951
```

VARIABLE	SUM OF LAG COEFS	STD ERROR	T-RATIO	MEAN LAG
LFIL	0.81199	0.72516E-01	11.197	1.0000
ARD	0.17632	0.55689E-01	3.1662	1.0000
LP	0.10082E-01	0.18445E-01	0.54661	13.160

VARIABLE NAME	ESTIMATED COEFFICIENT	STANDARD ERROR	T-RATIO 27 DF	P-VALUE	PARTIAL CORR.	STANDARDIZED COEFFICIENT	ELASTICITY AT MEANS
LFIL	0.81199	0.7252E-01	11.20	1.000	0.907	0.7524	0.8120
ARD	0.17632	0.5569E-01	3.166	0.998	0.520	0.2606	0.1763
LP	-0.12260	0.1911E-01	-6.415	0.000	-0.777	-0.3384	-0.1226
LP	0.13268	0.1953E-01	6.793	1.000	0.794	0.3394	0.1327
CONSTANT	0.21665	0.4433	0.4887	0.686	0.094	0.0000	0.2166

```
DURBIN-WATSON = 1.6148     VON NEUMANN RATIO = 1.6669     RHO =  0.13653
RESIDUAL SUM =  0.12471E-12  RESIDUAL VARIANCE =  0.77181E-03
RUNS TEST:   19 RUNS,    15 POSITIVE,    17 NEGATIVE, NORMAL STATISTIC =  0.7443
DURBIN H STATISTIC (ASYMPTOTIC NORMAL) =  0.84684
COEFFICIENT OF SKEWNESS =   0.1109 WITH STANDARD DEVIATION OF 0.4145
COEFFICIENT OF EXCESS KURTOSIS =  -1.0189 WITH STANDARD DEVIATION OF 0.8094

     GOODNESS OF FIT TEST FOR NORMALITY OF RESIDUALS - 10 GROUPS
OBSERVED  0.0   0.0   3.0   7.0   7.0   5.0   6.0   4.0   0.0   0.0
EXPECTED  0.3   0.9   2.5   5.1   7.2   7.2   5.1   2.5   0.9   0.3
CHI-SQUARE =     4.7952 WITH  3 DEGREES OF FREEDOM

JARQUE-BERA ASYMPTOTIC LM NORMALITY TEST
CHI-SQUARE =     1.5250 WITH  2 DEGREES OF FREEDOM

HETEROSKEDASTICITY TESTS
  E**2 ON YHAT:             CHI-SQUARE =       4.268 WITH 1 D.F.
  E**2 ON YHAT**2:          CHI-SQUARE =       4.317 WITH 1 D.F.
  E**2 ON LOG(YHAT**2):     CHI-SQUARE =       4.218 WITH 1 D.F.
  E**2 ON X (B-P-G) TEST:            CHI-SQUARE =      5.685 WITH   4 D.F.
  E**2 ON LAG(E**2)  ARCH TEST:      CHI-SQUARE =      1.356 WITH    1 D.F.
  LOG(E**2) ON X (HARVEY) TEST:      CHI-SQUARE =      1.519 WITH   4 D.F.
  ABS(E) ON X (GLEJSER) TEST:        CHI-SQUARE =      2.852 WITH   4 D.F.
```

TABLE 6 (continued)

RESIDUAL CORRELOGRAM
LM-TEST FOR HJ:RHO(J)=0, STATISTIC IS STANDARD NORMAL

LAG	RHO	STD ERR	T-STAT	LM-STAT	DW-TEST	BOX-PIERCE-LJUNG
1	0.1206	0.1768	0.6823	0.8316	1.6148	0.5105
2	0.3191	0.1768	1.8050	2.1196	1.1693	4.2028
3	-0.0970	0.1768	-0.5487	0.7566	1.8593	4.5557
4	-0.0010	0.1768	-0.0054	0.0077	1.6630	4.5558
5	0.0035	0.1768	0.0196	0.0245	1.5407	4.5562
6	-0.0574	0.1768	-0.3245	0.4091	1.6414	4.6940
7	0.1388	0.1768	0.7851	1.0700	1.2162	5.5322
8	0.0629	0.1768	0.3557	0.5152	1.3182	5.7114
9	0.0571	0.1768	0.3231	0.5412	1.2601	5.8657

LM CHI-SQUARE STATISTIC WITH 9 D.F. IS 4.977

SEQUENTIAL CHOW AND GOLDFELD-QUANDT TESTS

N1	N2	SSE1	SSE2	CHOW	G-Q	DF1	DF2
6	26	0.15528E-03	0.15612E-01	1.4152	0.20886	1	21
7	25	0.18658E-03	0.14608E-01	1.7975	0.12772	2	20
8	24	0.31967E-03	0.13292E-01	2.3364	0.15232	3	19
9	23	0.48792E-03	0.13277E-01	2.2610	0.16537	4	18
10	22	0.17435E-02	0.12001E-01	2.2711	0.49396	5	17
11	21	0.21412E-02	0.11672E-01	2.2378	0.48919	6	16
12	20	0.22545E-02	0.11033E-01	2.5007	0.43789	7	15
13	19	0.25666E-02	0.10965E-01	2.3759	0.40961	8	14
14	18	0.27004E-02	0.10842E-01	2.3707	0.35976	9	13
15	17	0.31733E-02	0.99995E-02	2.5607	0.38082	10	12
16	16	0.35858E-02	0.96380E-02	2.5338	0.37205	11	11
17	15	0.35861E-02	0.96378E-02	2.5338	0.31007	12	10
18	14	0.35901E-02	0.96316E-02	2.5349	0.25805	13	9
19	13	0.37930E-02	0.91410E-02	2.6892	0.23711	14	8
20	12	0.41664E-02	0.67908E-02	3.9682	0.28632	15	7
21	11	0.47074E-02	0.66937E-02	3.6423	0.26372	16	6
22	10	0.47109E-02	0.44279E-02	5.6332	0.31291	17	5
23	9	0.67747E-02	0.42261E-02	3.9349	0.35624	18	4
24	8	0.69278E-02	0.41767E-02	3.8572	0.26190	19	3
25	7	0.73916E-02	0.86829E-03	6.7008	0.85128	20	2
26	6	0.74613E-02	0.16656E-05	7.8861	213.32	21	1

CHOW TEST - F DISTRIBUTION WITH DF1= 5 AND DF2= 22

TABLE 7

QUARTERLY PATENT APPLICATION FILINGS NAIVE MODEL

SAMPLE 1976Q1 - 1992Q3

```
R-SQUARE =    0.9553      R-SQUARE ADJUSTED =    0.9522
VARIANCE OF THE ESTIMATE-SIGMA**2 =  0.18754E-02
STANDARD ERROR OF THE ESTIMATE-SIGMA =  0.43306E-01
SUM OF SQUARED ERRORS-SSE=  0.10877
MEAN OF DEPENDENT VARIABLE =    10.321
LOG OF THE LIKELIHOOD FUNCTION(IF DEPVAR LOG) = -539.228
 AKAIKE (1973) INFORMATION CRITERION- LOG AIC =  -6.2029
 SCHWARZ(1978) CRITERION-LOG SC =  -6.0328
```

VARIABLE	SUM OF LAG COEFS	STD ERROR	T-RATIO	MEAN LAG
LFIL	1.0324	0.31247E-01	33.040	2.1333

VARIABLE NAME	ESTIMATED COEFFICIENT	STANDARD ERROR	T-RATIO 58 DF	P-VALUE	PARTIAL CORR.	STANDARDIZED COEFFICIENT	ELASTICITY AT MEANS
LFIL	0.36832	0.5016E-01	7.343	1.000	0.694	0.3613	0.3683
LFIL	0.41114	0.6631E-01	6.200	1.000	0.631	0.3943	0.4111
LFIL	0.25296	0.6916E-01	3.657	1.000	0.433	0.2335	0.2530
SURGE	0.39772	0.3298E-01	12.06	1.000	0.846	0.3606	0.3977
CONSTANT	-0.31758	0.3219	-0.9866	0.164	-0.128	0.0000	-0.3176

```
DURBIN-WATSON = 2.1277     VON NEUMANN RATIO = 2.1620     RHO = -0.08533
RESIDUAL SUM = -0.76782E-13  RESIDUAL VARIANCE =  0.18754E-02
SUM OF ABSOLUTE ERRORS=    2.1372
R-SQUARE BETWEEN OBSERVED AND PREDICTED = 0.9553
R-SQUARE BETWEEN ANTILOGS OBSERVED AND PREDICTED = 0.9590
RUNS TEST:    39 RUNS,    32 POSITIVE,    31 NEGATIVE, NORMAL STATISTIC =  1.6537
DURBIN H STATISTIC (ASYMPTOTIC NORMAL) = -0.73832
COEFFICIENT OF SKEWNESS =  -0.2209 WITH STANDARD DEVIATION OF 0.3016
COEFFICIENT OF EXCESS KURTOSIS =  -0.6058 WITH STANDARD DEVIATION OF 0.5948
JARQUE-BERA ASYMPTOTIC LM NORMALITY TEST
CHI-SQUARE =      1.6056 WITH  2 DEGREES OF FREEDOM

HETEROSKEDASTICITY TESTS
  E**2 ON YHAT:          CHI-SQUARE =      0.823 WITH 1 D.F.
  E**2 ON YHAT**2:       CHI-SQUARE =      0.814 WITH 1 D.F.
  E**2 ON LOG(YHAT**2):  CHI-SQUARE =      0.833 WITH 1 D.F.
  E**2 ON X (B-P-G) TEST:         CHI-SQUARE =      0.924 WITH   4 D.F.
  E**2 ON LAG(E**2)  ARCH TEST:   CHI-SQUARE =      2.636 WITH   1 D.F.
  LOG(E**2) ON X (HARVEY) TEST:   CHI-SQUARE =      1.555 WITH   4 D.F.
  ABS(E) ON X (GLEJSER) TEST:     CHI-SQUARE =      1.098 WITH   4 D.F.
```

TABLE 8

QUARTERLY PATENT APPLICATION FILINGS AUTOCORRELOGRAM
SAMPLE PERIOD 1976Q1-1992Q3

FILINGS IN NATURAL LOGARITHMS

MEAN = 10.311
VARIANCE = 0.39003E-01
STANDARD DEV. = 0.19749

LAGS	AUTOCORRELATIONS												STD ERR
1 -12	0.75	0.84	0.71	0.76	0.65	0.66	0.58	0.55	0.51	0.45	0.42	0.35	0.12
13 -24	0.33	0.25	0.25	0.19	0.20	0.11	0.11	0.05	0.06	-.01	0.00	-.07	0.39

LAGS	PARTIAL AUTOCORRELATIONS												STD ERR
1 -12	0.75	0.62	0.03	0.11	-.07	-.06	-.04	-.09	0.01	-.09	-.01	-.06	0.12

FILINGS IN FIRST DIFFERENCES OF NATURAL LOGARITHMS

MEAN = 0.71013E-02
VARIANCE = 0.17626E-01
STANDARD DEV. = 0.13276

LAGS	AUTOCORRELATIONS												STD ERR
1 -12	-.77	0.49	-.39	0.32	-.22	0.16	-.11	0.04	0.03	-.07	0.10	-.12	0.12
13 -24	0.16	-.18	0.16	-.17	0.21	-.19	0.17	-.19	0.19	-.18	0.17	-.19	0.23

LAGS	PARTIAL AUTOCORRELATIONS												STD ERR
1 -12	-.77	-.26	-.31	-.15	-.01	0.04	0.10	0.00	0.08	0.00	0.05	0.00	0.12

MODIFIED BOX-PIERCE (LJUNG-BOX-PIERCE) STATISTICS (CHI-SQUARE)

LAG	Q	DF	P-VALUE	LAG	Q	DF	P-VALUE
1	40.96	1	.000	13	86.43	13	.000
2	57.57	2	.000	14	89.17	14	.000
3	68.31	3	.000	15	91.32	15	.000
4	75.79	4	.000	16	93.79	16	.000
5	79.23	5	.000	17	97.81	17	.000
6	81.04	6	.000	18	101.12	18	.000
7	81.89	7	.000	19	103.92	19	.000
8	82.01	8	.000	20	107.26	20	.000
9	82.07	9	.000	21	110.91	21	.000
10	82.45	10	.000	22	114.11	22	.000
11	83.23	11	.000	23	117.14	23	.000
12	84.40	12	.000	24	121.04	24	.000

TABLE 9
QUARTERLY PATENT APPLICATION FILINGS ARIMA(1,1,1) MODEL

R-SQUARE = 0.6673 R-SQUARE ADJUSTED = 0.6568
VARIANCE OF THE ESTIMATE-SIGMA**2 = 0.59744E-02
STANDARD ERROR OF THE ESTIMATE-SIGMA = 0.77294E-01
AKAIKE INFORMATION CRITERIA -AIC(K) = -5.0294
SCHWARZ CRITERIA- SC(K) = -4.9298

PARAMETER	ESTIMATES	STD ERROR	T-STAT
AR(1)	-0.60018	0.1144	-5.248
MA(1)	0.54758	0.1210	4.525
CONSTANT	0.13219E-01	0.4533E-02	2.916

RESIDUALS

LAGS	AUTOCORRELATIONS												STD ERR
1 -12	0.04	-.08	-.08	0.17	0.08	0.04	-.05	0.02	0.06	-.01	0.04	0.00	0.12

MODIFIED BOX-PIERCE (LJUNG-BOX-PIERCE) STATISTICS (CHI-SQUARE)

LAG	Q	DF	P-VALUE
3	1.04	1	.307
4	3.18	2	.204
5	3.64	3	.303
6	3.78	4	.437
7	3.95	5	.557
8	3.97	6	.681
9	4.21	7	.755
10	4.23	8	.836
11	4.38	9	.885
12	4.38	10	.929

ANALYSIS OF RESIDUALS

VALUES RANGE FROM -0.3171 TO 0.2922

SAMPLE MOMENTS OF RESIDUALS (USING THE DIVISOR 66) :

MEAN = -0.5406873E-03
VARIANCE = 0.5702537E-02
SKEWNESS = -0.7533300
KURTOSIS = 10.25777
STUDENTIZED RANGE = 8.069505

TABLE 10

SPECIFICATION TESTS FOR LAG LENGTH IN AUTOREGRESSIVE DISTRIBUTED LAG MODEL

Sample 1976Q1-1992Q3

FIL	AGDP	FEE	AIC	BSC	$\bar{R}^2$
UNRESTRICTED MODELS STARTING WITH FOUR LAGS					
4,8,12,16	4,8,12,16	0-1	14.550	15.005	0.9575
4,8,12	4,8,12	0-1	14.525	14.890	0.9576
4,8	4,8	0-1	14.402	14.684	0.9610
4	4	0-1	15.376	15.580	0.8913
RESTRICTING LAGS ON FILINGS TO THREE					
4,8,12	4,8,12,16	0-1	14.523	14.939	0.9581
4,8,12	4,8,12	0-1	14.492	14.820	0.9584
4,8,12	4	0-1	14.456	14.748	0.9593
RESTRICTING LAGS ON FILINGS TO TWO					
4,8	4,8,12,16	0-1	14.484	14.863	0.9591
4,8	4,8,12	0-1	14.495	14.824	0.9583
4,8	4	0-1	14.495	14.741	0.9566
RESTRICTING LAGS ON FILINGS TO ONE					
4	4,8,12,16	0-1	14.483	14.824	0.9585
4	4,8,12	0-1	14.561	14.853	0.9548
4	4,8	0-1	14.504	14.750	0.9562

TABLE 11

QUARTERLY PATENT APPLICATION FILINGS ECONOMETRIC MODEL
SAMPLE 1976Q1 - 1992Q3

R-SQUARE = 0.9657 R-SQUARE ADJUSTED = 0.9610
VARIANCE OF THE ESTIMATE-SIGMA**2 = 0.15860E+07
STANDARD ERROR OF THE ESTIMATE-SIGMA = 1259.4
SUM OF SQUARED ERRORS-SSE= 0.80886E+08
MEAN OF DEPENDENT VARIABLE = 31329.
LOG OF THE LIKELIHOOD FUNCTION = -500.582
AKAIKE (1973) INFORMATION CRITERION- LOG AIC = 14.402
SCHWARZ(1978) CRITERION-LOG SC = 14.684

VARIABLE	SUM OF LAG COEFS	STD ERROR	T-RATIO	MEAN LAG
FIL	0.47237	0.70209E-01	6.7280	5.6077
AGDPQ	8.8869	0.87921	10.108	3.4866
ERFEE	-2.2129	1.9548	-1.1320	1.4223

VARIABLE NAME	ESTIMATED COEFFICIENT	STANDARD ERROR	T-RATIO 51 DF	P-VALUE	PARTIAL CORR.	STANDARDIZED COEFFICIENT	ELASTICITY AT MEANS
FIL	0.28252	0.7324E-01	3.858	1.000	0.475	0.2554	0.2723
FIL	0.18986	0.6984E-01	2.718	0.996	0.356	0.1485	0.1765
AGDPQ	10.028	1.075	9.331	1.000	0.794	0.7252	1.3281
AGDPQ	-1.1407	0.4352	-2.621	0.006	-0.345	-0.1257	-0.1443
ERFEE	0.93453	2.062	0.4532	0.674	0.063	0.0227	0.0083
ERFEE	-3.1474	2.620	-1.201	0.118	-0.166	-0.0749	-0.0281
SURGE	9029.1	963.2	9.374	1.000	0.795	0.2629	0.0000
CONSTANT	-19197.	2350.	-8.168	0.000	-0.753	0.0000	-0.6128

DURBIN-WATSON = 2.2698 VON NEUMANN RATIO = 2.3089 RHO = -0.14273
RESIDUAL SUM = -0.76056E-10 RESIDUAL VARIANCE = 0.15860E+07
RUNS TEST: 32 RUNS, 31 POSITIVE, 28 NEGATIVE, NORMAL STATISTIC = 0.4151
DURBIN H STATISTIC (ASYMPTOTIC NORMAL) = -1.3261
COEFFICIENT OF SKEWNESS = -0.3554 WITH STANDARD DEVIATION OF 0.3112
COEFFICIENT OF EXCESS KURTOSIS = 0.0721 WITH STANDARD DEVIATION OF 0.6133

GOODNESS OF FIT TEST FOR NORMALITY OF RESIDUALS - 12 GROUPS

OBSERVED	1.0	1.0	3.0	3.0	8.0	12.0	11.0	12.0	7.0	0.0	1.0	0.0
EXPECTED	0.4	1.0	2.6	5.4	8.8	11.3	11.3	8.8	5.4	2.6	1.0	0.4

CHI-SQUARE = 6.9283 WITH 2 DEGREES OF FREEDOM

JARQUE-BERA ASYMPTOTIC LM NORMALITY TEST
CHI-SQUARE = 1.1824 WITH 2 DEGREES OF FREEDOM

HETEROSKEDASTICITY TESTS
E**2 ON YHAT: CHI-SQUARE = 0.161 WITH 1 D.F.
E**2 ON YHAT**2: CHI-SQUARE = 0.214 WITH 1 D.F.
E**2 ON LOG(YHAT**2): CHI-SQUARE = 0.116 WITH 1 D.F.
E**2 ON X (B-P-G) TEST: CHI-SQUARE = 4.440 WITH 7 D.F.
E**2 ON LAG(E**2) ARCH TEST: CHI-SQUARE = 0.007 WITH 1 D.F.
LOG(E**2) ON X (HARVEY) TEST: CHI-SQUARE = 3.168 WITH 7 D.F.
ABS(E) ON X (GLEJSER) TEST: CHI-SQUARE = 3.436 WITH 7 D.F.

TABLE 11 (continued)

RESIDUAL CORRELOGRAM
LM-TEST FOR HJ:RHO(J)=0, STATISTIC IS STANDARD NORMAL

LAG	RHO	STD ERR	T-STAT	LM-STAT	DW-TEST	BOX-PIERCE-LJUNG
1	-0.1415	0.1302	-1.0867	1.1731	2.2698	1.2419
2	0.2013	0.1302	1.5465	1.7207	1.5748	3.8015
3	-0.0294	0.1302	-0.2260	0.2428	2.0216	3.8572
4	0.1806	0.1302	1.3872	1.6586	1.5848	5.9914
5	-0.0819	0.1302	-0.6294	0.7070	2.0867	6.4388
6	0.1676	0.1302	1.2877	1.5290	1.5301	8.3474
7	-0.0864	0.1302	-0.6635	0.7395	2.0290	8.8637
8	0.0211	0.1302	0.1621	0.1881	1.8071	8.8952
9	-0.1311	0.1302	-1.0068	1.1457	2.0985	10.1318
10	-0.0724	0.1302	-0.5560	0.6596	1.9699	10.5166
11	0.0371	0.1302	0.2849	0.3321	1.7245	10.6198
12	-0.2013	0.1302	-1.5464	1.8709	2.2001	13.7235
13	-0.0476	0.1302	-0.3654	0.4322	1.8525	13.9006
14	-0.1110	0.1302	-0.8525	1.0121	1.9691	14.8857
15	-0.1405	0.1302	-1.0793	1.3003	2.0041	16.5008
16	0.0565	0.1302	0.4336	0.5269	1.5906	16.7675
17	-0.0252	0.1302	-0.1936	0.2347	1.7073	16.8220
18	-0.0283	0.1302	-0.2174	0.2646	1.6683	16.8923

LM CHI-SQUARE STATISTIC WITH 18 D.F. IS 14.162

TABLE 12

COMPARISON OF THE FORECAST PERFORMANCE OF THE ANNUAL MODELS

		NAIVE		ARIMA		ECONOMETRIC	
Year	Actual Level	1 Year	2 Year	1 Year	2 Year	1 Year	2 Year
1985	116,427	6,524		11,902		-1,180	
1986	121,611	8,432	10,629	5,574	12,043	-380	-1,470
1987	126,407	9,766	12,034	4,020	7,643	-2,778	-3,127
1988	137,069	13,764	19,125	9,444	12,145	4,229	1,857
1989	151,331	15,980	26,550	14,627	21,094	12,085	15,455
1990	163,571	11,047	26,435	12,150	23,155	10,386	21,199
1991	167,715	-3,548	12,637	2,041	12,093	13,654	23,941
ME		8,852	17,902	8,537	14,696	5,145	9,643
MPE		6.50	12.24	6.17	10.02	3.15	5.85
RMSE		10,613	19,092	9,590	15,697	8,172	14,588
MAE		9,866	17,902	8,537	14,696	6,385	11,175
MAPE		7.10	12.24	6.17	10.02	4.16	7.08

TABLE 13

COMPARISON OF THE FORECAST PERFORMANCE OF THE QUARTERLY MODELS

		NAIVE		ARIMA		ECONOMETRIC	
Year	Actual Level	1 Year	2 Year	1 Year	2 Year	1 Year	2 Year
1985	116,427	7,905		-377		2,487	
1986	121,611	6,712	10,809	7,019	2,919	-1,112	100
1987	126,407	2,170	7,570	4,375	9,747	-1,181	-1,673
1988	137,069	8,759	9,352	8,147	12,378	5,795	3,288
1989	151,331	7,327	12,804	12,345	19,951	8,413	13,158
1990	163,571	-1,147	3,873	8,164	21,006	4,983	12,195
1991	167,715	-7,006	-8,679	-716	7,283	3,631	7,097
ME		3,531	5,955	5,565	12,214	3,288	5,694
MPE		2.91	4.56	3.94	8.25	2.18	3.59
RMSE		6,470	9,272	7,116	13,839	4,647	8,019
MAE		5,861	8,848	5,878	12,214	3,943	6,252
MAPE		4.31	6.28	4.15	8.25	2.71	4.03

References

Acs, Z. J. and D. B. Audretsch, "Patents as a Measure of Innovative Activity," *Kyklos*, vol. 42 (1989): p. 171-80.

Bosworth, D. and Westaway, T., "The Influence of demand and supply side pressure on the quantity and quality of inventive activity," *Applied Economics*, vol. 16 (1984): p. 131-146.

Box, George and Gwilym Jenkins, Time Series Analysis: Forecasting and Control, Holden-Day, Inc., San Francisco, 1970.

CITIBASE, 1992, *Citibank Economic Database* (1946-present), Citibank, New York, N.A.

Durbin, J. and Watson, G. S., "Testing for Serial Correlation in least-Squares Regression," *Biometrica,* vol. 38 (1971): p. 159-177.

Fair, Ray C., Specification, Estimation, and Analysis of Macroeconometric Models, 1984, Harvard University Press.

Granger, C.W.J. and Newbold, P., Forecasting Economic Time Series, Second Edition, Academic Press, Inc., 1986.

Griliches, Z., "Patents: Recent Trends and Puzzles," *Brookings Papers on Economic Activity: Microeconomics,* (1989): p. 291-330.

Griliches, Z., "Patent Statistics as Economic Indicators: A Survey," *Journal of Economic Literature*, vol. 27 (December, 1990): p. 1661-1707.

Hall, B., Griliches, Z., and Hausman, J., "Patents and R&D: Is There a Lag?" *International Economic Review,* vol. 27, no. 2 (1986): p. 265-83.

Holt, C. C., "Forecasting Seasonal and Trends by Exponentially Weighted Moving Averages." Office of Naval Research, Memorandum, no. 52, 1957.

Judge, G., Hill, Carter Hill R., Griffiths, W.E., Lutkepohl, H., and Lee, T.C., Introduction to the Theory and Practice of Econometrics, Second Edition, John Wiley and Son, 1988.

OECD "International Surveys of the Resources Devoted to R&D by OECD Member Countries." Directories for Science, Technology and Industry. Organization for Economic Cooperation and Development. Paris, 1987.

Pavitt, K., "R&D, Patenting and Innovative Activity: A Statistical Exploration," *Research Policy*, vol. 11, no. 1 (1982): p. 33-51.

Scherer, F.M., "The Propensity to Patent", International Journal of Industrial Organization (1983): p. 107-128.

Schmookler, J., "The Level of Inventive Activity," *Rev. Econ. Statist.,* vol. 36, no. 2 (1954):p. 183-190

Stellwagen, E. A. and Goodrich, R.L., *Forecast Pro for Windows* (Business Forecast Systems, Inc., MA), 1992.

White, K.J., S.A. Haun, N.G. Horsman and S.D. Wong, 1993, *SHAZAM: Econometrics Computer Program - User's Reference Manual, Version 7.0* (Mcgraw-Hill, New York).

Wilson, J. Holton and Barry Keating, Business Forecasting, Irwin, IL., 1990.

LABOR AND PRODUCTIVITY

MODERATOR Norm Saunders, Bureau of Labor Statistics

PRESENTATION 1 Howard Fullerton, Bureau of Labor Statistics

PRESENTATION 2 Lea Swanson, Environmental Protection Agency

PRESENTATION 3 Lloyd Teigen, Economic Research Service

Labor and Productivity

Measuring Rates of Labor Force Dynamics

Howard N Fullerton, Jr., Office of Employment Projections, Bureau of Labor Statistics

Keywords: Entrants, leavers, numerator analysis, projections

1. Introduction

In its most recent description of its labor force projection, BLS attributed the slowing labor force growth to a lag in population growth, noting that most labor force growth is explained by population growth (Fullerton, 1991). In a later section in the same document, the slower growth is attributed to an increase in the number of persons leaving the labor force while the number of entrants remained virtually the same. Although these are not competing explanations of slower labor force growth, it is not obvious that they are the same explanation. The analysis of labor force growth and population growth was done by dividing the growth rate of the labor force into components. The analysis of entrants and leavers was done by calculating the number of entrants and leavers for the 1975-90 and 1990-2005 periods and comparing them, a much lower level of sophistication.

In this paper, we will first explain how BLS measures entrants and leavers; then examine alternate measures of the rate of entry or separation.

BLS measures labor force *entrants* by comparing a specific sex-race birth cohort in 2005 with itself in the base year (1990). For most purposes, entrants is a better or more useful measure than net change. To be in the labor force, one has to be at least 16; thus, in 2005, none of those under 30 could have been in the labor force in 1990. All 16 to 29 year-olds, by definition, must have entered the labor force between 1990 and 2005. Older cohorts are examined by five-year of age groups and compared with their labor force size in the base year. This was done by sex and by race or Hispanic origin. Older cohorts had more members in the labor force in 1990 than in 2005, a difference termed labor force *leavers*. The difference between the aggregated entrants and leavers is *net change*, and should equal the net change calculated by comparing the total labor force numbers for the two years.

2. An example: calculating female entrants

To make the calculation of net entrants more concrete, we have included a summary table of entrants for women with no race or Hispanic origin detail. Those 16 to 19 in 2005 were aged 1 to 4 in 1990; all their 2005 labor force would enter between 2001 and 2005. Starting with ages 40 and over in 1990, the numbers in the labor force at each age decreased; these are the leavers. What is striking is that women are entering the labor force for an extended period of time; the cohort born 1951-55, which would be 55 to 59 in 2005, has more than 150,000 entrants over the 1990-2005 period. The baby boom generation is split between entering and leaving; the older groups, born 1946-50 are leaving; those born 1951-65 are still entering. The baby boom generation would contribute 3 million entrants to the 27 million entrants and 1.3 million leavers to the 12.7 million leavers.

3. Net change

The 2005 labor force of women is larger than in 1990 age group by age group. This reflects increased participation as well as population increases. However, between 1975 and 1990, the passage of the baby boom resulted in a decrease in some age groups. For men, especially, labor force participation rates did not increase enough to offset the population decline. Thus, the net change would have been negative.

Table 1.	Entrants to and net change in the male labor force, 1975-90					
	Labor force	Entrants	Share of entrants	Labor force	Net change	Share of net change
	1975	1975-90	1975-90	1990	1975-90	Percent
Men, 16 and over	56,299	28,135	100.0	68,232	11,935	100.0
16 to 19	4,805	3,866	13.7	3,866	-939	-7.9
20 to 24	7,565	7,291	25.9	7,291	-274	-2.3
25 to 29	7,734	9,583	34.1	9,583	1,849	15.5
30 to 34	6,457	5,425	19.3	10,230	3,773	31.6
35 and over	29,737	1,970	7.0	37,264	7,527	63.1

Using share of net change yields the surprising and perhaps not very meaningful result that workers over the age of 35 accounted for 63 percent of net change, though only about 7 percent of entrants. Men at the younger ages accounted for 10 percent of net change, considered as a share of "net" or "new" entrants.

Calculating shares when some components are negative and some positive leads to difficulties between 1985 and 2000, the younger portion of the labor force was projected to decrease in size. Analysis of net change is fraught with difficulty, even if the components reviewed do not have any apparent negative elements. Of course, the advantage of using net change is that the underlying age detail is not needed.

BLS has calculated that 55.9 million persons entered the labor force between 1975 and 1990. It has projected that 55.8 million will enter over the 1990-2005 period, virtually the same number. It also has estimated that 24.9 persons left the labor force over the earlier period and that 29.8 million will leave over the 1990-2005 period. The 1975 labor force was 93.8 million, the 1990 labor force 124.8 million. Adding 56 million persons to a 94 million labor force should have a larger impact than adding 56 million persons to a labor force of 125 million. Investigating, or quantifying this observation should help us understand the relation between population growth and entrants, between population aging and increasing numbers of leavers. The number that we seek is a rate. We have been conducting numerator analysis, what we must do is analyze rates; to do that, we need a denominator.

We of course know our numerators, which are available by sex and race or Hispanic origin and by age. However, what are the denominators---what do demographers consider a rate?

4. What is a rate

Shryock and Siegel (1973) discuss the difference between rates and ratios: "The term 'rate' most appropriately applies to the number of demographic events in a given period of time divided by the population at risk during that period. The population at risk is usually only approximated. It may be the population at the middle of the period, (which is roughly the average population), the population at the beginning of the period, or a more complex definition." The cost of obtaining the more complexly calculated population at risk, or a mid-year estimate may be a significant factor in determining the denominator.

The people at risk to entering the labor force (the events we are measuring) are those not in the labor force, the population less the labor force, and those emigrating to the United States. Those at risk to leaving the labor force are those in the labor force.

We may also distinguish between "crude" and more refined rates. For example, the crude birth rate is the births in a year divided by the mid-year population. Not all the population (babies and older women) are at risk to having a child at all and we are counting both parents. Further, risks vary significantly across population groups, specifically, marital status groups. For our measures, a quick approach would be the total population as the at-risk denominator for entrants and the labor force for leavers.

For more refined rates of entry, we would use the not in the labor force population group, with an allowance for immigration. We would also provide age-sex-race or Hispanic origin-specific rates. Hispanic origin data is not available before 1980 and the refinements for race and Hispanic origin will not be pursued in this paper.

Demographers follow actuarial science in distinguishing between probabilities and central rates. Probability-type rates are based on the population at risk at the beginning of the period; central rates are based on a mid-period group at risk. Shryock and Siegel point out that when the population is open, specifically, receiving immigrants, then average populations or person-years are used.

The most difficult problem in providing a denominator for entrants is supplying an estimate of immigrants. Generally, the Census Bureau provides estimates of net immigration with their estimates of the components of change {reference}. Although there is an assumption of immigration and emigration in these estimates, they are generally not available nor are they available by age. This suggests using central rates rather than probability-type rates. An intermediate value would have immigrants embodied in the estimates.

5. Using "probability-type" denominators

The easiest set of rates to compute uses the total population for both entrants and leavers. If the same base is used for both groups, the difference is meaningful:

Table 2. Crude rates of entry to and departure from the labor force, 1975-90 and 1990-2005, probability-type measures
[Numbers in thousands]

	1975-90		Population		1990-2005		Population
	Entrants	Leavers	1975		Entrants	Leavers	1990
Total	55,915	24,901	215,465		55,798	29,851	249,415
Men	28,135	16,199	104,876		28,197	17,090	121,599
Women	27,780	8,702	110,589		27,601	12,761	127,816
	rates per thousand		difference		rates per thousand		difference
Total	260	116	144		224	120	104
Men	268	154	114		232	141	91
Women	251	79	173		216	100	116

The entrant rates declined from 260 per thousand for the 1975-90 period to 224 per thousand 1990-2005. This represents a decline of more than ten percent. Given that the number of entrants remained essentially the same, while the labor force grew by 26 million, the rates would be expected to drop. Female and male rates dropped about the same amount. Notice that their rates of entry, like their levels, are fairly close.

The rate of leaving the labor force increased modestly from 116 per thousand over the 1975-90 period to 120 per thousand. For the labor force as a whole, the slowing of labor force growth was influenced more by the slowdown in entrants than the increase in leavers. The male rate of leaving dropped by 10 percent. That is, men are projected to be *less* likely to leave the labor force over the 1990-2005 period than they did over the 1975-90 period. The rate of departure for women increased by 27 percent or they are projected to be more likely to leave the labor force than in the earlier period The increase was more than the amount of the decrease in the rate of entry. The female labor force is projected to grow more slowly both because the rate of entering dropped and because the rate of leaving increased.

Since both rates are calculated with the same base, it is meaningful to subtract the rates. This rate of change is lower for the projected period; measuring the slower growth of the labor force. The male labor force has been growing more slowly than the female labor force; this is projected to continue.

6. Refining the rates--departures

Only those in the labor force may leave it; for leavers, a more meaningful measure would use the labor force as a denominator:

Table 3. Leavers, labor force, and rates of leaving the labor force, 1975 and projected 1990-2005
[In thousands]

Sex and age	Labor force	Leavers 1975-90		Labor force	Leavers 1990-2005	
		Number	per thousand		Number	per thousand
Age in 1975 or 1990	1975			1990		
Total, 1 year or older	93,775	24,901	266	237,893	29,851	125
Men, 1 year or older	56,298	16,199	288	181,339	17,090	94
1 to 14						
15 to 19	4,805			3,866		
20 to 24	7,565			7,291		
25 to 29	7,734			9,583		
30 to 34	6,457	220	34	10,230	346	34
35 to 39	5,288	348	66	9,261	738	80
40 to 44	5,112	1,098	215	8,007	1,564	195
45 to 49	5,299	2,528	477	6,237	2,988	479
50 to 54	5,101	3,909	766	4,939	3,644	738
55 to 59	4,205	3,671	873	4,012	3,447	859
60 and	4,732	4,425	935	4,805	4,363	908
Women, 1 year or older	37,477	8,702	232	56,554	12,761	226
1 to 14						
15 to 19	4,065			3,544		
20 to 24	6,185			6,552		
25 to 29	4,965			7,837		
30 to 34	3,708			8,152		
35 to 39	3,275			7,640		
40 to 44	3,231			6,936		
45 to 49	3,376			5,339		
50 to 54	3,307			3,977	54	14
55 to 59	2,662	172	65	3,059	1,310	428
60 and	2,703	8,530	3,156	3,518	11,396	3,239

This table also shows a decline in the rate of leaving the labor force between the 1975-90 and 1990-2005 periods. The drop is substantial, by more than half. However, only the male rate dropped, by almost two-thirds. Female departures remain at the same level. This measures leavers from the labor force, not asking their destination, retirement, death, or emigration. If we were concerned about a specific outcome, we would have to take mortality and emigration into account.

This more refined measure points in a different direction than the crude rate, which indicated no change in the rate of leaving the labor force. The crude rate did preserve the relationship between men and women's rates.

The age-specific data indicate that men begin leaving the labor force sooner than men (or possibly that men complete their entry to the labor force sooner). The rates rise until the oldest age break, 60 and over. Since the usual age of entitlement to retirement occurs after 60, that we do not have information on the pattern of withdrawal is unfortunate. It is an aspect of the wide interval 15 years over which we calculate the changes. By this measure, few women permanently leave the labor force before age 60. The women considered to depart for child-rearing are outweighed in both time intervals by the number of women entering. One difference between the historic period and that being projected is that the most rapid increase in the rate of leaving the male labor force occurs between ages 45 to 49 and 50 to 54 in 1975-90. According to the projections, that will be between ages 40 to 44 and 45 to 49.

For entrants, the concern is not about destination, but about sources. Although most entrants come from the United States, and there few deaths at the ages of entry, some entrants do arrive from outside the country.

Table 4. Entrants, and not in the labor force, 1975-90 and projected 1990-2005
In thousands

Age and sex	Not in the Labor Force 1975	1975-90 Entrants number	1975-90 rate per thousand	Not in the Labor Force 1990	1990-2005 Entrants number	1990-2005 rate per thousand
Total, 1 year or older	118,541	55,915	472	120,618	55,798	463
Men, 1 year and older	46,967	28,135	599	51,315	28,197	549
1 to 4	6,627	3,866	583	7,608	4,575	601
5 to 9	8,972	7,291	813	9,247	7,989	864
10 to 14	10,534	9,583	910	8,805	8,269	939
15 to 19	5,952	5,425	911	5,243	4,820	919
20 to 24	2,075	1,697	818	2,453	1,817	741
25 to 29	791	273	345	1,071	546	510
30 to 34	501			671	92	137
35 to 39	367			638	89	140
40 and over	11,148			15,579		
Women, 1 year and older	71,574	27,780	388	69,303	27,601	398
1 to 4	6,342	3,544	559	7,258	4,218	581
5 to 9	8,623	6,552	760	8,817	7,266	824
10 to 14	10,112	7,838	775	8,385	7,227	862
15 to 19	6,400	4,087	639	5,101	3,953	775
20 to 24	3,492	1,455	417	2,835	1,706	602
25 to 29	3,695	1,971	533	2,737	1,734	634
30 to 34	3,465	1,631	471	2,854	1,272	446
35 to 39	2,658	702	264	2,435	219	90
40 and over	26,787			28,881	5	0

According to this more refined measure, the rate of entry between 1990 to 2005 remained the same. Both sex groups remained about the same, with women's entry increasing slightly, men's decreasing slightly. The crude rate of entry indicated a more greater drop in the rate of entry, with even less difference in the change in rates of men and women. The not in the labor force population (NILF) increased for men and decreased for women. This decrease reflects women's labor force growing more rapidly than their population. The continued growth of the male NILF population simply reflects population growth.

Within the age data, men completed their labor force entry by age 29, more technically, the size of the labor force cohorts did not grow for any cohorts aged 30 and over in the first year of the period. For women, this age of entry extended to 35 to 39. Judging the entry rates or levels on the basis of labor force rates for those aged 16 to 24 yields an underestimate of the number of women entering the labor force, 4.3 million in the cohorts aged 25 and over entered the labor force between 1975 and 1990. Although entry is projected to extend over a longer period, the number of women entering the labor force after age 24 is projected to be smaller. The number of men projected to enter the labor force at these "older" ages is projected to more than double.

References

Fullerton, Howard N, Jr.,(1991), "Labor force projections: the baby boom moves on," *Monthly Labor Review*, November, pp. 31-44, reprinted in *Outlook: 1990 - 2005*, (BLS Bulletin 2402), April 1992.

Gendell, Murray, and Siegel, Jacob S., (1992), "Trends in retirement age by sex: 1950-2005," *Monthly Labor Review*, July, pp. 22-29.

Shyrock, Henry J., and Siegel, Jacob S., and Associates, (1973), *The Methods and Materials of Demography*, (Washington, U.S. Bureau of the Census)

U.S. Bureau of the Census (1990), "United States Population Estimates, by Age, Sex, Race, and Hispanic Origin: 1980 to 1988," *Current Population Reports*, Series P-25, No 1045, (Washington), pp. 24-29, and Table 5, "Estimates of the Components of Population Change for Calendar Years: 1940 to 1988."

Pollution Prevention and Unemployment Prevention are Parallel Arms of a Single Strategy for Sustainable Development

Lea Swanson*, Environmental Protection Agency

Summary

This paper argues that pollution prevention and unemployment prevention are parallel arms of a single strategy for national competitiveness -- defined to include the achievement of quantitative and qualitative societal goals. The level of national competitiveness achieved ultimately depends on productivity. And productivity, in turn, can be increased by reducing the inputs required to achieve the same level of production. However, this is only half of the story. If the technology/process applied to increase productivity results in spoiled production inputs, the spoiled inputs cannot be used for future production. Similarly, if the technology/process results in reduced labor inputs, and that labor cannot be employed in another industry, the reduced input costs will be offset by increased expenditures required to support (welfare) and retrain displaced labor. A productivity increase is not realized if the resources released for use elsewhere, cannot be used. Moreover, productivity increases must increase efficiency and effectiveness if they are to add value to national competitiveness.

Drivers of Economic Prosperity

Competition drives national economic development: that is, competition -- at both an international and local level -- for trade, skilled manpower, innovative technology, and strategic advantage. Indeed, when Michael Porter, a recognized authority on competitiveness,[1] refers to phases of economic development (factor driven, investment driven, innovation driven, wealth driven) he talks in terms of the "stages of national competitiveness". He also links vigorous local competition between firms to the creation and maintenance of competitive advantage in the global arena. This link is critical to understanding Porter's definition of competition as dynamic and evolving; a process that has no equilibrium. Rather, he says "competition is a constantly changing landscape in which new products, new ways of marketing, new production processes, and whole new market segments emerge. Static efficiency at a point in time is rapidly overcome by a faster rate of progress."[2] The essential character in competition is thus innovation and change. And innovation here refers both to improvements in technology and better processes of production, with productivity improvements the logical corollary.

Porter uses the terms economic prosperity and competitiveness interchangeably. Competition, then, not only drives economic prosperity, it is also an end in itself. The level of competitiveness achieved by a nation depends, in turn, on the productivity with which national resources are employed. Productivity is a function of innovation, and innovation "requires sustained investment in research, physical capital, and human resources" to develop skills and knowledge. Mapping back full circle to the vigor of local competition, Porter asserts that rivalry drives innovation and "upgrading of competitive advantage". Rivalry is especially important to Porter's hypothesis

---

*Lea Swanson serves as Special Assistant for Policy, Planning and Evaluation, Office of the Assistant Administrator, U.S. Environmental Protection Agency. Ms Swanson earned her Masters of Public Administration at Harvard University. She is currently finalizing her thesis on national competitiveness and sustainable development to complete requirements for her Masters of Business Administration at Monash University (Australia).

[1] Porter, Michael E., teaches at the Harvard Business School, and is an advisor to leading companies all over the world. He is the author of Competitive Strategy (1980), Competitive Advantage (1985), and The Competitive Advantage of Nations (1990), all published by The Free Press, New York.

[2] Porter, Michael, E. The Competitive Advantage of Nations, The Free Press, New York, 1990, p.20.

because it creates a fear of failure and thereby overcomes inertia.[3] For Porter, corporate rivalry is the critical driver of national competitiveness.[4]

Recent analyses of national competitiveness offer a range of perspectives that support Porter, disagree with Porter, or reinforce Porter's conclusions but through a different theoretical framework. For example, Stephan Schmidheiny and the Business Council for Sustainable Development[5] fully concur with Michael Porter: competitiveness depends on a nation's capacity to innovate, which in turn depends on the intensity of local competition. Efficiency is highlighted as an important determinant of competitiveness. This lends support to Porter's emphasis on resource productivity as an important measure of competitiveness. Another important determinant of competitiveness cited by the Business Council is national infrastructure. "Macroeconomic stability; free, open markets; clear property rights; and political stability" are important elements of an attractive environment for investment.

The World Bank,[6] on the other hand, characterizes economic prosperity as the driver of national development. The report states that "economic development is an essential means for enabling national development, but in itself is a highly imperfect proxy for progress." Rather, national development is about improving the well-being of people. As such, the essential elements of national development are higher standards of living; better education, health, and opportunity; improved environmental protection, and a more equitable distribution of the burdens and benefits of national policies. The World Bank effectively turns Porter's concept of economic prosperity/competitiveness on its head, by describing it as a means to achieving broad societal goals, rather than an end in itself. Yet, somewhat paradoxically, the World Bank supports Porter's framework for national competitiveness: "By promoting specialization and competition and encouraging technological progress, open trade and investment policies raise productivity and improve efficiency -- including efficient use of environmental resources."[7] The difference lies in the qualifications that the World Bank adds to Porter's concept of economic prosperity/competitiveness as an end in itself. For the World Bank "what matters is that the overall productivity of the accumulated capital -- including its impact on human health and aesthetic pleasure, as well as on incomes -- more than compensates for any loss from depletion of natural capital." It is to this end that the World Bank identifies macroeconomic policies such as open trade and foreign investment as the drivers of national development.

A view that is diametrically opposed to the World Bank's perspective comes from a study by the McKinsey Global Institute[8], together with Nobel Laureate Robert Solow. The report cuts a straight line from productivity to competitiveness. There is no intervening process of competition emphasized here. Rather, productivity is the means to an end. According to the report, productivity, including service sector productivity, is the ultimate measure of international competitiveness. There are no broader societal goals to be achieved here. Insofar as

---

[3] ibid. I have focussed on only one of the four preconditions for competitive advancement discussed by Porter. The four preconditions are:

1. factor creation mechanisms
2. motivation
3. domestic rivalry
4. demand upgrading

[4] In my forthcoming (December 1993) MBA thesis entitled "Resilience: A Three Dimensional Metric of National Development", I do not concur with Porter's assertion that corporate rivalry is the critical driver of national competitiveness. I argue instead for the paramount role of cooperation in guiding competitive forces toward win/win solutions. For example, when I explore the role of corporate rivalry/worker skill and experience in driving corporate performance, it is only when I introduce the concept of co-evolution that the dynamic interaction of both competitive and cooperative forces are invoked. In the resultant scenario, mutually interdependent firms are able to maintain their independence and reap the benefits of cooperative efforts. Likewise when I explore the role of innovation and experimentation/regulations and standards, co-evolution as opposed to competitiveness offers a dynamic that utilizes both competitive and cooperative forces and underscores the value of strategic alliances. And cooperation is not limited in these examples to the level of the corporation. At the core of the concept of co-evolution is the notion of cooperative partnerships in all facets of activity -- between employees and management; between competing firms and industries; between the regulated community and the regulators; between government, the private sector, and the public, and between nations.

[5] Schmidheiny, Stephan, with The Business Council for Sustainable Development, Changing Course, The MIT Press, Cambridge, Ma., 1992.

[6] World Development Report 1992, Oxford University Press, New York, 1992.

[7] ibid.

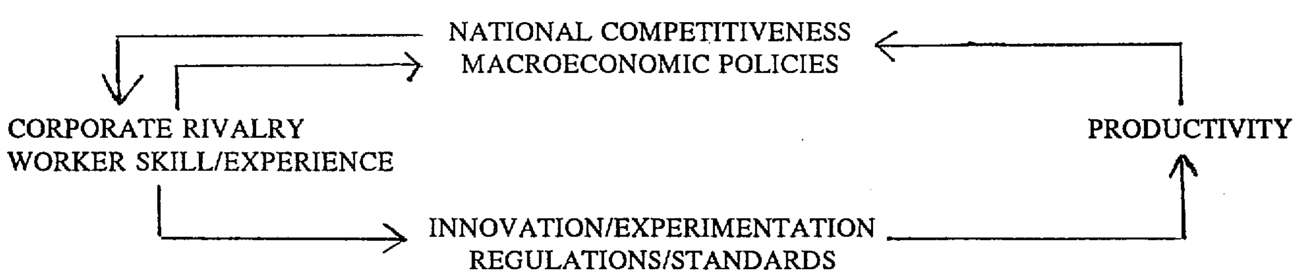
[8] Sylvia Nasar, "U.S. Rate of Output Called Best -- Study Affirms Lead Over Foreign Rivals," New York Times, October 13, 1992.

productivity describes, "not only a nation's standard of living, but also status in the world" -- productivity-cum-competitiveness is an end in itself.

On the topic of competitiveness, U.S. Secretary for Labor, Robert B. Reich, has quipped: "rarely has a term in public discourse gone so directly from obscurity to meaninglessness without an intervening period of coherence,"[9] That hasn't stopped Reich from offering an opinion on just what competitiveness means and how it should be measured. Reich firmly asserts that competitiveness is "the capacity of a nation's citizens to add value to the global economy without going into ever deeper debt."[10] It depends, he says, on the jobs that Americans perform which in turn determines standards of living and whether or not these standards are sustainable. National competitiveness, in this context, is synonymous with a nation's standard of living. Finally, the jobs that Americans perform depends on their respective skills and insights -- attributes that are a function of both training and experience. Of particular importance is the skill that finds the right fit between particular technologies and particular markets. "Experimentation", says Reich, "is the lifeblood of the high-value enterprise, because customization requires continuous trial and error." It is here that the corporation plays its vital role. The high-value corporation offers problem solvers, problem identifiers, and facilitators "membership in a small group engaged in a common task, sharing the risks of defeat and the potential rewards of victory." Thus Reich identifies the skill and experience of a nation's workforce (as distinct from Porter's corporate rivalry) as the critical driver of national prosperity and hence competitiveness.[11]

What's important here is that for Porter, national prosperity is synonymous with competitiveness, whereas for Reich, national competitiveness is synonymous with national prosperity. In each case the latter concept defines the former. Reich, not unlike the World Bank, defines national prosperity/competitiveness to include both the quantitative and qualitative terms imbued in standard of living and quality of life. Porter, Schmidheiny, and Solow, on the other hand, draw their definitions of national prosperity/ competitiveness from the narrower, and more immediate, quantitative perspective of economic competition.

Notwithstanding the differences in definitional perspectives, the components of the views discussed form a logical framework for exploring the process of national economic development. In summary, competitive forces, both global and local, drive a nation's economic development. However it is defined, national competitiveness, reflected in a country's macroeconomic policies, affects local corporate rivalry, and at the same time, is influenced by local competitive forces. Local rivalry, and worker skill and experience, drive innovation and experimentation. Acting within the parameters of a nation's regulatory environment, innovation, in turn, drives productivity improvements. The level of competitiveness achieved by a nation ultimately depends on national productivity.



Productivity

Productivity, then, reflects technology, innovation, invention, skills, and investment at both the macro- and micro-economic levels. It can be defined as the rate at which inputs (land, labor, capital, materials) are transformed into outputs, and it is usually measured by dividing GDP by total population. However, on recalling the World Bank's qualifications -- that the overall productivity of accumulated capital must more than compensate for any loss from the depletion of natural capital -- GDP per capita, as currently defined, is an inadequate measure of national

[9] Bob Davis, "Competitiveness is a big word in D.C., just ask the Vice-President," Wall Street Journal, July 1, 1992.

[10] Reich, R.B., "Who Is Us", Harvard Business Review, 1990.

[11] Reich, Robert B., The Work of Nations, Alfred A. Knopf Inc., U.S.A., 1991, p.119.

productivity.[12] Indeed the shortcomings of conventional measures of national income and productivity continue to be explored and documented by notable economists and environmentalists such as Herman Daly.[13]

The following is my own attempt, drawing on recent research and opinion[14], to state a transparent formula for measuring productivity and defining its relationship to national competitiveness. I defer here to the definition of national competitiveness discussed above that includes both quantitative and qualitative societal goals.

The Production Function:

National Production = Inputs (labor, capital, materials, natural resources) x productivity (the rate of transforming inputs into outputs)

In $ value terms:

Net Value of national production (Gross Value of Production - Value of Pollution) = Volume of production x marginal value to the nation of transforming inputs into outputs

Where:

National Margin = Gross value of national production - the total value of inputs (including labor and natural resources) - net benefits the nation must deliver to attract domestic and foreign customers and investors in the face of international competition[15] (such as investments in infrastructure, education, resource maintenance and restoration, as well as expenditure on welfare and environmental clean-up)

Now transparent, the production function reveals important characteristics of its various components:

1. The value of national production can be increased either by increasing the volume of production; decreasing pollution, or by increasing productivity.
2. Productivity increases can result from reducing the inputs required for production, and thereby the costs of inputs. This is usually accomplished through the application of innovative technologies.
3. If production inputs, for example labor, are reduced in one industry, but cannot be employed in another industry, the reduced input costs will be offset by increased net National benefits that must be delivered to support and retrain displaced labor.
4. If the technology or process applied to achieve increased productivity, results in spoiled production inputs (for example soil erosion, water pollution) then:
   a) pollution increases, leading to a decrease in the Net Value of Production

---

[12] Some of the Bank's current research in this area, carried out jointly with the United Nations Statistical Office (UNSO), is being conducted in two developing countries -- Mexico and Papua New Guinea. The UNSO Framework being used in these studies is a system for environmentally adjusted economic accounts. It tries to integrate environmental data with existing national accounts information, while maintaining national accounting concepts and principles as far as possible. Refer, Ernst Lutz and Mohan Munasinghe, "Accounting for the Environment", Finance & Development, March 1991.

[13] Daly, Herman, "Toward a Measure of Sustainable Social Net National Product", in Environmental Accounting for Sustainable Development, ed. Yusuf J. Ahmad, Salah El Serafy, Ernst Lutz, The World Bank, Washington D.C., 1989.

See also, by the same author, and John B. Cobb Jr., For the Common Good, Beacon Press, Boston Mass, 1989.

And also by the same author, and Robert Goodland, "Ten Reasons Why Northern Income Growth is not the Solution to Southern Poverty", The World Bank, February 1992.

Ernst Lutz, Salah El Serafy, and Roefie Hueting are among the recognized authorities on this subject. The following essays appear in Environmental Accounting for Sustainable Development, op.cit.

Serafy and Lutz, "Environmental and Resource Accounting: An Overview"

Serafy, "The Proper Calculation of Income from Depletable Natural Resources"

Hueting, "Correcting National Income for Environmental Losses: Toward Practical Solution"

[14] In addition to Herman Daly, op.cit, Ernst Lutz, Salah El Serafy, and Roefie Hueting, I refer to William P. Browne, et al, Sacred Cows and Hot Potatoes: Agrarian Myths in Agricultural Policy, Westview Press, Boulder Co, 1992. The authors articulate the important distinction between production and productivity; note the societal benefits of productivity increases, and emphasize that technological innovation is not to be blamed for unemployment.

[15] I have adapted a definition used by Pankaj Ghemawat, Professor of Business Administration, Harvard Business School, to describe an organization's margin. Refer his book entitled Commitment: The Dynamic of Strategy, The Free Press, New York, 1991.

b) the spoiled inputs, cannot be used for future production.[16] Again, the reduced input costs will be offset by increased net National benefits that must be delivered to restore, if possible, the spoiled resources, or to search for substitutes.

5. If labor retraining, or resource restoration, are required, there will be a time lag before productivity increases are realized. This lag may be indefinite depending on the nature and extent of retraining and/or restoration required.

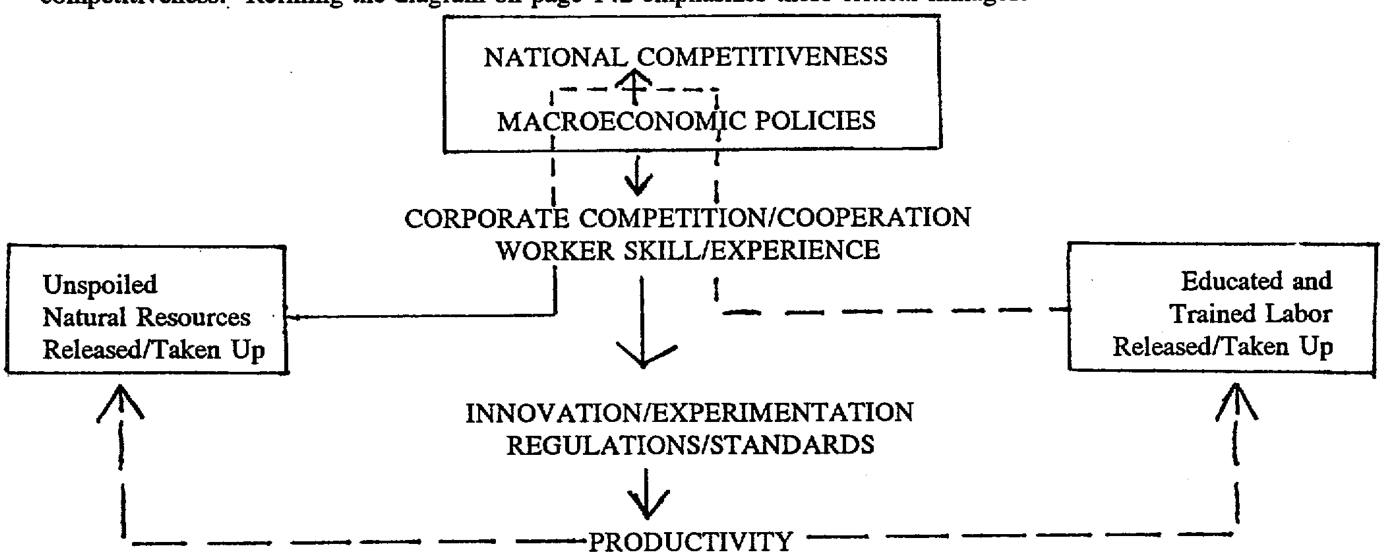
The conditions that must hold for productivity increases to be realized are therefore more complex than is apparent from the simple equation that divides GDP by population. In particular, **a productivity increase is not realized if the resources released for use elsewhere, cannot be used.**

Even if a productivity increase is realized, however, this does not automatically signal enhanced national competitiveness. While the dollar value of productivity, national margin, serves as a measure of national efficiency, it may not serve as a measure of national effectiveness. National margin does not necessarily reflect the quality of production, or whether or not the investments are worthwhile.[17] Productivity increases must increase efficiency and effectiveness, and thereby enhance quality of life, if they are to add value to national competitiveness in its richer definition. This means they must:

i) achieve worthwhile increases in the total volume of production in industries with unmet demand. An increase in the volume of production in industries where total production exceeds demand -- such as subsidized industries -- does not add value to national competitiveness. Instead, it results in overproduction which is a value-subtracting exercise, OR

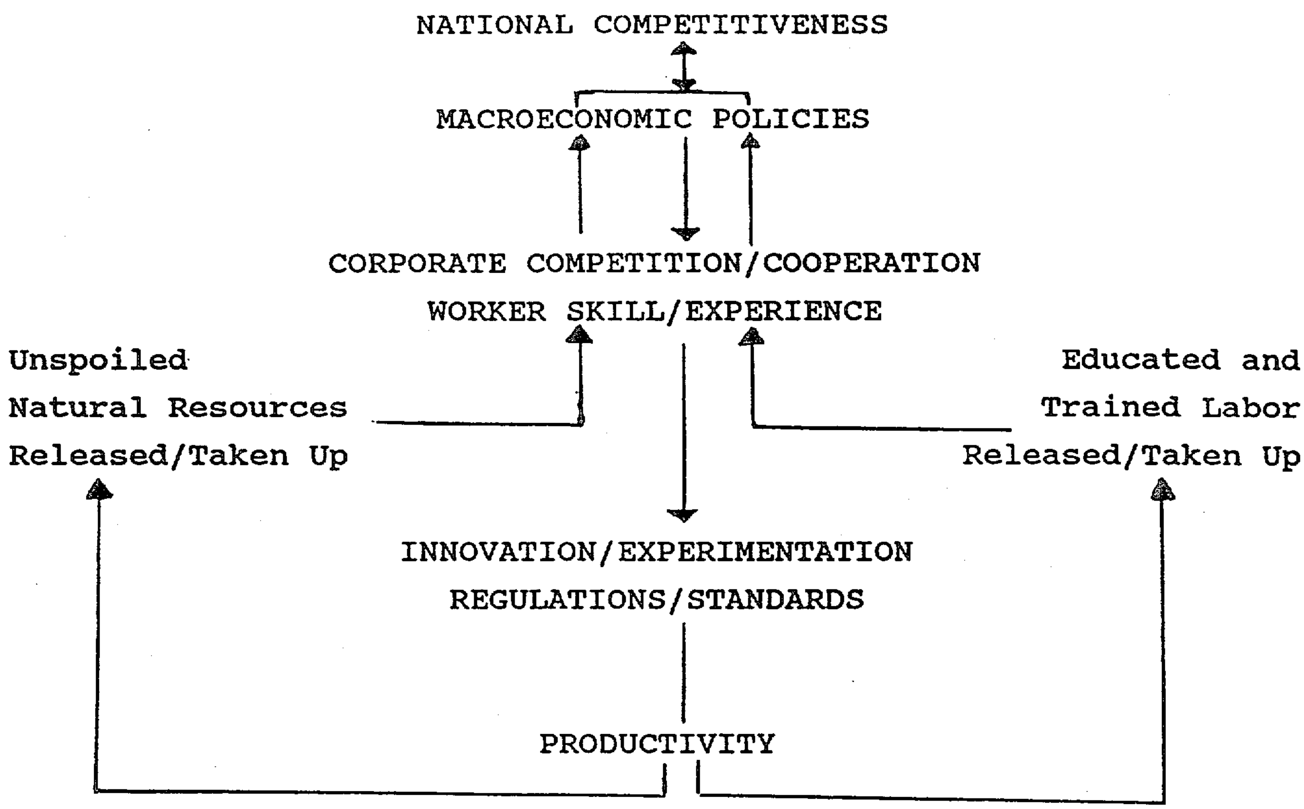
ii) release useful inputs to production (unspoiled natural resources, educated labor) from a mature or declining industry to be used in emerging, or growing industries for the production of high quality goods and services

National competitiveness, which includes the achievement of quantitative and qualitative societal goals, ultimately depends on productivity -- as now defined. It is measured by expanding national accounts to include investments in environmental and human assets that contribute to economic activities. Similarly, imputed costs of resource depletion, costs of degradation, and costs of worker displacement that are directly related to productive activities and the generation of value added, are also included. Strategic thinking and acting to prevent both pollution and unemployment spurs innovation, and enables productivity gains to be fully realized. Hence pollution prevention and unemployment prevention emerge as parallel arms of a single, sustainable strategy to advance national competitiveness. Refining the diagram on page 142 emphasizes these critical linkages:



[16] This point is made in Sacred Cows and Hot Potatoes: Agrarian Myths in Agricultural Policy, William P. Browne, et al, Westview Press, Boulder Co., 1992.

[17] I refer here to the definitions of efficiency and effectiveness adopted by Osborne and Gaebler in their book Reinventing Government, A William Patrick Book, Addison-Wesley Publishing Company, Inc., Reading, Ma., 1992.


NATIONAL COMPETITIVENESS
MACROECONOMIC POLICIES
CORPORATE COMPETITION/COOPERATION
WORKER SKILL/EXPERIENCE
Unspoiled
Natural Resources
Released/Taken Up
Educated and
Trained Labor
Released/Taken Up
INNOVATION/EXPERIMENTATION
REGULATIONS/STANDARDS
PRODUCTIVITY

Employment and Income: Analysis and Prospects

Lloyd D. Teigen, USDA Economic Research Service

Abstract

Per capita income won't rise above 1990 levels, unless high-wage jobs are created faster than population projections. The reason for the peak in 1990 was record employment per capita. Creating new jobs in the high-wage (high-productivity) sectors of the economy will increase both real average wages and real per capita income. But in the absence of new high-wage jobs, the prospects are falling real wages, fewer jobs per capita, and falling per capita income, not prospects we'd like to see realized. This paper investigates sector specific trends in the relationship between employment per capita and relative wage rates.

Summary

Per capita income won't rise above 1990 levels, unless high-wage jobs are created faster than population growth. Real disposable personal income per capita peaked early in 1990, not to be surpassed until 1993-I and then by only a quarter of 1 percent. Real wage and salary disbursements per capita peaked in 1989 and have not recovered (1992 is less than 1987). The reason for the peaks in 1989 and 1990 is record employment per capita. From 1990 to 1992, employment per capita (EPC) has dropped 3.5 percent.

Relating EPC to real wages, a negative response is apparent. At lower real wages, more people are employed per capita. For the economy at large, two regimes are evident: when manufacturing was the dominant sector, the response was inelastic; when services became dominant, the response became elastic. At the 1990 EPC, the elasticity on the 1969-92 curve is one -- denoting the maximum attainable labor income per capita under that curve. However, the relationship between total wage payments and employment doesn't show enough curvature to fully corroborate the unit elasticity finding.

Table 1: Employment per 1,000 population, 1983-92.

	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992
Agriculture	14.4	14.1	13.3	13.1	13.2	12.9	12.9	12.7	12.8	12.6
Mining	4.1	4.1	3.9	3.2	3.0	2.9	2.8	2.8	2.7	2.5
Construction	16.8	18.5	19.6	20.0	20.5	20.9	21.0	20.5	18.5	18.0
Total Manufacturing	78.7	82.0	80.8	78.8	78.4	79.0	78.6	76.5	73.0	71.2
Durable Goods Manufacturing	45.7	48.6	48.1	46.6	46.0	46.4	46.2	44.5	42.0	40.5
Nondurable Goods Manufacturing	33.0	33.4	32.7	32.3	32.4	32.5	32.4	32.0	31.1	30.7
Transportation and Public Utilities	21.1	21.8	22.0	21.8	22.1	22.6	22.8	23.2	22.8	22.5
Wholesale Trade	22.6	23.6	24.1	24.0	24.2	24.7	25.2	24.8	24.0	23.4
Retail Trade	66.6	69.9	72.7	74.4	76.0	77.9	79.0	78.7	76.2	74.9
Finance, Insurance, and Real Estate	23.3	24.1	25.0	26.1	27.0	27.1	27.1	26.9	26.4	26.1
Services	84.1	88.0	92.3	95.8	99.8	104.8	109.6	112.4	112.1	113.2
Federal Government	11.8	11.9	12.1	12.0	12.1	12.1	12.1	12.3	11.7	11.6
State and Local Government	55.9	55.9	56.7	57.3	57.9	58.8	59.8	60.9	61.0	61.1
Total Armed Forces	9.1	9.0	9.0	9.0	9.0	8.7	8.6	8.1	7.8	7.0
Total Employment	408.5	422.9	431.3	435.7	443.1	452.4	459.5	460.1	449.2	444.1

Source: [1], Table B-41, Table B-30, and B-29.

In the upper chart, higher employment per capita generates more wage and salary disbursements per capita. In the lower chart, increasing jobs per capita --beyond 1990 levels-- by reducing real wages can't be counted on to increase real (wage-based) income. But, first, employment per capita has to rise to 1990 levels. That's why the 1990 level of per capita real income may be hard to surpass.

But, one equation doesn't tell the whole story. Each sector of the economy has its own story, and they are not marching in lock step. There are sectors where EPC is rising (services, wholesale & retail trade, finance insurance & real estate (FIRE), state & local government); where EPC has no trend (transportation & public utilities, construction, federal government); and where EPC is falling (agriculture, mining, manufacturing, and armed forces). These trends affect the prospect for per capita income, since incomes vary by sector: 2 construction jobs equal 3 service jobs, 4 manufacturing jobs equal 4 wholesale or 9 retail jobs, etc.

Sectoral evidence corroborates the aggregate relationship. About half of US jobs are in sectors with elastic-to-unitary demand for labor, about a fourth are in sectors with inelastic demand, and about a fourth in sectors without an acceptable relationship. Estimation revealed a highly elastic response of service EPC to relative wage rates. Employment in wholesale trade and FIRE also had an elastic response to relative wage rates. Retail EPC had an elasticity near one. Other sectors, notably agriculture, manufacturing, and transportation & public utilities were quite inelastic to wage rates. No acceptable relationships were found for mining, construction, or government.

To increase real per capita income, jobs have to be created faster than population growth. To do this, the relationships have to shift outward. Creating new jobs in the high-wage (high-productivity) sectors of the economy will increase both real average wages and real per capita income. But in the absence of new high-wage jobs, the prospects are falling real wages, fewer jobs per capita, and falling per capita income, not prospects we'd like to see realized.

To restore per capita employment to 1990 levels, the challenge is to create 3.6 million jobs now and 100,000 jobs a month hereafter. This means creating more than 10,000 new jobs a day beyond those lost to layoffs to recoup the loss in a year. That's a tall order, especially when individual companies announce plant closings and layoffs affecting tens of thousands in a single day.

This paper's major contribution has been to identify, and to some extent quantify, a problem in the US macroeconomy's performance. The empirical research I've done does not identify all the policy instruments needed to expand employment per capita. It does suggest that falling real wages expand per-capita employment. Real wages can decline from two factors: nominal wage rates may fall, or product prices in the employing sector are allowed to rise. The rise in product prices is a form of "beneficial inflation," if the labor demand elasticity is large enough to produce a significant growth of employment. ["Beneficial" in a partial equilibrium sense that abstracts from the effects on interest rates and, subsequently, on the capital market.

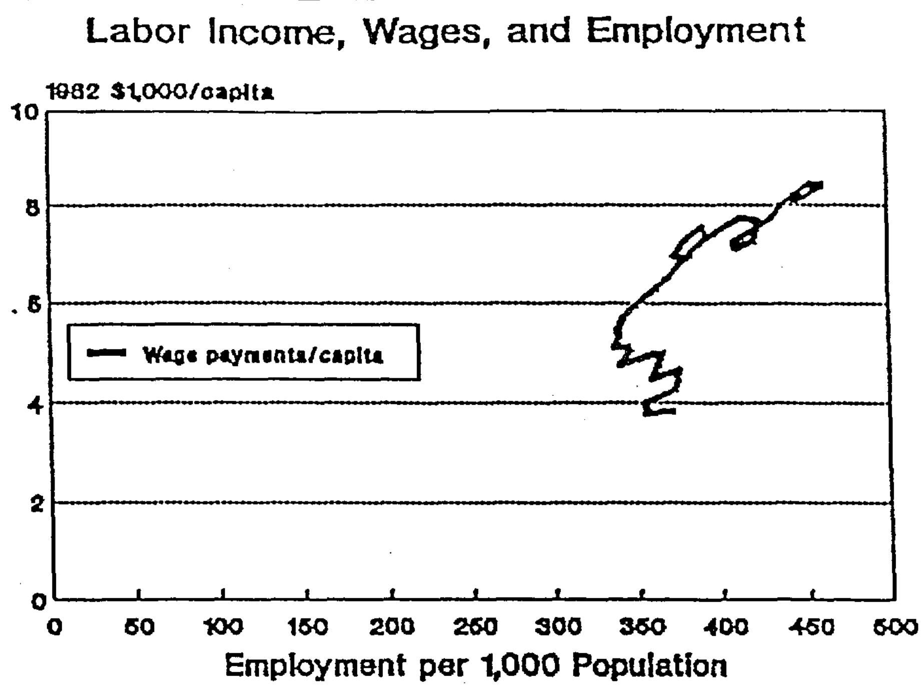
Empirical Evidence -- in Aggregate

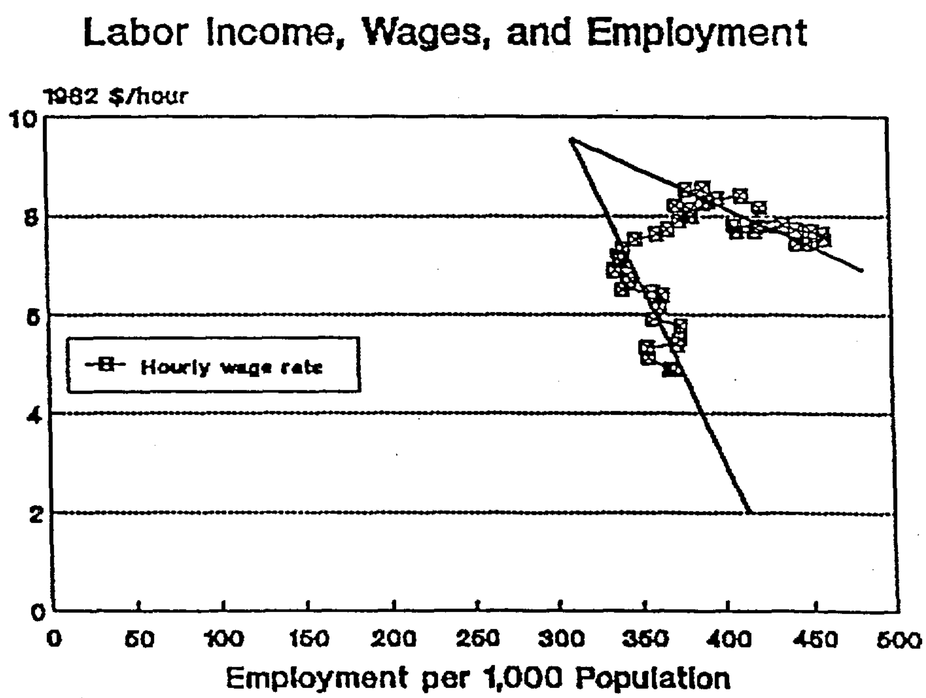
My model of labor demand expresses employment as a function of wage rates and variable input prices, each deflated by an index of output prices for the respective sector. Because of aggregation, the coefficients change with the (unobserved) number of firms [in the sector, or using a technology] per 1,000 population. The lower half of the chart on page 118 illustrates this labor demand function. Two distinct trends are evident. The response is inelastic between 1947 and 1964, and elastic between 1969 and 1992.[1] Variations from the straight lines may be attributed to random disturbances, to cross price terms, and to intra-sample changes in the aggregate coefficients linked to individual technologies.

[1] What can we infer from the increasing elasticity of labor demand between the two epochs? In the early epoch, manufacturing dominated the employment picture. In the later period, service industries dominated. The more negative the second derivatives of the production function, the more inelastic the demand for labor. Interpret this as saying that both labor and capital are highly specialized, and that adding a little more of each greatly diminishes the productivity of both. That sounds to me like a manufacturing plant with a specific product line.

As workers shifted to service jobs, the nation's productive capacity became less specialized. Added capital or labor is more like existing capital or labor, and the marginal productivity does not diminish as rapidly. Labor demand consequently becomes more elastic. As the shift to the service sector becomes complete, specialties re-emerge in the work force, and the marginal productivity of labor and capital begins to fall off more rapidly. The elasticity of labor demand begins to fall -- out of the elastic range -- into the unitary range.

By including the average work-week, the unemployment rate, and the labor force participation rate for females, reasonably stable responses could be estimated over the entire period, 1948-92. Shorter work-weeks and high female participation correlate with greater employment per capita. Higher unemployment correlate with fewer workers per capita. Female participation rates explain the subsector trends moreso than the national aggregate. The elasticities with respect to relative wages in most of these cases were less than one.


Labor Income, Wages, and Employment
1982 $1,000/capita
10
8
6
4
2
0
Wage payments/capita
0
50
100
150
200
250
300
350
400
450
500
Employment per 1,000 Population


Labor Income, Wages, and Employment
1982 $/hour
10
8
6
4
2
0
Hourly wage rate
0
50
100
150
200
250
300
350
400
450
500
Employment per 1,000 Population

Representing the period 1947 to 1964, the elasticity is -.228 at ($6.13, 356), based on:

$$\frac{\text{Employment}}{\text{1,000 Popn}} = 436.6 \underset{(-5.64)}{-13.246} \frac{\text{Wage Rate}}{CPI_{82}} \qquad R^2 = .62,\ S = 8.5$$

Representing the period 1969 to 1992, the elasticity is -1.217 at ($7.94, 415), based on:

$$\frac{\text{Employment}}{\text{1,000 Popn}} = 920.2 \underset{(-5.63)}{-63.597} \frac{\text{Wage Rate}}{CPI_{82}} \qquad R^2 = .59,\ S = 18.2$$

The maximum labor income per capita occurs where the wage elasticity of labor demand is -1. This occurs at ($7.23, 460) under the 1969-92 relationship. By comparison, the 1990 data point was ($7.52, 460). Labor income per capita in 1990 was in fact larger than the maximum labor income possible under the smooth linear relationship.

To attempt estimation of the cross-price effects, two components of the Producer Price Index (PPI) were examined. These were the PPI for intermediate materials, supplies, and components; and the PPI for crude materials for further processing. The PPI for intermediate materials had a significant influence over 1947-64. But, neither had a significant influence during 1969-92.

During 1947-64, higher priced intermediate materials were associated with greater per capita employment. Qualitatively, intermediate materials substituted for labor or vice versa.

The estimated response remained inelastic during 1947-64. The wage elasticity was -0.20 and the cross-price elasticity was +0.41, at the point ($6.13, 1.04, 355). The estimated response is:

$$\frac{\text{Employment}}{\text{1,000 Popn}} = 279.3 \underset{(-4.54)}{-11.622} \frac{\text{Wage Rate}}{CPI_{82}} \underset{(1.88)}{+141.34} \frac{\text{PPI-Intermed.}}{CPI_{82}}, \qquad \text{with } R^2 = .69 \text{ and } S = 7.9 .$$

Under this specification, labor appears to be an inferior input. Higher output prices (viz. CPI_{82}) evoke lower employment per capita. A 10 percent rise in the CPI appears to reduce employment by 2.1 percent.

Finding no significant cross-price effect for 1969-92 was discouraging. Except for lower real wages, the only way to increase employment per capita is a random (or otherwise unidentified) event. Per capita employment can increase only at the expense of lower real wages. Given where the U.S. is on that relationship, labor income per capita cannot increase much with the additional per capita employment. Unless the relationship shifts dramatically, per capita labor income is not likely to exceed the levels of 1988-90.

Subsector Demand for Labor

Subsectors of the economy are more homogeneous than the economy as a whole. Thus, subsector estimates of labor demand are expected to have greater precision than estimates for the national aggregate. To help explain the shift in the national labor demand function, from an inelastic function before 1964 to a more elastic function after 1970, we expect sectors with growing per-capita employment to be more elastic than those with declining employment. That is, labor demand in the service, finance, and trade sectors is expected to be more elastic than labor demand in agriculture, mining, and manufacturing.

Because each sector represents a different aggregation of products or services, different price deflators and different wage concepts were used in each labor demand equation*.

The estimation revealed a highly elastic response of service employment per capita to relative wage rates. Employment in wholesale trade and in the financial sector (FIRE) also had an elastic response to wage rates. Per capita retail employment had an elasticity near one. All other sectors, notably manufacturing, agriculture, and transportation & public utilities were quite inelastic to wage rates. The elasticities are summarized in the following full-page box.

*All equations mentioned in this paper were estimated using Classical Least Squares. No attempt was made to correct for serial correlation or simultaneous equation bias. The Durbin-Watson statistic in virtually all equations was less than one. This is as much symptomatic of omitted variables or incorrect functional form as it is characteristic of the purely random part of the relationship.

Summary of Sectoral (per capita) Employment Elasticities (with respect to after-tax wages relative to sectoral product prices)

Sector (sample)	Elasticity	at (wage, employment)	Wage Concept
Agriculture			
1947-73	-1.34	($3.085, 31.1)	TPNAG-hr/PRFI
	-4.04	($4.313, 14.5)	
1974-92	-0.43	($4.542, 14.1)	TPNAG-hr/PRFI
	-0.40	($4.313, 14.5)	
	-1.00	($7.520, 10.1)	
Wholesale Trade			
1964-92	-1.59	($270.9, 21.7)	WHSL-wk/CPIU
	-1.00	($220.8, 28.0)	
Retail Trade			
1964-92	-0.98	($151.7, 63.3)	RETL-wk/CPIU
	-1.00	($153.2, 62.7)	
1969-92	-0.79	($146.2, 66.3)	RETL-wk/CPIU
	-1.00	($166.0, 59.3)	
Finance, Insurance, and Real Estate			
1964-92	-1.42	($236.2, 21.4)	FIRE-wk/CPservice
	-1.00	($201.6, 25.8)	
Services			
1964-92	-2.49	($208.4, 75.7)	SERV-wk/CPservice
	-1.00	($146.1, 132.0)	
1964-92	-2.34	($6.737, 75.7)	TPNAG-hr/CPservice
	-1.00	($4.810, 126.3)	
Transportation & Public Utilities			
1964-92	-0.09	($372.3, 21.9)	TPU-wk/CPpub.trans
Manufacturing (weekly wage rate not deflated by product price)			
Total Manufacturing			
1964-92	-0.18	($211.8, 87.0)	after-tax wage
Durable goods			
1964-92	-0.18	($211.8, 50.9)	after-tax wage
Nondurable goods			
1964-92	-0.19	($211.8, 36.1)	after-tax wage

No acceptable results for Mining, Construction, and Government.

The mnemonics are translated below:

PRFI	Prices Received by Farmers Index (1977 = 1.00, 1992 = 1.39), used in agriculture sector demand.
CPIU	Consumer Price Index Urban Wage-earners (1982-84 = 1.00, 1992 = 1.402), used in wholesale and retail trade sectors.
CPservice	Consumer Price Index for all services (1982-84 = 1.00, 1992 = 1.52), used in service sector and FIRE (finance, insurance, and real estate) sector.
CPpub.trans	Consumer Price Index for public transportation (1982-84 = 1.00, 1992 = 1.507), used in the transportation and public utility sector.
TPNAG	Total Private Nonagricultural Sector average wage [-hr = per hour], [-wk = per week] (1992 = $10.59/hour, or $364.30/week).
WHSL	Wholesale Trade average wage (1992 = $435.48/week).
RETL	Retail Trade average wage (1992 = $205.63/week).
FIRE	Finance, Insurance, and Real Estate average wage (1992 = $387.36/week).
SERV	Services average wage (1992 = $342.55/week).
TPU	Transportation and Public Utilities average wage (1992 = $523.41/week).
MFG	Manufacturing average earnings (1992 = $469.45/week).

Where the elasticity equals one, wage income in that sector (real income per 1,000 population) is at its maximum. The retail sector passed through that point when employment and wages reached 1976-77 levels. Subsequent employment increases came as a result of declining real wages that reduced total payments to retail workers (per 1,000 US population).

The finance, insurance, and real estate sector passed through that point when employment exceeded 1985 levels. The 1992 employment in FIRE (at 26.1 per 1,000 population) is just slightly higher than the point of maximum per capita wage income. Real per capita income originating in FIRE is just less than its maximum.

Neither wholesale trade nor services has yet passed the point where employment demand has unit elasticity, but both are close to that point. A 10-percent fall in real wages from 1992 levels in wholesaling would take us there. A 12 to 18 percent drop in the real wages in services would take that sector to the point of unit elasticity.

Agriculture, manufacturing, and transportation-public utilities are all in a region of inelastic demand for labor (per 1,000 population). That means that whenever real wages fall, total income in those industries also decreases, despite the small increase in employment there. Or, if real wages rise, total income rises, despite the loss of jobs. No acceptable relationships were found for the mining, construction, and government sectors of the economy. These 3 sectors account for slightly less than one-fourth of US employment, 100.2 workers per 1,000 population in 1992. Average wages in these sectors are substantially above the average in the rest of the economy.

In summary, slightly more than half the jobs are in sectors with elastic-to-unitary demand, a little less than one-quarter of the jobs are in sectors with inelastic demand for labor, and about a quarter of the jobs are in sectors without an acceptable relationship. The consequence is that the aggregate relationship is unlikely to be very elastic. Thus, increasing the number of jobs per capita can no longer be counted on to increase the (wage-based) income per capita. To increase per capita income, real wages have to increase or new jobs be created in the high wage (high productivity?) sectors of the economy.

Conclusion

I wish I could have identified the policy levers or program variables needed to solve the problem, and create the requisite number of high wage jobs. However, my econometrics did not focus on policy variables. Moreover, my Department is not the major actor on labor market policies. Thus, I must content myself with identifying the problem and bringing that to the attention of folks, like you, with the influence and information needed to solve the problem.

To repeat, the problem I've identified is: Per capita income is down, because the per capita wage payments are down. Per capita wage payments are down because employment per capita is down. Per capita wage payments can rise when lower wages increase employment only if labor demand is elastic. But, the evidence is that labor demand is becoming increasingly inelastic. Thus, to raise incomes, wages, and employment the entire demand schedule has to shift outward -- and that means creating more new jobs in the high-wage (high-productivity) economy.

Sectoral Labor Demand Equations, 1964-92, except as noted.

Sector[b]	Intercept	Slope	Wage Concept[a]	R^2	S.E.
Agric.[1]	72.88	-13.547	TPNAG-hr/PRFI	.901	3.86
	(25.44)	(-15.11)			
		e =	-1.34 at ($3.085, 31.1)		
Agric.[2]	20.26	-1.347	TPNAG-hr/PRFI	.918	0.34
	(44.94)	(-13.78)			
		e =	-0.43 at ($4.542, 14.1)		
Wholesale	56.08	-0.127	WHSL-wk/CPIU	.786	1.11
	(16.18)	(-9.95)			
		e =	-1.59 at ($270.9, 21.7)		
Retail	125.34	-0.409	RETL-wk/CPIU	.906	3.20
	(32.12)	(-16.09)			
		e =	-0.98 at ($151.7, 63.3)		
Retail[3]	118.51	-0.357	RETL-wk/CPIU	.913	2.55
	(34.06)	(-15.17)			
		e =	-0.79 at ($146.2, 66.3)		
FIRE	51.60	-0.128	FIRE-wk/CPservice	.930	1.08
	(32.14)	(-18.98)			
		e =	-1.42 at ($236.2, 21.4)		
Services	264.08	-0.904	SERV-wk/CPservice	.881	7.69
	(19.74)	(-14.16)			
		e =	-2.49 at ($208.4, 75.7)		
Services	252.62	-26.262	TPNAG-hr/CPservice	.917	6.41
	(24.56)	(-17.32)			
		e =	-2.34 at ($6.737, 75.7)		
Transportation & Public Utilities	23.88	-0.005	TPU-wk/CPpub.trans	.235	0.60
	(34.06)	(-2.88)			
		e =	-0.09 at ($372.3, 21.9)		
Total Mfg	103.07	-0.076	MFG-wk[4]	.837	3.34
	(68.62)	(-11.76)			
		e =	-0.18 at ($211.8, 87.0)		
Durable Mfg	59.99	-0.043	MFG-wk[4]	.712	2.71
	(49.33)	(-8.18)			
		e =	-0.18 at ($211.8, 50.9)		
Nondur. Mfg	43.08	-0.033	MFG-wk[4]	.946	0.79
	(121.59)	(-21.79)			
		e =	-0.19 at (211.8, 36.1)		

a/ All wage rates multiplied by (1-tx-ss) to obtain after-tax wages.
b/ Dependent variable is sectoral employment per 1,000 US population.
1/ 1947-73
2/ 1974-92
3/ 1969-92
4/ Manufacturing wages were not deflated by a product price index.

Sectoral Labor Demand Equations, 1947-64, except as noted.

Sector[b/]	Intercept	Slope	Wage Concept[a/]		R^2	S.E.
Nondur. Mfr	73.54	-5.653 (-20.10)	TPNAG-hr/CPIU		.962	0.71
	elasticity =		-0.68	at ($5.27, 43.8)		
Durable Mfr	74.48	-3.751 (-2.74)	TPNAG-hr/CPIU		.320	3.46
	elasticity =		-0.36	at ($5.27, 54.7)		
Total Mfr	148.02	-9.405 (-6.00)	TPNAG-hr/CPIU		.692	3.97
	elasticity =		-0.50	at ($5.27, 98.5)		
Mining	13.50	-1.623 (-23.60)	TPNAG-hr/CPIU		.972	0.20
	elasticity =		-1.96	at ($5.27, 4.9)		
Agriculture	116.85	-15.074 (-17.22)	TPNAG-hr/CPIU		.949	2.22
	elasticity =		-2.12	at ($5.27, 37.5)		
Nonag Self-employed & Household workers[1/]	63.83	-5.899 (-6.88)	TPNAG-hr/CPIU		.518	4.08
	elasticity =		-1.22	at ($5.94, 28.8)		
Transportation & Public Utilities[2/]	39.59	-2.797 (-13.64)	TPNAG-hr/CPIU		.809	0.97
	elasticity =		-0.72	at ($5.94, 23.0)		

a/ All wage rates multiplied by (1-tx-ss) to obtain after-tax wages.
b/ Dependent variable is sectoral employment per 1,000 US population.
1/ 1947-92. Includes unpaid family workers. This represents the difference between the employment concept in table B-30 and that in B-41 of the Economic Report of the President.
2/ 1947-92.

FORECAST EVALUATION

SESSION I

MODERATOR Thomas Lienesch, Bureau of Economic Analysis

PRESENTATION 1 Vernon Dale Jones and Stuart Bretschneider, The Maxwell School, Syracuse University

PRESENTATION 2 John Kitchen, Council of Economic Advisors and Economic Research Service

SESSION II

MODERATOR Steve MacDonald, Economic Research Service

PRESENTATION 1 David Stallings and LaWanda Musgrove, World Agricultural Outlook Board

PRESENTATION 2 Robin Reichenbach, Energy Information Administration

Forecast Evaluation

Organizational Pressures on Forecast Evaluation: Managerial, Political and Procedural Influences

Vernon Dale Jones and Stuart Bretschneider, The Maxwell School, Syracuse University

Abstract

This paper proposes a theory to explain why some forecasting organizations institutionalize forecast accuracy evaluation while others do not. The theory considers internal and external aspects of managerial, political and procedural factors as they effect forecasting organizations. The theory is then tested using data from a survey of the Federal Forecasters Group. Though some support for the theory is developed, multiple alternative explanations for results and the "public" nature of the sample organizations prevent wide-scale generalization. The results suggest that larger organizations are more likely to have some form of forecast evaluation compared to smaller units. The institutionalization of forecast accuracy evaluation is closely linked to internal managerial and procedural factors, while external political pressure tends to reduce the likelihood of institutionalization of evaluation of forecast accuracy.

Keywords: Institutional, organizational, managerial, political, and procedural pressures and influences on forecasting; implementation of forecasts; users of forecasts.

Introduction

Forecasting is applied to almost all decision-making and policy-making in organizations. Managers in government and business organizations, large and small organizations, and technically sophisticated and technically unsophisticated organizations, all rely on some kind of forecasting. Over the past few years, the scope of forecasting and its importance to organizations have grown to include planning, decision-making, as well as other management issues beyond the traditional technical concerns over which forecasting methods should be used.

As the scope and importance of forecasting in organizations grow, there are a variety of environmental pressures, both external and internal to organizations, which influence the forecasting process. Users of forecasts want forecasts which will enable them to succeed in an environment which is increasingly complex, interdependent, and uncertain. To produce accurate and credible forecasts, forecasters need an appreciation of factors which influence the forecasting process as conducted in their organizations. The purpose of this paper is to develop and test a theory to explain why some public organizations decide to implement formal processes of forecast evaluation. More specifically, since forecasting tends to be implemented by technical support units within larger organizations, this research focuses on why units responsible for generating forecasts useful to other elements of the larger organization are likely to formally evaluate themselves.

The first section of the paper reviews the literature associated with influences on forecasting. The next develops a formal model along with a series of empirical hypotheses that explain how external and internal organizational factors influence the likelihood a forecasting subunit will enact a formal forecast evaluation process. The third section of the paper describes the data collection process. Data to test aspects of the model come from the Federal Forecasting Practices Study Project. The empirical findings of the study are summarized in the fourth section. Finally, the paper concludes with a summary of our findings including recommendations for future research.

Forecasting Research

The Need to Focus on Broader Forecasting Issues

Research on forecasting emphasizes the development and testing of new forecasting methodologies. Furthermore, most of the empirical work focuses on the relative accuracy of techniques, ways of combining forecasts and forecast methods to improve accuracy, methods for measuring accuracy, and techniques for achieving accuracy in different applied settings (Schultz, 1992). More attention needs to be focused on how managerial, political, and procedural factors influence the forecasting process in organizations; implementation of forecasts; uses of forecasts; and credibility of forecasts. A few forecasting researchers in recent years have taken note of this deficiency (Bretschneider and Gorr, 1989, 1991; Makridakis, 1991; and Schultz, 1992).

Influences on Forecast Accuracy

Bretschneider and Gorr (1987) argued that forecast accuracy of state government revenue forecasts was directly related to the level of political conflict/cooperation present in state government. They also argued that the organizational design of the forecasting process directly influenced the accuracy of the forecasts. Bretschneider, Gorr, Grizzle, and Klay (1989) tested and extended this theory. Their empirical results suggest that organizational, political, and forecast process variables influence accuracy as much as the forecasting methods. In particular, they found that forecast accuracy increases when independent forecasts from competing agencies are used. Furthermore, accuracy increases even more when formal procedures exist to combine competing forecasts generated by executive and legislative branch agencies. The use of outside expert advisors has a surprisingly negative effect on forecasting accuracy of state government revenues. Shkurti and Winefordner (1989) examined the relationship between forecasting and the political process in state revenue forecasting in Ohio. Their findings supported the earlier work and found that forecast accuracy is enhanced in an environment with strong institutional pressures and highly partisan politics.

Influences on Forecast Bias

Kamlet, Mowery, and Su (1987) found no evidence that partisan politics biased short-term economic forecasts by the executive branch or Congress.

On the other hand, Larkey and Smith (1984, 1989) and Feenberg, Gentry, Gilroy, and Rosen (1989) discovered a conservative bias or underestimation in state revenue forecasts explained by the need to generate a hedge or buffer against future revenue uncertainty. Cassidy, Kamlet, and Nagin (1989) found no systematic relationship between revenue forecast errors and state political and institutional factors but instead attributed forecast bias directly to the level of error (uncertainty) in economic projections used to forecast revenues. These researchers attributed a modest tendency towards conservative forecasting to economic uncertainty.

Furthermore, their research uncovered such a small amount of state revenue estimation bias that they cst doubt on previous findings of conservative bias. In a reexamination of bias in state government revenue forecasting, Bretschnieder and Gorr (1992) found that organizational design, political environment, and economic uncertainty factors all contributed to bias in forecasting state sales tax revenues. This study uncovered complex interaction effects between recessionary cycles (economic uncertainty) and the level of political conflict. The nature of bias and the impact of politics changed significantly depending upon whether a state government was at the beginning, middle, or end of the economic and political cycle.

Evaluations Within Public Organizations

A rich prescriptive literature exists on public program evaluation and evaluation research (Rossi and Freeman, 1985; and Dynes and Marvel, 1987). However, this literature is concerned with program conceptualization and design interventions, monitoring of program implementation, and assessment of program efficiency, effectiveness, and impact (Downs and Larkey, 1986). Additionally, the public strategic management and strategic planning literature addresses strategic management and planning influence on and contribution to public sector productivity (Nutt and Backoff, 1987; Bryson, 1988; and Halachmi, 1992). Little work has been done on why some organizations evaluate their activities while others do not. This study fills a gap in the literature by studying public organizations and the influences upon their decision to implement formal processes of forecast evaluation.

Research Theory and Hypotheses

Noting that little research has been conducted on organizational aspects of forecasting, Bromiley (1987) suggests areas for consideration. This study relates to three of his suggested areas by considering potential pressure factors from an organization's external and internal environment which influence the formality of forecast evaluation.

In this study, the unit of analysis is the forecasting unit of a large U.S. federal government agency. Therefore, "external" influences refer to those pressures emanating from outside the forecasting unit. This could include the large agency, Congress, the Office of the President, or even agency clients. "Internal" influences refer to those forces originating from within the forecasting unit.

Figure 1 organizes these ideas into a model, in which both internal and external pressures to evaluate forecasts may derive from managerial, political or procedural sources. Managerial pressures originate from the forecasting unit's organizational structure and role. For example, the level of resources committed to the forecasting unit, or its formal position within the agency are likely to influence both willingness and ability to carry out formal evaluation.

Political influences refer to both formal and informal efforts to manipulate the forecasting unit's activities and behaviors. For example, conscious policies regarding the level of importance placed on forecasting, the extent of usage of the forecasts by other organizations, and the degree of oversight of forecasting activities.

Procedural influences are task and technology characteristics of the forecasting unit's day-to-day activities. For example, characteristics of the forecasts being generated such as the level of aggregation in data, data quality, time horizon of forecasts (e.g. short vs. long), complexity of the forecast methods in use by the unit, and the stability of phenomena forecasted.

The dashed box in the model represents a formal forecasting group within an organization. The external and internal pressures influence the forecasting process by directly affecting the forecaster and the forecaster's development of forecasts. This occurs because the forecaster's valuation of criteria for forecast quality is influenced which in turn influences the forecasts produced. Environmental pressures alter the level of perceived importance placed on different criteria by the forecaster for evaluating the quality or usefulness of a forecast. Certain criteria can be assessed using formal evaluation while others mitigate the use of evaluation. For example, as external political pressure to manipulate forecasts increase, formal evaluation is likely to be less important since formal evaluation would tend to highlight the effects of political manipulation. On the other hand, as criteria related to technical accuracy or other scientific/professional norms become more important, then formal evaluation could provide valuable information towards enhancement of the forecasting process and the forecasts themselves.

Hypotheses

Six hypotheses have been formulated based upon our model. There is one hypothesis for each external and internal organizational factor influencing the forecasting process.

> *Hypothesis 1.* External managerial. If the forecasts are generated by the organization for its own use for external purposes, then the forecasts produced by the forecasting group will be subjected to greater scrutiny and the more formal forecast accuracy evaluation will be.

Forecasts are intended for external or internal managerial purposes by the organization producing them. Possible purposes of forecasts include financial management, management of direct service delivery of public goods, production planning, auditing the activities of others, and policy analysis and evaluation. Forecasts developed for external managerial purposes will generally be presented to more people, some outside in addition to those inside the organization. The visibility of forecasts produced for external managerial purposes is greater since the impacts go beyond the organization. Thus, the forecasting group's activities and forecasts are expected to come under greater scrutiny. Such closer examination is likely to result in greater pressure to make use of forecast evaluation in order to ensure higher quality forecasts.

> *Hypothesis 2.* Internal managerial. The greater the level of agency level resource support for forecasting, the more formal forecast accuracy evaluation will be.

An organization which places a high value on forecasting will provide more resource support for forecasting. Such support is best reflected in the relative amount of resources given to the forecasting group. The annual budget or the number of employees are two such measures of support. If an organization makes the decision to increase resources to the forecasting function at the expense of other suborganizational functions, then it is likely to require a return from its investment. Thus, the agency will produce external pressure on the forecasting unit to conduct formal evaluations of its activities to confirm that resources dedicated to forecasting are effective.

> *Hypothesis 3.* External political. If there is external organizational interest in or use of the forecasts by or for political influence, then accuracy as a criteria is likely to be less important. Consequently, an increase in external political influence is likely to reduce the likelihood that a forecasting unit will formally evaluate its forecast accuracy.

External organizations sometimes are interested in or use the forecasts generated by a forecasting group. For example Senate members or their staffs, House members or their staffs, the Office of Management and Budget, judicial organizations, the press, and industrial groups are all potential consumers of forecasts generated by different executive agencies like USDA or the Department of Labor. For forecasting groups which routinely provide forecasts to external groups, it is expected they will experience great pressure to generate accurate forecasts. On the other hand, as competing political forces come into play, there is a growing potential that forecasts will be used to support preconceived policy positions without regard to the potential accuracy of the forecast. Thus, the direction of effect for external political influence depends on whether the external actors are motivated by overall accuracy or preconceived policy positions.

> *Hypothesis 4.* Internal political. The greater the internal use of the forecasts and the greater the pressure for the forecasting group to deliver accurate forecasts, the more formal forecast accuracy evaluation will be.

An organization sometimes uses for internal purposes its own forecasts produced by the forecasting group. The organization's management or technical staffs use the forecasts for a variety of reasons. For forecasting groups which provide their forecasts to internal users and experience pressure from internal users for accurate forecasts, it is expected that forecast accuracy evaluation will be more formal.

> *Hypothesis 5.* External procedural. If the forecasting group faces competition by external forecasting sources, then the forecasts produced by the forecasting group will be subjected to greater scrutiny and the more formal forecast accuracy evaluation will be.

External procedural factors such as laws, regulations, user preferences, and competition influence the forecasting process. Competition has a potentially strong influence on the forecasting process. An organization's forecasting group which competes with other external sources of forecasts will seek to generate accurate forecasts to remain a primary source of forecasts. Therefore, the forecasting group's procedures will come under greater scrutiny in order to determine forecast accuracy and consequent competitiveness. Thus, more formal forecast accuracy evaluation will occur.

> *Hypothesis 6.* Internal procedural. The broader the scope and greater the complexity of the forecasting process, the more formal forecast accuracy evaluation will be.

Numerous methods can be used to forecast and include such categories as judgmental, time series extrapolation, single-equation econometrics regression, multi-equation/simultaneous equation econometrics regression, intentions survey, and combination. The more complex the methods used and the more frequent the more complex methods used, then the greater the procedural complexity for the forecasting group. Because of the greater procedural complexity, the organization seeks greater control over the forecasting process to ensure accuracy. Thus, it is expected that forecast accuracy evaluation will in turn be more formal.

Data Collection and Measurement

The Federal Forecasting Practices Study Project was sponsored by The International Institute of Forecasters, Syracuse University's Technology and Information Policy Program, and Carnegie Mellon University's School of Urban and Public Affairs.

Study Sample

Data were collected in 1990 from a mail survey of members of the Federal Forecasters professional organization. Members of the Federal Forecasters either have forecasting as part of their job responsibilities in the federal government or have an interest in forecasting. Membership is voluntary and the principle activity of the organization is an annual conference where ideas are exchanged and discussed through keynote speeches and presentation of papers. The sample frame consisted of 259 members listed in the organization's 1989 and 1990 directories. Thus, the sample for the study was not a random sample of all the forecasters working in the federal government. Nevertheless, the sample did represent a diversity of federal agencies by type and size and provides data for investigating managerial, political, and procedural influences on forecast evaluation. Respondents were from a broad representation of federal agencies including the Departments of Agriculture, Commerce, Defense, Energy, Health and Human Services, and Labor.

Study Survey

The overall response rate for the survey was 60%. Several types of responses were received other than completed surveys. These included individuals who returned uncompleted surveys because they no longer had job responsibilities which included forecasting or because others in their organization had already filled out a survey. Of the total number of respondents, over 75% were actively involved with developing or reviewing forecasts for their organizations. A total of 115 useable survey responses were obtained from 25 agencies for a 45% sample.

Additionally, the nature of the forecasting groups represented varies significantly. Not only were a wide variety of agencies represented in the sample, these organizations also varied significantly in size. Though the mean forecasting group employed nearly 13 full-time equivalent staff members and had a budget of over $800,000, the median group only employed five technical staff members with a budget of $300,000. It is clearly evident the sample contained a few very large forecasting groups, while most employ ten or fewer technical staff members.

Discussion of Analysis and Results

Variables

The dependent variable was the whether a forecasting unit had "no forecast accuracy evaluation process," " an informal forecast accuracy evaluation process," or "a formal forecast accuracy evaluation process."

Independent variables representing the six external and internal influences on the establishment of a formal evaluation process were developed. First, we considered managerial influences. To capture the effect of **external managerial** pressure we used a survey question designed to tap the extent to which the agency made use of forecasts for the purpose of policy analysis and evaluation. The variable is based upon combining two measures: the importance of the effect of proposed policy change and the importance of the effect of implemented policy change as purposes for forecasts. As the above hypothesis 1 suggests, a major external application of accurate forecasting is the use of forecasts for policy analysis by the agency and program evaluation. As the importance of forecasts generated by the unit for both policy analysis and policy implementation increases, the need for accurate prediction also increases. The increased need for accuracy results in more pressure on the forecasting unit to evaluate forecast accuracy. To capture the effects of **internal managerial** pressure we measured the number of full-time equivalent forecasting positions. This measure captured the impact of resource commitment by the agency to the forecasting process as well as the organizational slack resources the forecasting unit might have available to commit to additional activities. Increased resources enhance both potential and likelihood the forecasting unit can and will devote resources to forecast accuracy evaluation.

Political influences can derive from both the agency/forecasting unit or externally from Congress or other executive branch agencies. A variable measuring the extent of **external political** pressure and pressure to adjust forecasts captured the effects of external political pressure of the forecasting unit. The variable comes from combining two external use factors (use by the Senate and use by the House) to create a measure of use and then interacting the result with a measure of political pressure to adjust forecasts. While the survey question focusing on political pressure to adjust forecasts tapped the aggregate effect of external political pressure, interacting it with use of forecasts by both houses of Congress focuses the effect on external political influence. As external political pressure mounts, there is an increased potential to make use of forecasts for political ends. This is likely to reduce the likelihood accuracy as a criteria will be important. Thus, this factor will tend to decrease the likelihood a forecasting unit will have a formal forecast accuracy evaluation process. **Internal political** pressure was captured by interacting the measure of the internal use of forecasts with the measure of political pressure to adjust forecasts. Internal political pressure for using forecasts by internal staff and pressure to adjust forecasts by other agencies of similar quantities (the presence of competition), and use of single-equation econometrics regression methods are the independent variables measured. Again, as this variable is designed to pick up the extent forecasts are being manipulated, the factor is political. The extent to which the origin of the activity is internal to the agency or forecast unit, the less likely this variable taps preconceived political positions. Thus, this factor is expected to be positively related to the existence of a formal evaluation process within the forecasting unit.

Procedural influence derive from institutional and process-based rules. **External procedural** influence is represented by whether other agencies produce forecasts of the same phenomenon. If institutional/procedural conditions have led to the existence of competition, then the forecasting unit is likely to be more motivated to evaluate their own forecasts. Finally, to tap the effects of **internal procedural** influence, an independent variable should focus on the technology of forecasting in use within the forecast unit. A major distinction in forecasting methodology occurs once the unit moves to the use of correlative/causal models for generating forecasting. Thus, a binary variable from the survey that indicates that the forecasting unit is using single-equation econometrics regression methods discriminates between units making use of causal models and those that do not. We anticipate that as more sophisticated methods are used that professional values and training will combine to enhance the likelihood that formal evaluation will take place.

Table 1 presents the means and standard deviations for the measures used for the forecast accuracy evaluation independent variables. Table 2 summarizes how the combined and/or interaction variables were derived.

The multinomial logistic regression procedure assumes that the dependent variable is a multi-level categorical measure, which is then modeled as a logarithmic function of an odds ratio. There are two odds ratios in this analysis with both having the probability of "formal forecast accuracy evaluation" in the denominator. The first estimates the conditional probability of "no forecast accuracy evaluation" relative to "formal forecast accuracy evaluation" (log y1/y3) and the second estimates the conditional probability of "informal forecast accuracy evaluation" relative to "formal forecast accuracy evaluation" (log y2/y3) . The parameters of the multinomial logistic regression model are estimated based on a maximum likelihood procedure.

A negative sign for an estimate indicates that an increase in the predictor value decreases the value of the odds ratio. This implies that the numerator probability would be expected to decrease relative to the denominator probability. In other words, a negative sign indicates that an increase in the independent variable decrease the likelihood of "no forecast accuracy evaluation" relative to "formal forecast accuracy evaluation." This also implies a decrease in "informal forecast accuracy evaluation" relative to "formal forecast accuracy evaluation".

On the other hand, a positive sign for an estimate indicates that an increase in the predictor value increases the value of the odds ratio. This implies that the numerator probability would be expected to increase relative to the denominator probability. Therefore, a positive sign indicates that as the independent variable increase the likelihood of "no forecast accuracy evaluation" or "informal forecast accuracy evaluation" will also increase relative to "formal forecast accuracy evaluation."

Results

Tables 3 and 4 present the results for estimating the multinomial logistic regression equations. Table 3 displays the results for the log odds of no evaluation versus formal evaluation, while Table 4 lists the results for informal evaluation versus formal evaluation. Given that only 8 forecasting units indicated no form of evaluation, it is not surprising that only one term in the model was statistically significant. We will therefore focus our attention on the results presented in Table 4.

The logarithm of full-time equivalent staff positions is essentially a measure of the size and resource endowments of the forecasting unit. This term is statistically significant in both models and suggests that as the forecasting unit becomes larger and has greater resources it is more likely to have a formal forecast accuracy evaluation process. Clearly larger organizations have the resources and capacity to enact more specialized activities such as evaluation. This suggests capacity, not motivation.

The use of single-equation econometrics regression methods is also statistically significant in explaining the log odds for informal evaluation relative to formal evaluation. The direction of effect associated with a forecasting unit that makes use of more sophisticated techniques and causal modeling is to increase the likelihood that the unit will implement a formal forecast accuracy evaluation process. There are several possible explanations for this result. The more sophisticated methods and causal techniques are typically carried out by technically trained and professionally oriented staff. Thus, by background and training, individuals within the forecasting organization may be heavily influenced by cannons of professional behavior. For example, econometricians working on causal models have long been motivated to conduct ex-post sample evaluations of their models (McNees, 1992).

Table 1

Means and Standard Deviations for Measures Used for Forecast Accuracy Evaluation Independent Variables (n=97)

Variable	Mean	Standard Deviation
Number of full-time equivalent forecasting positions	12.70	17.50
Importance of effect of proposed policy change as purpose of forecasts	3.49[1]	1.70
Importance of effect of implemented policy change as purpose of forecasts	3.42[1]	1.67
Forecast use by Senate members or staff	2.42[2]	1.08
Forecast use by House members or staff	2.41[2]	1.09
Forecast use by agency technical staff	3.25[2]	1.14
Political pressure to adjust forecasts	1.63[3]	1.30
Use of aggregate single-equation econometrics regression methods	3.10[3]	1.56
Use of disaggregated single-equation econometrics regression methods	2.67[3]	1.76
Another agency forecasting similar quantities (competition)	0.58[4]	0.49

Means and Standard Deviations for Measures Used for Forecast Accuracy Evaluation Independent Variables (n=97)

Variable	Value	Frequency	Percentage
Level of forecast accuracy evaluation	None	8	8.5
	Informal	51	54.3
	Formal	35	37.2

1 1 = "insignificant" and 5 = "essential."
2 1 = "no use" and 5 = "heavy use."
3 1 = "very rarely" and 5 = "almost always."
4 Based on a nominal scale where 0 = "no" and 1 = "yes."

Table 2

Combined and Interaction
Forecast Accuracy Evaluation Independent Variables

Type Variable	Variable	Factors
Combined	Purpose of forecasts for policy analysis and evaluation	Importance of effect of proposed policy change as purpose of forecasts + Importance of effect of implemented policy change as purpose of forecasts
Combined and Interaction	External political pressure from internal use and pressure to adjust forecast	(Forecast use by Senate members or staff + Forecast use by House members or staff) * Political pressure to adjust forecasts
Interaction	Internal political pressure from internal use and pressure to adjust forecasts	Forecast use by agency technical staff * Political pressure to adjust forecasts
Combined	Use of single-equation econometrics (regression) methods	Use of aggregate single-equation econometrics methods + Use of disaggregated single-equation econometrics methods

Table 3.

Logistic Regression Results for the Probability of "No Forecast Accuracy Evaluation" versus "Formal Forecast Accuracy Evaluation" (n=94)

	Estimate	Standard Error	Chi-Square	Probability
Purpose of forecasts for policy analysis and evaluation	-0.24	0.20	1.40	0.24
Logarithm of full-time equivalent positions	-1.33	0.60	4.92	0.03*
External political pressure from external use and pressure to adjust forecasts	-0.46	0.32	2.09	0.15
Internal political pressure from internal use and pressure to adjust forecasts	0.005	0.25	0.00	0.98
Forecasting by other agencies (competition)	0.99	0.77	1.63	0.20
Use of single-equation econometrics models	-0.32	0.21	2.33	0.13

* Statistically significant at the 5 percent level.
** Statistically significant at the 10 percent level.

Table 4.

Logistic Regression Results for the Probability of "Informal Forecast Accuracy Evaluation" versus "Formal Forecast Accuracy Evaluation" (n=94)

	Estimate	Standard Error	Chi-Square	Probability
Purpose of forecasts for policy analysis and evaluation	-0.14	0.11	1.59	0.21
Logarithm of full-time equivalent positions	-0.98	0.29	11.18	0.0008*
External political pressure from external use and pressure to adjust forecasts	0.11	0.06	3.09	0.08**
Internal political pressure from internal use and pressure to adjust forecasts	-0.09	0.07	1.35	0.24
Forecasting by other agencies (competition)	-0.43	0.28	2.29	0.13
Use of single-equation econometrics models	-0.20	0.11	3.67	0.06**

* Statistically significant at the 5 percent level.
** Statistically significant at the 10 percent level.

One of the more interesting results presented in Table 4 are those associated with the effects of external political influence on the likelihood a forecasting unit will formally evaluate the forecast accuracy of their forecasts. In the original hypothesis, it was felt that to the extent external political influence was motivated by preconceived political positions, the effect would be to reduce the likelihood of formal evaluation. Ex-post accuracy would not be as meaningful to external political users as short-term political support for a pre-existing position. The empirical results provides some support for this view. The coefficient is positive suggesting that increasing external political influence increases the log odds that informal evaluation will occur relative to formal evaluation. The term is statistically significant but only at the 10% level.

The results in Table 4 suggest a mix of internal and external influences on the forecasting unit. Managerial and procedural factors tend to originate internally and have a positive influence while political influence is external and negative. Generally, internal factors have a greater impact on the establishment of a formal process than external factors. Size and resource endowments have the greatest impact on the likelihood that the forecasting unit will formally evaluate forecast accuracy.

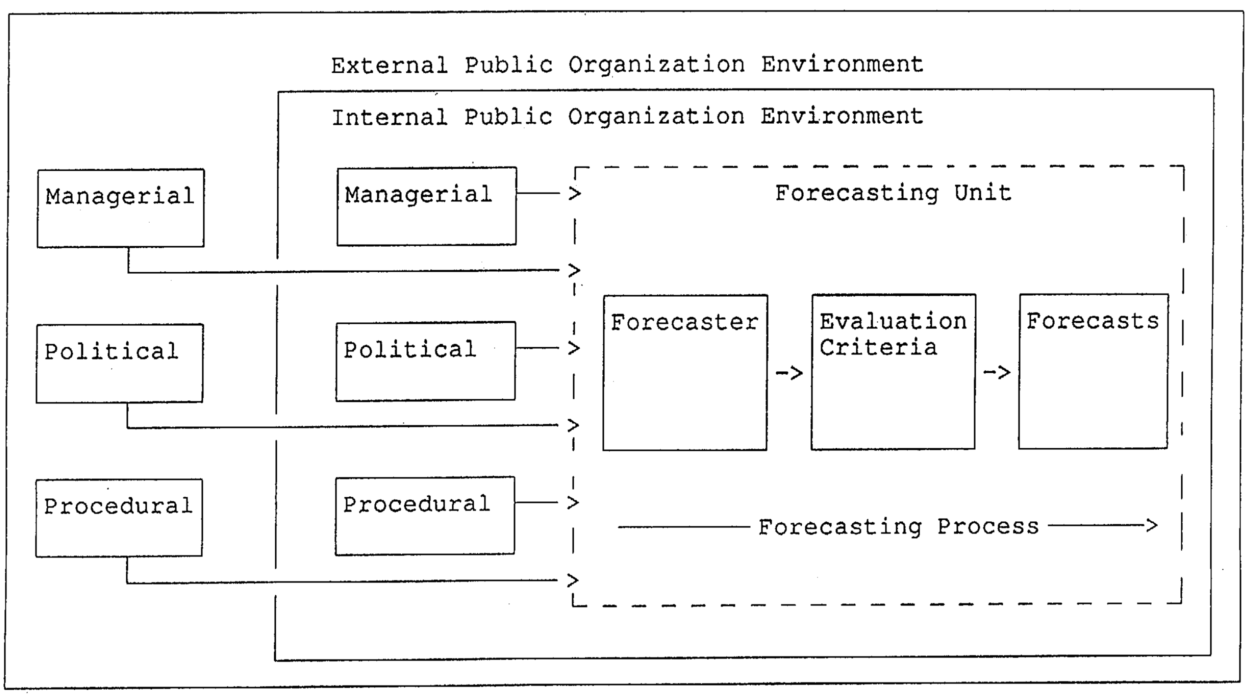
Conclusions

In considering the organizational impacts on implementation of formal process for the evaluation of forecast accuracy, several empirical results emerge. First, a mix of internal and external influences exist. The internal managerial and procedural influences tend to have a dominant and positive effect. "Good" management and process includes some form of evaluation. It is useful to note that over 90% of the sample had some form of accuracy evaluation in place. Larger organizations and those that make use of more sophisticated techniques of forecasting are more likely to establish formal forecasting accuracy evaluation processes than smaller units using less sophisticated techniques. This is not a surprise. Though part of the explanation is simple capacity to do formal evaluation, a part of the explanation is also related to the professionalization of the forecasting unit.

More sophisticated techniques can only be implemented by individuals with advanced training and high levels of professional training. As individuals obtain such training, usually in professional master's degree programs and doctoral degree programs, they become more heavily influenced by the norms of behavior associated with an academic discipline or professional group. In the context of forecasting, particularly with causal models, such norm invariably include the use of forecast accuracy evaluation.

More surprising, though, are results related to external influences on the use of formal forecast accuracy evaluation. Here we find that the principle external influence on the establishment of a formal evaluation is the extent of use by Congressional staff interacted with the level of pressure to adjust or change forecasts. This type of political influence suggests that the forecasts are being used within the context of the political process. Such forecasts are more likely to have short-term political value in the context of specific debates and have little long-term value. Thus, accuracy of a forecast, which can only be determined in a longer run context, has little value. Higher levels of external political influence reduce the likelihood a forecasting unit will formally evaluated accuracy. A similar conclusion was reached by Kraemer et. al. (1987), while studying the impact of various policy models used in the federal government.

Figure 1
External and Internal Organizational Factors Influencing Forecasting



References

Bretschneider, Stuart I. and Gorr, Wilpen L. "Economic, Organizational, and Political Influences on Biases in Forecasting State Sales Tax Receipts." International Journal of Forecasting, 7 (1992): 457-466.

Bretschneider, Stuart I. and Gorr, Wilpen L. "Forecasting Practices in the Federal Government." Federal Forecasters Conference 1991: Papers and Proceedings, (1991): 17-24.

Bretschneider, Stuart I. and Gorr, Wilpen L. "Editorial: Forecasting as a Science." International Journal of Forecasting, 5 (1989): 305-306.

Bretschneider, Stuart I. and Gorr, Wilpen L. "State and Local Government Revenue Forecasting." In The Handbook of Forecasting: A Manager's Guide, 2nd ed., 118-134. Edited by Spyros Makridakis and Steven C. Wheelwright. New York: John Wiley and Sons, 1987.

Bretschneider, Stuart I.; Gorr, Wilpen L.; Grizzle, Gloria; and Klay, Earle. "Political and Organizational Influences on the Accuracy of Forecasting State Government Revenues." International Journal of Forecasting, 5 (1989): 307-319.

Bromiley, Philip. "Do Forecasts Produced By Organizations Reflect Anchoring and Adjustment?" Journal of Forecasting, 6 (1987): 201-210.

Bryson, John M. Strategic Planning for Public and Nonprofit Organizations. San Francisco: Jossey-Bass Publishers, 1988.

Cassidy, Glenn; Kamlet, Mark S.; and Nagin, Daniel S. "An Empirical Examination of Bias in Revenue Forecasts by State Government." International Journal of Forecasting, 5 (3) (1989): 321-333.

Downs, George W. and Larkey, Patrick D. The Search for Government Efficiency: From Hubris to Helplessness. Philadelphia: Temple University Press, 1986.

Dynes, Patrick S. and Marvel, Mary K. Program Evaluation: An Annotated Bibliography. New York: Garland Publishing, Inc., 1987.

Feenberg, Daniel R.; Gentry, William; Gilroy, David; and Rosen, Harvey S. "Testing the Rationality of State Revenue Forecasts." Review of Economics and Statistics, 71 (2) (1989): 300-308.

Halachmi, Arie. "Strategic Management and Productivity." In Public Productivity Handbook, 551-563. Edited by Marc Holzer. New York: Marcel Dekker, Inc., 1992.

Kamlet, Mark S.; Mowery, David C.; and Su, Tsai-Tsu. "Whom Do You Trust? An Analysis of Executive and Congressional Economic Forecasts." Journal of Policy Analysis and Management, 6 (3) (1987): 365-384.

Kraemer, Kenneth L., Dickhoven, Siegfried, Tierney, Susan F., and King, John L., Datawars: The Politics of Modeling in Federal Policy Making, New York: Columbia University Press, 1987.

Larkey, Patrick D. and Smith, Richard A. "Bias in the Formulation of Local Government Budget Problems." Policy Sciences, 22 (1989): 123-166.

Larkey, Patrick D. and Smith, Richard A. "The Misrepresentation of Information in Government Budgeting." In Advances in Information Processing in Organizations, Vol. 1, 63-92. Edited by L. S. Sproull and P. D. Larkey. Greenwich, CT: JAI Press, 1984.

Makridakis, Spyros. "Editorial: Forecasting in the 21st Century." International Journal of Forecasting, 7 (1991): 123-126.

McNees, Stephen, How Large are Economic Forecast Errors?, New England Economic Review, (July/August 1992): 1-18.

Nutt, P. C. and Backoff, R. W. "A Strategic Management Process for Public and Third-Sector Organizations." Journal of the American Planning Association, 53 (1987): 44-57.

Rossi, Peter H. and Freeman, Howard E. Evaluation: A Systematic Approach. 3rd ed. Beverly Hills, CA: Sage Publications, 1985.

Schultz, Randall L. "Editorial: Fundamental Aspects of Forecasting in Organizations." International Journal of Forecasting, 7 (1992): 409-411.

Shkurti, William J. and Winefordner, Darrell. "The Politics of State Revenue Forecasting in Ohio, 1984-1987: A Case Study and Research Implications." International Journal of Forecasting, 5 (1989): 361-371.

Relative Information in Public and Private Sector Macroeconomic Forecasts

John Kitchen, Council of Economic Advisers and Economic Research Service

Introduction

Public sector macroeconomic forecasts--including those of the Administration and the Congressional Budget Office (CBO)--play an important information role for policy makers and the general public. Macroeconomic forecasts are a specific mechanism, and perhaps one of the more obvious ones, through which the Federal government plays the role of an anticipatory government. The forecasts provide information on the outlook for the economy as well as serving as a basis for calculating budget estimates--revenues, outlays, and deficits. Hence, forecasts of revenues, outlays, and deficits can only be as good as the underlying macroeconomic forecasts.

Macroeconomic forecasts can be used to evaluate how well the Federal government has functioned as an anticipatory government--at least in application to economic outlook and budget issues. One way to proceed in such an evaluation is to examine relative forecast accuracy. Alternative public sector forecasts--Administration and CBO forecasts--can be compared. Similarly, comparisons can be made between public sector and private sector macroeconomic forecasts. In fact, it is typically the case that the relative optimism or pessimism of government forecasts is evaluated in comparison to private sector "consensus" forecasts. In the early 1980s, Administration forecasts were criticized as being overly optimistic (through comparison with private forecasts) and the label of "rosy scenario" was born. For example, the 1981 Administration real GNP growth forecast had a cumulative 6-year forecast error that amounted to about 4 percent of GNP (although the forecast error in 1982 due to the recession more than accounted for the total). This issue helps to illustrate that proper functioning of the anticipatory role of government is necessary to maintain policy credibility.

Key questions emerge. Which of the forecasts was "better" at predicting the performance of the economy? Did one forecast carry more information than other forecasts? Did the consensus private forecast outperform public sector forecasts? The research reported in this paper is an attempt to answer such questions.

One might interpret the last question concerning the relative accuracy of public versus private forecasts as an attempt to determine whether the Federal government would have been better off over time to have used the private sector consensus forecast as a basis for its budget estimates, rather than having produced its own macroeconomic forecasts. However, as will be discussed later in the paper, such a view would be based on a misunderstanding of the manner in which Administration forecasts are produced. Administration forecasts are "policy" forecasts that incorporate the effects of the Administration's policy proposals. Hence, at times there may be a valid reason--a policy proposal that was not adopted--for a positive bias of Administration real growth forecasts. A recent example (not included in the sample to be analyzed in this paper) is the Clinton Administration forecast for real GDP growth from February 1993. That forecast was a policy forecast that included estimated short-run beneficial effects that would result from passage of a proposed stimulus package. When the stimulus package was not included as part of the final budget proposal, the Administration's expectation for real growth for 1993 and 1994 declined. In contrast, CBO forecasts are based on a current services baseline and do not include prospective policy changes, while private forecasts presumably incorporate prospective policy changes according to the (subjective) probability that they will pass into law.

Data and Methodology

Data for forecasts of key macroeconomic variables were collected for the 1980 to 1992 period for Administration, CBO, and the Blue Chip consensus forecasts. Data were available for 1-year-ahead forecasts and 2-year-ahead forecasts. This means there are relatively small sample sizes: 13 observations for the 1-year-ahead forecasts and 12 observations for the 2-year-ahead forecasts. For the Administration forecast, the source was the Budget of the U.S. Government; for the CBO, The Economic and Budget Outlook; and, for the Blue Chip consensus, Blue Chip Economic Indicators. While the Budget usually was not released until early February, Administration forecasts typically were made in December with adjustment possible into early January. A similar schedule typically applied to CBO forecasts. Comparable Blue Chip consensus forecasts were taken from January issues of Blue Chip Economic Indicators. Data for the actual values of the forecasted variables were the first-reported final estimates. For example, the first-reported final estimates of real GNP growth and GNP deflator change for a given year were available at the end of March of the following year when the final estimate of GNP for the fourth quarter was released. There are advantages and disadvantages to this approach. Among the advantages is the fact that the approach reduces problems that might arise from methodological or base-year inconsistencies between the forecasts and actual values. A disadvantage is that the first final estimates do not always incorporate all relevant information, such as occurs when annual revisions are made to the NIPAs or when the accounts are rebenchmarked.[1] Given the alternatives, I chose to use the first-reported final values because they would yield the most legitimate comparison for the information set available when the forecasts were made. Of course, that choice also meant more work at assembling the data set.

Due to differences in published forecast presentations, only a subset of macroeconomic variables can be studied: real GNP growth, GNP deflator inflation, and the unemployment rate.[2] CPI inflation could not be used in the comparisons because, prior to 1992, Administration forecasts were CPI-W forecasts while CBO and Blue Chip forecasts were for the CPI-U (a simple mechanical adjustment between the two measures does not exist). Similarly, there was not a consistent set of interest rate data available for either the short-term or long-term interest rates for the sample period used--Administration and CBO forecasts reported Treasury security rates (and somewhat inconsistently for the CBO) while the Blue Chip survey results typically reported private rates such as commercial paper and corporate bond rates.

The empirical analysis used to compare the alternative forecasts is based on a procedure described by Fair and Shiller (1990) for comparing information in forecasts from alternative econometric models. To examine whether one model's forecast carried different information from an alternative model, Fair and Shiller regressed the actual change in GNP on the forecasted change from the two models. For example,

$$Y_t = a + b\,Y_{1t} + c\,Y_{2t} + e_t$$

where Y_t is the actual value of the variable in period t, Y_{1t} is the forecast of the variable in period t from model 1, and Y_{2t} is the forecast of the variable in period t from model 2. If neither model contained useful information for forecasting Y_t, then both coefficients b and c should be zero. If both models carried independent information, then both b and c will be nonzero. If one model carried all of the information of the other model plus additional information, then the coefficient on the model with more information would be nonzero and the other slope coefficient would be zero.

The application of the Fair and Shiller procedure for analyzing information carried by alternative public and private sector forecasts is straightforward. Regressions of the forecasted variable on the alternative public and private sector forecasts provide information on the relative information carried by the forecasts.

Results

Tables 1 and 2 present results testing for bias in the alternative forecasts as well as root-mean-square error (RMSE) statistics. To examine bias for the alternative forecasts, in each case the forecast error was regressed on a constant. The first column in each table shows the coefficient estimate for the constant term. The standard error and Durbin-Watson statistics from that regression are reported in the second and third columns. The RMSE is reported in the fourth column. If a forecast was unbiased, then the constant coefficient would be zero and the standard error of the regression would equal the RMSE.

For one-year-ahead real GNP growth forecasts, Administration forecasts exhibited the least bias but the greatest variation as revealed by the higher standard error of the regression and the higher RMSE; CBO forecasts had the smallest standard error and RMSE. For the unemployment rate, similar results were observed across the forecasts, with the CBO having a slightly smaller standard error and RMSE. For the GNP deflator change, all of

[1]Comparisons of results for the alternative actual values may be interesting. Estimations using the most-recent reported values have not yet been performed.

[2]Prior to 1992, Administration forecasts reported the total unemployment rate rather than the civilian unemployment rate--the difference attributable to military personnel. The forecasts for the total unemployment rate were raised by 0.1 percentage point to attain the equivalent civilian unemployment rate.

TABLE 1--BIAS AND RMSE FOR 1-YEAR-AHEAD FORECASTS

	CONSTANT	SE	DW	RMSE
GNP GROWTH				
ADMINISTRATION	0.008 (0.022)	1.241	2.31	1.241
BLUE CHIP	-0.200 (0.652)	1.105	2.30	1.125
CBO	-0.100 (0.347)	1.038	2.18	1.043
UNEMPLOYMENT RATE				
ADMINISTRATION	0.092 (0.665)	0.501	2.39	0.510
BLUE CHIP	0.092 (0.656)	0.507	2.06	0.516
CBO	0.092 (0.698)	0.477	2.28	0.486
GNP DEFLATOR INFLATION				
ADMINISTRATION	0.523 (2.800)	0.673	1.02	0.866
BLUE CHIP	0.600 (4.544)	0.476	1.40	0.785
CBO	0.462 (3.197)	0.520	1.69	0.708

NOTES: The first three columns show results from regressions of the forecast on a constant.
t-statistic in absolute value in parentheses.
SE is standard error of the regression.
DW is the Durbin-Watson statistic.
RMSE is the root mean square error for the forecast.

TABLE 2--BIAS AND RMSE FOR 2-YEAR-AHEAD FORECASTS

	CONSTANT	SE	DW	RMSE
GNP GROWTH				
ADMINISTRATION	1.242 (2.047)	2.101	2.19	2.469
BLUE CHIP	0.733 (1.345)	1.889	2.07	2.039
CBO	0.867 (1.685)	1.782	2.08	1.994
UNEMPLOYMENT RATE				
ADMINISTRATION	-0.142 (0.390)	1.259	1.64	1.267
BLUE CHIP	-0.092 (0.241)	1.317	1.25	1.321
CBO	-0.050 (0.146)	1.190	1.78	1.191
GNP DEFLATOR INFLATION				
ADMINISTRATION	1.008 (3.458)	1.051	1.44	1.508
BLUE CHIP	1.223 (3.304)	1.335	1.26	1.892
CBO	1.177 (4.003)	1.060	1.19	1.659

Notes: See Table 1.

the forecasts exhibited positive bias. Evidently, the extent of disinflation of the 1980s generally was unanticipated by public and private forecasters. The CBO again had the lowest RMSE, while the Administration had the highest.

For the two-year-ahead real GNP growth forecasts, each of the forecasts show evidence of positive bias, with the Administration forecast having the largest and most significant bias of about 1 1/4 percentage points. Administration real growth forecasts also had the highest RMSE. Blue Chip and CBO forecasts showed similar positive biases and RMSEs. For the unemployment rate, each of the forecasts had a small negative bias--consistent with the overestimate of real growth, although a typical Okun's law relationship would suggest absolutely larger negative biases given the size of the positive biases for real GNP growth.[3] CBO forecasts for the unemployment rate had the smallest RMSE. For the GNP deflator change, the forecasts had positive and significant biases in the range of 1 to 1 1/4 percentage points. Administration forecasts had the smallest bias and RMSE for the deflator change.[4]

The bias and RMSE estimates generally would indicate the CBO forecasts were the "best" of the three forecasts considered. However, the regression procedure for comparing alternative forecasts described above has some advantages over a simple comparison of bias and RMSE:

> First, if the RMSEs are close for two forecasts, little can be concluded about the relative merits of the two. Second, even if one RMSE is much smaller than the other, it may still be that the forecast with the higher RMSE contains information not in the other forecast. There is no way to test for this using the RMSE framework. (Fair and Shiller, p. 376)

Tables 3, 4 and 5 show forecast comparisons using the regression procedure described above to test for the relative information carried by alternative forecasts.[5] Table 3 shows the comparisons for real GNP growth forecasts. For the one-year-ahead forecast, the regression of real GNP growth on the Administration forecast and the Blue Chip forecast shows a significant coefficient[6] on the Blue Chip forecast and a coefficient on the Administration forecast that is not significantly different from zero. Hence, these results indicate that the Blue Chip forecast carried all of the information contained in the Administration forecast plus additional information. A similar result is observed for the comparison of Administration forecasts with CBO forecasts, with the CBO forecast carrying more information. The comparisons between CBO and Blue Chip forecasts indicate, at most, a marginal information advantage for CBO forecasts. The results for 2-year-ahead forecasts are similar to those for the 1-year-ahead forecasts, suggesting that the Blue Chip and CBO forecasts carried more information than Administration forecasts. However, one difference is that the coefficient on the Administration forecast is negative (significant at the 7 percent level) in the comparison with the CBO forecast. This result indicates that by including information from CBO forecasts, Administration forecast carried "counter-information," in the sense that the sign of the effect of the additional information provided by the Administration forecast varied inversely with the forecasted value. Inspection of the forecast errors showed that the Administration forecast error tended to have the same sign as the CBO forecast error, but the Administration forecast error was larger in absolute magnitude. That result is confirmed by the fact that the Administration forecast had the largest RMSE in Table 2 for 2-year-ahead forecasts.[7]

---

[3] An alternative explanation is that there was an overestimate of the rate of growth of potential GNP.

[4] The positive biases on real GNP growth and the GNP deflator change go a long way to explaining why we have had the Federal budget deficit problems we have had in recent years and over the past decade. Across the various forecasts, the total effect on nominal GNP growth for the 2-year-ahead forecasts is in the 2 to 2 1/4 percentage point range. As an example of the budget effect, current budget sensitivities indicate that a one-time lower nominal GNP growth shock of that magnitude would result in higher cumulative effect on the deficit over 6 years of around $130 billion, with the deficit in the sixth year alone being over $30 billion higher.

[5] Estimation was performed using ordinary least squares with the variance-covariance matrix adjusted for heteroskedasticity of the errors. The two-year-ahead forecasts may introduce a moving average process to the error term in the regression estimates. Only the errors for the regressions for the two-year-ahead GNP deflator change showed evidence of an MA(1) process. Estimation with correction for the moving average process in that case have not yet been performed.

[6] Significance determined at least the 5 percent level.

[7] In the cases that follow, negative coefficients of relatively high significance also appear to result from relatively larger (in absolute magnitude) forecast errors.

TABLE 3--FORECAST COMPARISONS: REAL GNP GROWTH

	CONSTANT	ADMIN	BLUE CHIP	CBO	SE	DW
1-YEAR-AHEAD FORECASTS						
ADMIN - BLUE CHIP	-0.071	-0.141	1.299		1.175	2.06
	(0.121)	(0.288)	(2.657)			
ADMIN - CBO	-0.126	-0.847		2.007	1.031	1.82
	(0.260)	(1.303)		(2.939)		
BLUE CHIP - CBO	-0.204		-0.182	1.326	1.101	1.98
	(0.417)		(0.165)	(1.193)		
2-YEAR-AHEAD FORECASTS						
ADMIN - BLUE CHIP	-1.803	-0.095	1.466		2.057	2.10
	(1.182)	(0.169)	(1.302)			
ADMIN - CBO	-1.078	-1.214		2.426	1.763	1.75
	(0.901)	(2.080)		(3.054)		
BLUE CHIP - CBO	-2.515		0.037	1.488	1.907	2.10
	(1.277)		(0.021)	(0.941)		

Notes: The results in this table are from regressions of the form $Y_t = a + b Y_{1t} + c Y_{2t} + e_t$ as described in the text.
t-statistics in absolute value in parentheses

TABLE 4--FORECAST COMPARISONS: UNEMPLOYMENT RATE

	CONSTANT	ADMIN	BLUE CHIP	CBO	SE	DW
1-YEAR-AHEAD FORECASTS						
ADMIN - BLUE CHIP	0.621 (0.742)	0.797 (1.440)	0.104 (0.198)		0.519	2.23
ADMIN - CBO	0.438 (0.326)	0.338 (0.225)		0.588 (0.360)	0.515	2.20
BLUE CHIP - CBO	0.250 (0.266)		0.099 (0.174)	0.853 (1.324)	0.516	2.17
2-YEAR-AHEAD FORECASTS						
ADMIN - BLUE CHIP	7.250 (2.393)	2.540 (2.915)	-2.543 (2.265)		1.104	1.68
ADMIN - CBO	1.952 (0.819)	-0.279 (0.270)		1.006 (0.912)	1.254	1.51
BLUE CHIP - CBO	5.143 (2.021)		-2.037 (2.060)	2.302 (2.786)	1.071	2.11

Notes: See table 3.

TABLE 5--FORECAST COMPARISONS: GNP DEFLATOR INFLATION

	CONSTANT	ADMIN	BLUE CHIP	CBO	SE	DW
1-YEAR-AHEAD FORECASTS						
ADMIN - BLUE CHIP	-0.519	-0.033	1.017		0.519	1.44
	(1.177)	(0.110)	(2.847)			
ADMIN - CBO	-0.074	-0.220		1.145	0.525	1.94
	(0.223)	(0.792)		(3.911)		
BLUE CHIP - CBO	-0.322		0.543	0.414	0.498	1.40
	(0.645)		(1.197)	(1.154)		
2-YEAR-AHEAD FORECASTS						
ADMIN - BLUE CHIP	0.278	1.708	-0.925		0.947	1.74
	(0.291)	(2.794)	(1.897)			
ADMIN - CBO	0.076	0.055		0.705	1.016	1.35
	(0.076)	(0.066)		(0.872)		
BLUE CHIP - CBO	0.770		-1.231	1.867	0.774	1.98
	(1.536)		(2.764)	(4.114)		

Notes: See table 3.

For the unemployment rate forecast comparisons, the regression results for the 1-year-ahead forecasts indicate that none of the forecasts carried distinct information relative to the other forecasts. For the 2-year-ahead forecasts, no information advantage existed between Administration and CBO forecasts. For comparisons with Blue Chip forecasts, both the Administration and CBO forecasts carried information while the negative signs (of relatively high significance) on the Blue Chip forecast indicates a similar relative information role as was described above for the Administration 2-year-ahead GNP growth forecast (note the higher RMSE for the Blue Chip forecasts in Table 2).

For the GNP deflator percent change forecast comparisons, Blue Chip and CBO forecasts each carried more information relative to Administration forecasts for the 1-year-ahead forecasts. There is very marginal evidence that CBO and Blue Chip forecasts each carried independent information. For the 2-year-ahead forecasts, Administration and CBO forecasts each carried more relevant information than the Blue Chip forecasts, while Administration and CBO forecasts did not carry independent information. The coefficients on the Blue Chip forecasts again had negative coefficients of relatively high significance (again note the higher RMSE for the Blue Chip forecasts in Table 2).

Alternative Explanations

Every effort was made to assure consistency across forecasts. Even so, additional explanations for differences in forecast performance exist, including (1) differences in timing of completion of forecasts, (2) differences in information sets, and (3) differences in policy assumptions on which the forecasts are based. Although the timing and information differences can be important, the difference in policy assumptions is likely the single most important factor for the performance of Administration forecasts relative to alternative forecasts.

The differences in policy assumptions could be considered a difference in information sets with forecasts conditional on the given information set. However, in the case of the Administration forecast, the information set used for the forecast was constrained to include proposed policy changes--whether they were likely to be adopted or not. Similarly, the CBO was constrained to use a current services approach in which policy proposals were not included. Hence, those information sets were not necessarily the information sets that those responsible for preparing Administration and CBO forecasts considered to be most likely.

The incorporation of policy assumptions in forecasts can be described in notation as:

$$F_t = B_t + \pi_t P_t$$

where F_t is the forecast of the variable of interest, B_t is a current services baseline forecast, π_t is the (subjective) probability of adoption of the Administration's proposed policy changes, and P_t is the effect on the forecasted variable attributable to the proposed policy changes. The subjective probability π_t takes on the following values across the different forecasts:

Administration:	$\pi_t = 1$
Blue Chip:	$0 \leq \pi_t \leq 1$
CBO:	$\pi_t = 0$

Let's consider again the results described above. The results indicate that the CBO generally tended (although there were some exceptions) to have forecasts that had a lower bias, a lower RMSE, and that were the most informative (based on the regression approach). One might interpret that result simply as a superior forecasting performance on the part of the CBO. An alternative explanation is that lack of action on the part of Congress, specifically, and the Federal government, generally, to adopt Administration policy proposals made status quo forecasts more reliable. For example, Administration forecasts had the largest bias (positive) and RMSE for real GNP growth in the 2-year-ahead forecasts. Those results should not be surprising *given the incorporation of Administration policy proposals with a probability of one* along with the fact that, over the years, *many of the Administration's proposals were not adopted.* Not only would the forecast be upwardly biased, but also the variation of the forecast would be higher given the inclusion of the additional forecasted component. Note that it is impossible to determine the relative role of the subjective probabilities for the Blue Chip forecasts, because the subjective probabilities are unknown.

Conclusions

The results and analysis presented in this paper indicate that there is no clear-cut answer to the question of which of the alternative public or private sector forecasts had superior performance over the 1980-92 period. A direct interpretation of the results would indicate that, *ex post*, the CBO forecasts of that period were more accurate and carried more information relative to the Administration and Blue Chip forecasts. However, such interpretations are clouded by the fact that Administration forecasts were "policy" forecasts that directly and completely incorporated the forecasted effects of proposed policy changes. The incorporation of the additional forecasted policy component in Administration forecasts (with probability 1) and Blue Chip forecasts (with varying probabilities) distort the statistical measures used for comparison of the forecasts. The fact that Blue Chip consensus forecasts did not outperform the CBO forecasts is notable, suggesting either that CBO forecasts were inherently more accurate and informative, or that the Blue Chip estimates of the probability of policy adoption were inaccurate.

References

Fair, Ray C. and Robert J. Shiller, "Comparing Information in Forecasts from Econometric Models," American Economic Review, June 1990, pp. 375-389.

Budget of the U.S. Government, Office of Management and Budget, 1980-92.

The Economic and Budget Outlook, Congressional Budget Office, 1980-92.

Blue Chip Economic Indicators, Capital Publications, January issues, 1980-92.

Evaluation of Long-Term Forecasts for Selected U.S. Agricultural Commodities

David Stallings and LaWanda Musgrove, World Agricultural Outlook Board, U.S. Department of Agriculture

The U.S. Department of Agriculture (USDA) issues estimates of total expenditures for the following fiscal year in January of each year, as part of the Government-wide budget estimates prepared by the Office of Management and Budget (OMB). The USDA has a wide variety of programs, ranging in responsibility from food stamps to farm programs to forestry. The most volatile spending component of the USDA budget in recent years has been outlays for farm commodity programs. Net outlays in this area over the past 10 years have varied from almost $26 billion in fiscal 1986 down to $6.5 billion in fiscal 1990.[1] Thus, an accurate picture of next year's Department expenditures depends largely on the accuracy of the forecasts of those variables that influence commodity program outlays.

There are three sets of variables that directly influence the amount spent on commodity programs. First are the program parameters themselves: the array of support prices, land setaside, and export programs that may require government checks to be written. Second is the supply and use of U.S. program commodities. The larger the difference between supply and use, other things equal, the greater the level of expenditures required to provide income support. Third is the state of the national economy. High real interest rates, for example, discourage private stockholding. Slow income growth often means a slow change in demand for livestock products, and hence the feedstuffs that are part of that process.

This report focuses on forecasts of the supply and use of major program commodities for the five-year budget "baseline" period. Program parameters are, for the most part, known well in advance of planting and harvesting for the first year. These can therefore almost always be incorporated into the one year out forecast. However, fiscal conditions changed markedly during the period covered by this study, 1970-92, resulting in sharp changes in commodity programs, such as support prices. The five-year baseline forecasts that are discussed here do not make any attempt to predict changes in the laws governing program payments. Only the effects of the current law, given forecast conditions, are considered.

The value of consensus

No one knows the future, even the near future. The simple reason is that all the information that is necessary to "know" the future is unavailable. Thus, people make assumptions about the most important determinants of a variable in question, and then estimate the magnitude of the effects of changes in these determinants on that variable.

There can, *a priori*, be wrong assumptions, incomplete empiricism, or faulty reasoning. Wrong assumptions mean that the wrong determinants are chosen, or that the magnitude of change in those determinants is poorly described.

Forecasts are often made with the help of empirical models, often developed by economists, but which increasingly include non economic information such as weather. Despite the increasing sophistication of such models, no econometric model (even one including weather) can include all variables that may be relevant.

A typical empirical model might express the relationship of wheat production (Q) to expected price (P^e) in the following way:

$$Q = \alpha + \beta P^e$$

A rise in P^e, other things equal, would have the effect of raising Q by an *average* of β. However, we know that not all farmers will react to a known P^e in the same way, and they will not react the same way from year to year. The judgment of the empirical model, Q, will be tempered by other variables that do not enter into the model: expected weather, expected costs, changes in taxes, farm value, export prospects, relative risk-aversion, policy

[1] *Agricultural Outlook*, various issues, Table 34.

changes, and technology. There is no way, for example, to systematically factor in the former Soviet Union as an export market.

The baseline forecast was designed in a way so as to minimize the problem of (1) poor assumptions, (2) empirical limitations, and (3) illogical reasoning. The consensus approach to forecasting is an effort that recognizes where limitations arise, and consensus is an attempt to minimize those limitations. It also creates an institutional process whereby learning may occur.

Consensus forecasts recognize that "models" of all types, formal and informal, abstract from reality. Thus, expertise from a variety of agencies is brought to bear in order to try to incorporate all relevant information. Simplifying, one may describe each agency's role as part of the puzzle:

- The Economic Research Service (ERS) will identify the most important economic effects, and implications for prices, quantity supplied, and quantity demanded. This may include information on such diverse areas as exchange rates, oil prices, the effects of domestic and foreign agricultural policy, and economic growth.

- The Agricultural Stabilization and Cooperatives Service (ASCS) will describe the current policy environment. How many farmers will join commodity programs? Econometric models are also important here, tempered by specialist expertise. There is considerable interaction between ERS and ASCS, and a good bit of crossover in function. The result is more complete economic analysis.

- The Foreign Agricultural Service (FAS) will provide an overview of foreign production and trade, with implications for U.S. exports. Will the Commonwealth of Independent States (CIS) buy grain? Where does the Export Enhancement Program (EEP) stand?

- The Agricultural Marketing Service (AMS) and the Federal Crop Insurance Corporation (FCIC) often provide relevant anecdotal detail.

- The Office of Budget and Policy Analysis assists in interpreting current farm legislation.

- The World Agricultural Outlook Board convenes the forums which issue the supply and use forecasts.

No one of these organizations, obviously, has all the answers to the question of what will happen. The only way that these pieces can be fit together is by cooperation. The question then becomes one of whether or not the cooperation, and consensus, of this very diverse group, produces a result that stands up to serious scrutiny. The question then may be restated by asking whether or not all the information available is used efficiently.

Evaluating forecasts

Whenever forecasts differ from actual outcomes, the use of the term "error" is ubiquitous. A forecast "error" is defined here as the forecast minus the correct answer. The question is really one of why the difference occurs. The forecasting exercise that has used reasonable assumptions, recognized (and accounted for) incomplete empiricism, and eliminated faulty reasoning may not be in "error." The drought in July-August 1988 was not the result of poor empiricism in January of 1988. A projection of drought before it actually occurs, especially one with the severity of that in 1988, cannot be considered "reasonable," given current knowledge. Thus, the term "error" must be used cautiously.

The baseline projections contain many variables. There are, unfortunately, also many indicators of forecast accuracy. Choosing all available indicators is cumbersome, and may, in some circumstances, lead to ambiguous results. A preferred indicator, for evaluating overall performance, would be useful across all variables. There are four criteria that will be used here to select a "single" indicator:

1. The measure must be unit-free, so that we can discuss stocks, prices, and production without changing our frame of reference.

2. An unbiased examination of forecasts minimizes outliers -- these would be unusual events (such as export embargoes and droughts) which may dominate particular years and variables. An outlier may distort an indicator, and thus imply that a forecast in a more normal year is less useful than is actually the case.

3. A useful indicator would be one that can be used over the entire forecasting period, to determine whether or not the estimates are converging or diverging from the "correct" result as the horizon comes closer.

4. Last, we want an indicator that compares the forecast with some alternative, such as a "naive" forecast. This is a very inexpensive way of determining if the baseline consensus process is better than the simplest forecast.

There are several popular measures of forecast evaluation that appear frequently in the literature. The question is whether or not they meet the criteria set out above.[2] Two that meet all the above criteria are the relative absolute error (RAE) and Theil's U2 statistic. The RAE is calculated as:

$$RAE = \frac{\sum_{i=1}^{n} \left| \frac{(F_{it} - A_{it})}{(A_{i(t-1)} - A_{it})} \right|}{n}$$

where:

F_{it} = the current projection or estimate,
A_{it} = the actual value, and
$A_{i(t-1)}$ = the actual value in the period immediately before the forecast, as it is known at the time of the forecast.

An RAE of zero means that the projection or estimate was exactly correct. A value of one tells the forecast user that the "naive" forecast is as close to the actual value as is the projection or estimate. A value greater than one suggests that the "naive" forecast is better. Any value between zero and one indicates that the forecast adds enough information to improve on the naive model. We should also expect that the RAE will approach zero as we come closer to the forecast horizon. The RAE suffers, however, if there is no change from the previous year to the current year (as often occurs with target prices, for example). Such a circumstance leads the value to become infinite, which seriously distorts the average.

An indicator that is equivalent to the RAE is Theil's inequality statistic (also known as Theil's U2):

$$U2 = \frac{\left(\frac{\sum_{i=1}^{n} (F_{it} - A_{it})}{n} \right)^{1/2}}{\left(\frac{\sum_{i=1}^{n} (A_{i(t-1)} - A_{it})}{n} \right)^{1/2}} = \frac{RMSE}{RMSE_{rw}}$$

The numerator is the same as the root mean squared error (RMSE) and the denominator is the equivalent of an RMSE, using a "naive" (or random walk, subscript *rw*) for the forecast. The interpretation is the same as the RAE. If the U2 is greater than one, then the "naive" (or random walk) forecast is superior. Values less than one tell the user that the current forecast is superior to a random walk. The influence of an outlier is minimized, as with the RAE; it is contained in both the numerator and denominator.

The efficiency of the forecasts may be tested by estimating the following via regression:

---

[2] An excellent review of single indicators of forecast error is contained in J. Scott Armstrong and Fred Collopy "Error measures for generalizing about forecasting methods: empirical comparisons," *International Journal of Forecasting* 8 (1992) pp. 69-80.

$$A_t = \alpha + \beta F_t + u_t,$$

where A_t is the actual outcome, F_t is the forecast of that outcome, and u_t is a random error. An efficient forecast would have the estimator of α equal zero, and that for β equaling one. Deviations from these "perfect" values are evidence of inefficiency; better forecasts would result from knowledge of these parameters' values. Because both α and β must equal their "ideal" for efficiency to be established, the test must jointly restrict the two coefficients (Holden and Peel).[3] One may either use a Wald test or an F-test to test the null hypothesis of $\alpha=0$ and $\beta=1$. The Wald test is a Chi-square distribution and will be the one discussed below.

However, Holden and Peel also show that even if α and β differ from their "ideal," this does not necessarily indicate that a projection is biased. Instead, regressing the forecast error on a constant term provides the necessary and sufficient condition for determining bias, with the t-statistic telling the result. This test is equivalent to restricting β to one. Finding bias means that, even if α and β do not statistically differ from their respective "ideals," that the results can be called into question. However, if both efficiency and unbiasedness can be demonstrated, then the results strongly suggest a good forecasting record; information is being used economically.

The results

Forecasts are issued in January of each year. The current crop year has, at that time, not been completed. Thus, many components of supply and use are still estimated, such as production, and some are still subject to significant revision, such as market prices, exports, and ending stocks. This current crop year is labeled "Current year" in the results detailed in Tables 1-4. The first year out has crop years that begin June 1 for wheat and barley; August 1 for rice, cotton, and peanuts; September 1 for corn, sorghum, and soybeans; and December 16 for honey. Forecasts issued in January of 1992 then had the current crop year as 1991/92; the year ahead forecast as 1992/93 (beginning June 1, 1992, for wheat and barley, for example); the two years ahead forecast would be for the 1993/94 crop year (June 1, 1993 for wheat and barley); and the three years ahead forecasts would be for the 1994/95 crop year (June 1, 1994 for wheat and barley).

This study covers the January forecasts from 1970 through 1992. Long-term (more than one year out) forecasts were first issued in 1975. There are then 23 observations for current year projections, 22 for one-year projections, 16 for two-year projections, and 15 for three-year projections. The "actual" value is assumed to be that reported as the immediate past year for each of the forecasts. Thus the "actual" 1991/92 outcome is assumed to be that reported in the budget baseline document of January 1993.[4]

Supply and use categories, and prices, were chosen to maintain as much consistency as possible between commodities, for comparison and easy summarization. Area planted and harvested, obviously, however, does not apply to honey; the appropriate variable is number of colonies. The announced loan rate is the applicable loan rate for each commodity.[5] There are no explicit market price forecasts made for cotton (forbidden by law), and no target prices for soybeans, honey, or peanuts. Imports are significant components of supply and use for only honey and peanuts, of the crops covered herein. Commodity Credit Corporation (CCC) stocks are those bought by the USDA, voluntarily or not, for purposes of price stabilization or export programs such as PL-480.[6]

---

[3]*The Manchester School of Economic and Social Studies* 58 (June 1990), pp. 120-7.

[4]The 1991/92 values would have been issued in the November 1992 issue of *World Agricultural Supply and Demand Estimates* for major program commodities. Thus, the estimates are two months old at the time the budget baseline is completed. This is the case for all the "actual" observations as well as the estimates.

[5]The quota loan rate for peanuts, for example.

[6]A basic discussion of price stabilization and stock functions appears in *The Basic Mechanisms of U.S. Farm Policy Parts One, Two, and Three*, from the U.S. Department of Agriculture, Economic Research Service, Miscellaneous Publications numbers 1470 (May 1989), 1476 (December 1989), and 1477 (December 1989). A discussion of USDA storage programs appears in *Storage Subsidy Programs* from the U.S. Department of Agriculture, Economic Research Service, Staff Report AGES 9075, December 1990, pp. 4-5. Export programs are described in *Agricultural Export Programs: Background for 1990 Farm Legislation*, by Karen Z. Ackerman and Mark E. Smith, from the U.S. Department of Agriculture, Economic Research Service, Staff Report AGES 9033, May 1990.

Ending stocks are the difference between total supply (beginning stocks, imports, and production) and total use (domestic use plus exports). Errors for ending stocks are therefore not independent of errors in supply and use categories. Forecast errors from ending stocks will be "carried over," so to speak, into total supply for the subsequent years. Other things equal, then, forecast errors will automatically increase for supply and use variables as the forecast horizon recedes. Furthermore, changes in stocks are also strongly associated with changes in market prices. Thus, errors in stock level forecasts will produce errors in price forecasts. No attempt is made to adjust for these cumulative processes.

Table 1 summarizes a basic indicator of forecast accuracy, absolute percentage error. This indicator should demonstrate progressively higher errors as the forecast horizon recedes. The largest percentage errors, other than for CCC stocks, are for exports and ending stocks. Exports have varied widely in the 1980's because of significant changes in exchange rates, foreign economic growth rates, and U.S. export programs. All are difficult to forecast. The large CCC stock forecast percentage errors occur because of great volatility in relatively small holdings. Government-held stocks serve a variety of public policy purposes (export promotion, food security, price stability), not all of which are (or can be) well-identified more than a year in the future. Percentage errors in almost all categories are systematically higher as the forecast horizon recedes.

Low percentage errors for ending stocks in the current year for barley and wheat correspond to low percentage errors in market price forecasts for those two grains. Higher percentage errors in ending stocks for corn, sorghum, rice, and soybeans are also associated with higher average percentage errors in prices. These lower errors for wheat and barley, of course, could be the result of more of the crop year being completed at the time that the forecast is made.[7] The relative size of these errors, and their correspondence, persist until the second out year.

Table 2 shows Theil's inequality statistic (U2) for the same set of variables. The random walk assumption has it here that the previous year's value is maintained throughout the forecast horizon. For example, 1990/91 is the random walk assumption for forecasts issued in January 1992 for the "current" year 1991/92 and subsequent "out" years. The values of U2 show that the baseline forecasts add value for the current and year-ahead projections, with the exceptions of bee colonies (area harvested for honey) and CCC stocks. Most of the error in bee colonies comes from the fact that no statistics were collected between 1982 and 1986. In 1986, the number of colonies was determined to have dropped by 25 percent (from 4.3 million to 3.2 million). Current data now reflect only beekeepers with at least five colonies.[8]

Market price forecasts for those commodities with the greatest budget exposure (all but soybeans; cotton still being excluded) are better than "naive" forecasts for the current year and first year out, according to the U2 statistic. Other things equal, these market price forecasts mean that the supply and use estimates for the following fiscal year are better than a "naive" forecast. Two years out, the only observations of prices that suggest using a naive forecast are soybeans and honey. The two and three year estimates for exports, in general, have the least acceptable U2 values, other than those for CCC stocks.

Efficiency statistics are summarized in Table 3. The null hypothesis being tested is the Wald test of $\alpha=0$ and $\beta=1$. The designation "not efficient" means that the null hypothesis could not be accepted even at the 1 percent level. A "10" is the strongest acceptance of efficiency. The "year ahead" forecasts begin the systematic process by which exports and ending stocks are the least efficient forecasts, in general. These lead to price forecasts also being largely inefficient. By the time that the third year ahead is reached, only yield projections are mostly efficient.

Bias statistics appear in Table 4. These results, are, generally, better than those for efficiency. There is almost uniform unbiasedness to the forecasts. The notable exception is a downward bias for area for peanuts. Peanuts operate under a supply control program based on poundage. Thus, there is no reason to economize on land resources; less productive land may be brought into use rather than applying expensive soil enrichment resources.

[7]See fn. 4.

[8]See *Honey: Background for 1990 Farm Legislation*, by Frederic L. Hoff and Jane K. Phillips, from the U.S. Department of Agriculture, Economic Research Service, Staff Report AGES 89-43, September 1989.

Table 1. Absolute percentage errors for program commodity supply and use, forecasts issued 1970-1992.

	Corn	Sorghum	Barley	Wheat	Rice	Soybeans	Upland cotton	Honey	Peanuts
Current year									
Area planted	0.3	1.4	0.4	0.3	1.3	0.6	2.4	--	1
Area harvested	0.9	2.2	0.8	0.4	1.4	0.7	1.8	4.9	0
Yield	0.8	2.5	1.3	0.3	0.6	1.5	1.8	6.8	1
Production	1.3	2.5	1.5	0.4	1.6	2.0	1.9	5.0	1
Domestic use	3.5	10.5	5.1	5.6	5.8	7.5	3.9	4.3	7
Exports	12.2	20.1	24.6	5.0	5.8	8.3	11.3	9.0	18
Ending stocks	24.5	34.9	11.1	11.2	25.1	29.9	15.0	13.6	23
CCC inventory	86.1	84.0	122.6	57.8	96.7	234.6	245.2	29.8	--
Announced loan		0.2	5.0	4.9	4.8	1.2	0.6	0.3	0.10
Market price	7.6	7.1	2.9	2.8	5.7	6.8	--	6.9	4
Target price	0.0	0.0	0.0	0.0	0.0	--	0.1		
Imports								4.4	35
Year ahead									
Area planted	3.5	6.8	10.8	3.7	8.7	4.1	6.7	--	5
Area harvested	3.7	7.0	11.8	4.3	8.1	4.4	7.7	7.1	5
Yield	11.6	12.2	7.8	7.5	3.7	7.2	9.8	15.6	12
Production	14.2	15.3	15.2	8.2	7.7	9.7	14.0	12.3	11
Domestic use	8.1	16.9	9.1	7.7	7.7	10.7	10.3	10.6	11
Exports	26.1	21.7	64.2	16.7	13.1	14.8	26.8	35.2	34
Ending stocks	59.7	63.1	30.9	35.2	49.7	60.3	55.2	33.8	20
CCC inventory	693.5	353.8	421.1	117.1	888.1	928.0	1,485.3	77.2	--
Announced loan		3.2	13.5	6.3	8.2	6.0	3.5	0.8	2.92
Market price	19.8	18.5	16.0	15.1	21.8	24.4	--	13.4	5
Target price	1.8	3.2	8.3	2.2	1.4	--	1.5		
Imports								39.7	55
Two years ahead									
Area planted	8.4	14.7	13.4	10.3	18.6	8.8	16.5	--	10
Area harvested	8.7	14.8	14.8	12.6	18.8	9.0	17.0	9.8	10
Yield	10.7	12.0	8.5	7.9	5.0	7.3	9.8	17.5	15
Production	15.6	19.3	18.7	16.7	15.0	13.2	18.3	14.3	15
Domestic use	7.0	24.3	11.7	14.3	10.1	7.6	13.1	13.4	11
Exports	32.9	18.3	43.4	20.9	22.1	21.8	27.8	32.7	43
Ending stocks	66.7	73.3	43.6	42.2	51.6	39.8	51.0	37.2	24
CCC inventory	205.4	192.4	224.0	138.7	2,802.4	155.2	196.5	98.9	--
Announced loan		13.5	13.5	18.3	14.4	12.8	11.3	4.4	6.35
Market price	20.0	22.7	17.9	22.0	33.2	19.9	--	12.3	9
Target price	5.5	7.7	14.2	5.8	3.5	--	4.9		
Imports								39.1	74
Three years ahead									
Area planted	8.8	15.4	14.1	12.7	23.8	10.6	20.4	--	10
Area harvested	9.0	14.6	16.0	17.9	23.9	10.8	20.9	13.6	10
Yield	11.0	11.2	10.9	9.2	6.7	7.5	10.5	18.8	18
Production	18.4	18.8	19.9	19.3	20.0	14.6	19.1	16.2	16
Domestic use	6.3	23.1	11.4	21.0	10.3	7.3	16.8	12.1	13
Exports	38.2	19.8	68.5	22.1	31.3	28.0	33.9	44.7	45
Ending stocks	65.9	80.5	63.9	54.7	57.4	38.3	34.2	40.2	24
CCC inventory	240.3	229.9	206.6	203.9	3,406.6	243.5	190.9	130.8	--
Announced loan		18.3	18.4	23.5	21.2	18.6	16.4	7.8	11.09
Market price	22.9	24.6	24.6	25.0	40.7	22.3	--	15.5	13
Target price	8.4	12.0	20.2	9.5	6.8	--	5.5		
Imports								50.0	79

-- No information available or not relevant.

Table 2. Theil's inequality statistic for program commodity supply and use, forecasts issued 1970-1992.

	Corn	Sorghum	Barley	Wheat	Rice	Soybeans	Upland cotton	Honey	Peanuts	Exceeding one?
Current year										
Area planted	0.05	0.15	0.07	0.04	0.11	0.10	0.18	--	0.29	0 of 8
Area harvested	0.11	0.18	0.09	0.04	0.13	0.13	0.13	1.44	0.20	1 of 9
Yield	0.06	0.24	0.19	0.06	0.16	0.17	0.16	0.53	0.17	0 of 9
Production	0.07	0.15	0.13	0.05	0.14	0.15	0.10	0.37	0.15	0 of 9
Domestic use	0.52	0.61	0.78	0.53	0.88	0.91	0.53	0.47	0.57	0 of 9
Exports	0.76	1.04	0.47	0.32	0.46	0.54	0.41	0.22	0.73	1 of 9
Ending stocks	0.39	0.45	0.43	0.38	0.53	0.79	0.35	0.52	0.87	0 of 9
CCC inventory	0.91	0.89	0.87	0.96	1.04	0.53	1.08	0.91	--	2 of 8
Announced loan	0.08	0.65	0.79	0.55	0.26	0.19	0.11	0.01	0.21	0 of 9
Market price	0.58	0.60	0.19	0.15	0.28	0.48	--	0.89	0.59	0 of 8
Target price	0.00	0.00	0.00	0.00	0.00	--	0.05			0 of 6
Imports								0.17	0.32	0 of 2
Year ahead										
Area planted	0.34	0.46	0.78	0.31	0.59	0.54	0.40	--	0.97	0 of 8
Area harvested	0.32	0.46	0.77	0.34	0.54	0.57	0.43	1.11	1.00	2 of 9
Yield	0.76	0.84	0.79	1.00	0.68	0.72	0.81	0.72	0.93	0 of 9
Production	0.58	0.63	0.80	0.57	0.51	0.71	0.52	0.65	0.81	0 of 9
Domestic use	0.83	0.86	0.83	0.74	0.82	0.93	1.00	0.91	0.93	0 of 9
Exports	1.02	0.89	0.93	0.73	0.92	0.84	0.74	0.88	0.99	1 of 9
Ending stocks	0.72	0.65	0.72	0.68	0.71	0.90	0.85	0.85	0.65	0 of 9
CCC inventory	1.19	0.90	1.09	1.00	0.96	0.98	1.35	0.95	--	3 of 8
Announced loan	0.38	0.73	0.77	0.59	0.63	0.36	0.16	0.19	0.41	0 of 9
Market price	0.81	0.80	0.73	0.67	0.84	1.25	--	0.74	0.66	1 of 8
Target price	0.20	0.51	1.21	0.27	0.34	--	0.36			1 of 6
Imports								0.74	0.48	0 of 2
Two years ahead										
Area planted	0.84	0.80	0.74	0.74	0.96	0.86	0.87	--	1.47	1 of 8
Area harvested	0.82	0.74	0.75	0.79	0.97	0.89	0.94	1.22	1.45	2 of 9
Yield	0.79	0.84	0.79	1.13	0.83	0.58	0.96	0.75	1.26	2 of 9
Production	0.75	0.71	0.80	0.91	0.86	0.79	0.95	0.81	1.10	1 of 9
Domestic use	0.65	0.99	0.82	0.92	0.65	0.70	0.98	1.28	0.93	1 of 9
Exports	1.19	0.81	1.00	0.83	1.17	0.93	0.78	0.70	1.05	4 of 9
Ending stocks	0.96	0.66	1.04	0.81	1.01	0.80	1.07	0.78	0.83	3 of 9
CCC inventory	1.50	0.80	1.24	1.23	1.45	0.85	0.62	1.12	--	5 of 8
Announced loan	0.67	0.67	0.84	0.66	1.05	0.69	0.31	0.39	0.63	1 of 9
Market price	0.87	0.89	0.97	0.89	0.96	1.18	--	1.23	0.87	2 of 8
Target price	0.36	0.67	1.07	0.43	0.51	--	0.46			1 of 6
Imports								0.86	0.44	0 of 2
Three years ahead										
Area planted	1.16	0.84	0.71	1.04	1.16	0.92	0.89	--	1.33	4 of 8
Area harvested	1.06	0.71	0.68	1.09	1.19	0.95	0.93	1.26	1.28	5 of 9
Yield	0.76	0.74	0.97	1.04	0.88	0.75	0.76	0.87	1.13	2 of 9
Production	0.87	0.66	0.75	0.99	1.02	0.90	0.83	1.16	0.88	2 of 9
Domestic use	0.59	0.82	0.75	0.96	0.54	0.67	1.01	1.30	1.16	3 of 9
Exports	1.29	1.09	0.97	0.88	1.32	1.21	0.78	0.90	1.14	5 of 9
Ending stocks	1.01	0.65	1.30	0.96	0.95	0.82	0.75	0.84	0.77	2 of 9
CCC inventory	1.95	0.79	1.24	1.50	1.27	0.93	0.61	1.13	--	5 of 8
Announced loan	0.70	0.70	0.88	0.70	1.17	0.80	0.47	0.49	0.81	1 of 9
Market price	0.94	0.99	1.28	1.13	1.19	1.68	--	1.34	1.12	6 of 8
Target price	0.50	0.84	0.94	0.52	0.63	--	0.34			0 of 6
Imports								1.00	0.49	1 of 2

--Not available or not relevant.

Table 3. Forecast efficiency for program commodity supply and use, forecasts issued 1970-1992.[1]

	Corn	Sorghum	Barley	Wheat	Rice	Soybeans	Upland cotton	Honey	Peanuts
Current year									
Area planted	10	10	10	10	10	5	10	10	10
Area harvested	NE	1	10	10	10	10	10	NE	1
Yield	10	10	10	10	10	10	10	1	10
Production	1	10	5	NE	10	10	10	10	10
Domestic use	10	10	10	5	10	NE	10	1	5
Exports	1	NE	10	10	10	10	10	5	10
Ending stocks	10	10	.10	5	10	NE	10	10	N
CCC inventory	NE	1	1	NE	NE	NE	NE	10	--
Announced loan	NE	10	10	10	10	10	10	10	10
Market price	NE	NE	10	10	10	NE	--	NE	10
Target price	10	10	10	10	10	--	10		
Imports								10	N
Year ahead									
Area planted	10	10	10	10	5	10	10	10	10
Area harvested	10	10	10	10	10	10	10	NE	10
Yield	10	5	10	1	10	10	10	10	N
Production	10	10	10	10	10	10	10	NE	1
Domestic use	10	NE	10	5	10	NE	10	NE	N
Exports	NE	NE	NE	1	NE	5	10	NE	N
Ending stocks	10	10	NE	NE	10	NE	NE	10	10
CCC inventory	NE	NE	NE	NE	NE	NE	NE	1	--
Announced loan	NE	10	5	10	5	10	10	NE	10
Market price	NE	NE	NE	1	NE	NE	--	NE	10
Target price	10	5	NE	5	10	--	10		
Imports								5	10
Two years ahead									
Area planted	1	1	10	NE	NE	NE	5	10	N
Area harvested	1	10	10	1	NE	NE	5	NE	N
Yield	10	1	5	NE	10	10	10	10	N
Production	5	10	5	NE	NE	1	10	1	1
Domestic use	10	NE	10	10	10	10	10	NE	N
Exports	NE	1	1	1	NE	NE	5	NE	N
Ending stocks	NE	10	NE	NE	NE	5	1	10	5
CCC inventory	NE	NE	NE	NE	NE	NE	10	NE	--
Announced loan	NE	NE	NE	NE	NE	NE	NE	NE	1
Market price	NE	NE	NE	NE	NE	NE	--	NE	10
Target price	N	NE	NE	NE	1	--	5		
Imports								NE	10
Three years ahead									
Area planted	NE	NE	10	NE	NE	NE	NE	10	N
Area harvested	NE	10	10	NE	NE	NE	NE	NE	N
Yield	5	5	NE	NE	10	10	5	10	N
Production	1	10	5	NE	NE	NE	1	NE	1
Domestic use	10	NE	5	10	10	NE	10	NE	N
Exports	N	NE	NE	NE	NE	NE	NE	NE	N
Ending stocks	NE	10	NE	NE	NE	1	10	1	1
CCC inventory	NE	1	NE	NE	NE	1	10	NE	--
Announced loan	NE	NE	NE	NE	NE	NE	NE	NE	N
Market price	NE	NE	NE	NE	NE	NE	--	NE	1
Target price	NE	NE	NE	NE	NE	--	NE		
Imports								NE	10

NE = "Not efficient"

--Not available or not relevant.

[1]The test summarized is the Wald test, a joint test of the slope equaling one and the intercept zero. A "10" means this null hypothesis cannot be rejected even at the 10 percent level. "Not efficient" means that the null hypothesis is rejected. The 10 percent level is the most stringent criteria for acceptance of the null hypothesis.

Table 4. Forecast bias for program commodity supply and use, forecasts issued 1970-1992.[1]

	Corn	Sorghum	Barley	Wheat	Rice	Soybeans	Upland cotton	Honey	Peanuts
Current year									
Area planted	NB	-10	NB	NB	NB	10	NB	--	N
Area harvested	-1	-5	NB	NB	NB	NB	NB	10	N
Yield	NB	NB	-10	NB	NB	NB	-10	-1	N
Production	-1	NB	NB	-10	NB	NB	NB	NB	N
Domestic use	NB	NB	NB	NB	NB	NB	NB	NB	1
Exports	NB	-5	NB	NB	NB	NB	-10	NB	N
Ending stocks	NB	NB	NB	-10	NB	NB	10	NB	N
CCC inventory	NB	NB	NB	NB	NB	NB	NB	NB	--
Announced loan	NB	NB	NB	NB	NB	NB	NB	NB	N
Market price	NB	NB	NB	NB	NB	NB	--	NB	N
Target price	NB	NB	NB	NB	NB	--	NB		
Imports								NB	N
Year ahead									
Area planted	NB	NB	NB	NB	NB	NB	NB	--	-1
Area harvested	NB	NB	NB	NB	NB	NB	NB	5	-1
Yield	NB	NB	NB	NB	NB	NB	NB	NB	5
Production	NB	NB	NB	NB	NB	NB	NB	5	N
Domestic use	NB	10	NB	NB	NB	NB	NB	10	N
Exports	NB	-10	NB	NB	NB	NB	NB	NB	N
Ending stocks	NB	NB	NB	NB	NB	NB	NB	NB	N
CCC inventory	NB	NB	5	NB	NB	NB	10	NB	--
Announced loan	-10	NB	-5	NB	NB	NB	NB	-5	N
Market price	NB	NB	-10	NB	NB	NB	--	NB	-1
Target price	NB	NB	NB	NB	NB	--	NB		
Imports								NB	N
Two years ahead									
Area planted	NB	NB	NB	NB	NB	NB	NB	--	-5
Area harvested	NB	NB	NB	NB	NB	NB	NB	5	-5
Yield	NB	NB	NB	NB	NB	NB	-10	NB	1
Production	NB	NB	NB	NB	NB	NB	NB	5	N
Domestic use	NB	NB	NB	-10	NB	NB	NB	NB	N
Exports	NB	NB	-10	NB	NB	NB	NB	NB	N
Ending stocks	NB	NB	NB	NB	NB	NB	NB	NB	N
CCC inventory	NB	NB	NB	NB	NB	NB	NB	NB	--
Announced loan	NB	NB	-10	NB	1	NB	NB	-10	N
Market price	NB	NB	NB	NB	NB	NB	--	NB	N
Target price	NB	NB	NB	NB	10	--	NB		
Imports								NB	N
Three years ahead									
Area planted	NB	1	NB	NB	NB	NB	NB	--	-5
Area harvested	NB	10	NB	NB	NB	NB	NB	5	-1
Yield	NB	NB	NB	NB	NB	NB	-5	NB	1
Production	NB	NB	NB	NB	NB	NB	NB	10	N
Domestic use	NB	NB	NB	-10	NB	NB	NB	NB	N
Exports	NB	NB	NB	NB	NB	NB	NB	NB	N
Ending stocks	NB	NB	NB	NB	NB	NB	NB	NB	N
CCC inventory	NB	NB	NB	10	NB	NB	NB	NB	--
Announced loan	NB	NB	NB	NB	1	NB	NB	-10	N
Market price	NB	NB	NB	NB	NB	NB	--	NB	N
Target price	NB	NB	NB	NB	10	--	NB		
Imports								NB	N

NB = "Not Biased"

-- Not available or not relevant.

[1]The test summarized is based on a regression of the forecast error (forecast minus actual) on a constant. The t-test is used to determine whether or not the constant (average error) is significantly different from zero. A value of minus 10 in the table means that the average is significantly different from zero at the 10 percent level and that the forecast is biased downward. A positive 10 means that the forecast is biased upwards, and differs from zero at the 10 percent level.

The role of forecasts in policy and the role of policy in forecast error

The January baseline forecasts are, as mentioned earlier, based on current law; no attempt is made to forecast policy changes outside current legislation. However, these forecasts can serve as warnings to policy makers that changes are needed. One example, leading to the payment-in-kind (PIK) program, occurred in 1983.

Supply and use forecasts for corn in January of 1983 implied (Table 5) that the huge stock buildup in 1982/83, resulting from large production and carryover, would send market prices below the loan rate for most of 1982/83 and at least part of 1983/84. This would mean that the government would be forced to accept grain at the loan rate, make high program payments (the difference between the target price and loan rate per bushel), and carry large grain stocks into the near future. The purchase of an additional 1.1 billion bushels of corn at the loan rate would cost $2.6 billion dollars alone. Significant budgetary pressure demanded that something be done, both for the current year and the year ahead.

The decision was made to "pay" farmers not to produce in 1983/84 by giving them grain carried over from 1982/83.[9] The result would be to effectively lower current stock levels, raising current prices above the loan rate. At the same time, lower production in 1983/84 would reduce the supply pressure in that crop year.

The PIK program helped to reduce production and raise prices in both 1982/83 and 1983/84 (partially assisted by a drought in much of the Corn Belt).[10] One additional effect of PIK was to significantly increase the forecast error for 1983/84, as a result of the abrupt policy shift. The percentage error on ending stocks for 1983/84, for example, is almost 300 percent. For CCC stocks, the error is over 1000 percent. Whatever one believes about the efficacy of the PIK program, the forecast "error" introduced by the change in policy clearly affected the appearance of accuracy in Table 1.[11] The question is whether or not good-faith forecasts affect policy decisions. Clearly, they can. One could argue that it is, in fact, one of their functions.

Conclusions

The USDA prepares, for internal use, long-term forecasts of program commodity supply and use each January, as part of a government-wide effort to estimate obligations for the next fiscal year. These supply and use forecasts are the result of consensus by several different agencies. Absolute percentage errors tend to increase as the forecast horizon recedes. This is partly due to the propagation of past errors, assuming that ending stocks are a residual term after the current year. Errors are particularly large for exports, ending stocks, and government-held (CCC) stocks.

Theil's U2 statistic indicates that, up until exports and ending stocks in the second year out, USDA forecasts add value to what is known about the past. The forecasts are useful. However, the lack of efficiency, especially starting with exports and propagated from there to ending stocks in the first year ahead, indicates that crucial information is not being included. Despite being inefficient, there is very little sign of bias in the forecasts. This may imply that there is no systematic "fix" to the problem of inefficiency.

The PIK program was used as an example of how projections may have been used to revise existing policies or implement new ones. Thus, an accurate *ex ante* forecast can be put to use to offset "problems" before they become severe. However, a forecast that leads to such use implies, *ex post,* that the forecaster is very poor at the job!

---

[9]A description of the program choices faced by farmers subsequent to the announcement of PIK, but before its actual implementation, may be found in *Analyzing the 1983 Payment in Kind Program at the Farm Level* by Fred J. Benson and Paul Hasbagen, University of Minnesota, Department of Agricultural and Applied Economics, Staff Paper series number P83-7, January, 1983.

[10]One could argue, from Table 5, that the effect of PIK was to reduce area planted from 70 to 60 million acres, while the drought reduced yields by almost 30 percent. Thus, the drought effectively turned a three-year program into a one-year program. We are grateful to Paul Westcott of ERS for this observation.

[11]See "Agricultural Policy," by Richard Rizzi, in *Agricultural Outlook* (January-February 1983), pp. 20-1, for a further example of how projections for the 1982/83 and 1983/84 crop year were affected.

Table 5. Supply and use forecasts for corn, issued in January 1983, and actual outcome

	History[1]		Forecast			Actual		
	1980/81	1981/82	1982/83	1983/84	1984/85	1982/83[2]	1983/84[3]	1984/85[4]
				Million acres				
Area planted	84.0	84.2	81.9	70.0	74.0	81.9	60.2	80.4
Area harvested (grain)	73.0	74.6	72.8	60.0	64.0	73.2	51.4	71.8
				Bushels per acre				
Yield	91.0	109.9	114.4	114.6	114.8	114.8	81.0	106.7
				Million bushels				
Beginning stocks	1,617	1,034	2,366	3,497	2,884	2,286	3,120	723
Production	6,645	8,201	8,330	6,876	7,347	8,397	4,166	7,656
Imports	1	1	1	1	1	1	2	3
Total supply	8,263	9,236	10,697	10,374	10,232	10,684	7,288	8,382
Domestic use	4,874	4,903	5,100	5,265	5,450	5,674	4,700	5,165
Exports	2,355	1,967	2,100	2,225	2,300	1,870	1,866	1,838
Total use	7,229	6,870	7,200	7,490	7,750	7,544	6,566	7,003
Ending stocks	1,034	2,366	3,497	2,884	2,482	3,140	722	1,379
CCC stocks	510	1,673	2,723	2,280	1,735	1,150	201	240
				Dollars per bushel				
Loan rate	2.25	2.40	2.55	2.65	2.55	2.55	2.65	2.55
Market price	3.11	2.45	2.30	2.65	2.65	2.70	3.20	2.65
Target price	2.35	2.40	2.70	2.86	2.86	2.70	2.86	3.03

[1]As of January 1983. [2]As known in January 1984. [3]As known in January 1985. [4]As known in January 1986.

Evaluating the Accuracy of the Short-Term Hydroelectric Generation Model Forecast

Robin D. Reichenbach, Energy Information Administration

The Short-Term Hydroelectric Generation Model (STHGM) was developed by the Energy Information Administration (EIA) to provide up to a 27-month forecast of utility net hydroelectric generation for publication in the quarterly Short-Term Energy Outlook (STEO) reports.

Utility net generation is gross generation minus plant use. Utility net hydroelectric generation includes generation from conventional hydroelectric facilities as well as pumped storage plants. Conventional facilities include storage, run-of-river, and diversion facilities. Storage facilities feature reservoirs created by dams and typically have additional functions such as flood control and recreation. Run-of-river facilities use natural stream flow along with a small dam. Some run-of-river facilities also impound water behind the dam to store enough energy for use during peak electric demand hours. Diversion facilities involve a man-made channel or aqueduct with sufficient slope for the water flow to drive a turbine. Pumped storage facilities repeatedly recycle water by pumping water discharged from the turbines to a lower retaining pool back into an upper storage facility for peak power production. Pumped storage facilities have a negative net generation, since the electricity consumed to pump the water exceeds the amount produced.

Hydroelectric generation provided 9 percent of total U.S. generation in 1992, and historically has provided as much as 14 percent. In 1992, approximately 3,500 hydroelectric facilities existed in the United States, including 144 pumped storage facilities.

Average national net hydroelectric generation for 1970 to 1992 was 275 billion kilowatthours; however, annual hydroelectric generation fluctuated a great deal (Figure 1). Between 1970 and 1992, generation ranged from a low of 220 billion kilowatthours in 1977 to a high of 332 billion kilowatthours in 1983. This variation was caused by several factors, including the dependency of hydroelectric generation on precipitation, the relative share of conventional versus pumped storage capacity, and regulations governing reservoir levels for alternative water uses.

Seasonality appeared in the monthly data for 1990 and 1991 (Figure 2) and was also apparent for other historical years. Seasonality appeared in 1992 as well, although it was less obvious. In all three of these years, generation decreased in April, peaked in mid-summer, and then bottomed out in late summer. Many explanations can be found for this seasonality. One reason is the late summer decrease in generation due in part to storage facilities maintaining higher water levels, particularly in recreational areas.

The hydroelectric generation data were also analyzed for trends in the data. From 1970 to 1992, conventional hydroelectric generating capacity increased by 18 gigawatts, from 53 gigawatts to 71 gigawatts. During the same time period, pumped storage hydroelectric generating capacity increased by 15 gigawatts, from 4 gigawatts to 18 gigawatts. However, even with these additions to capacity, generation from 1970 to 1992 remained stationary. No trend appeared in the generation data and, therefore, was not included in the model. It is believed that the increase in conventional generation was offset by the increase in net negative generation from the pumped storage facilities.

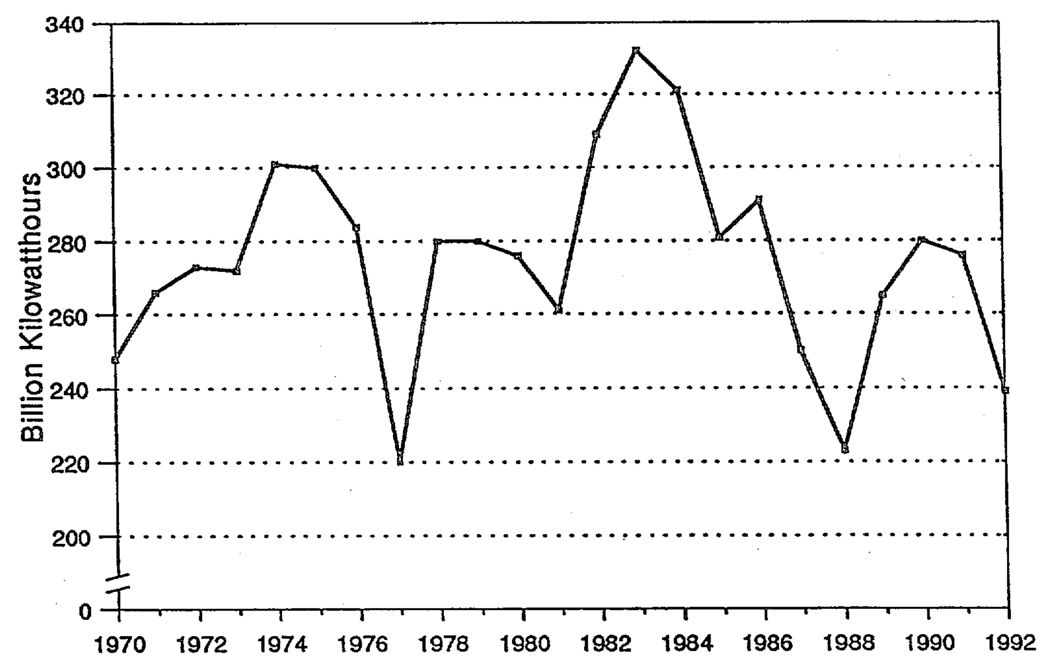
Census Division concentrations of hydroelectric generating stations are quite varied. Hydroelectric resources are not evenly distributed across the Nation, but are generally concentrated where precipitation and mountains combine to provide large water volumes. The Nation's most concentrated hydropower regions are the Pacific Division with 50 percent of total national hydroelectric generation in 1992, the Middle Atlantic Division with 11 percent, and the Mountain Division with 11 percent, although some major facilities are located in other areas (Figure 3). These three areas generated 172 billion kilowatthours of hydroelectric power in 1992 or 72 percent of the Nation's 239 billion kilowatthours of net hydroelectric generation. The most concentrated States (those generating more than 15 billion kilowatthours in 1992) were California, New York, Oregon, and Washington. All but New York are located in the Pacific Division.

With that background information, I will now describe the Short-Term Hydroelectric Generation Model. The Short-Term Hydroelectric Generation Model forecasts national net hydroelectric generation by utilities using an autoregressive integrated moving average (ARIMA) time series model with deterministic seasonality and precipitation as an explanatory variable, otherwise known as a transfer function model.[1] The model is based on 23 years of monthly historical data from 1970 to 1992.

---

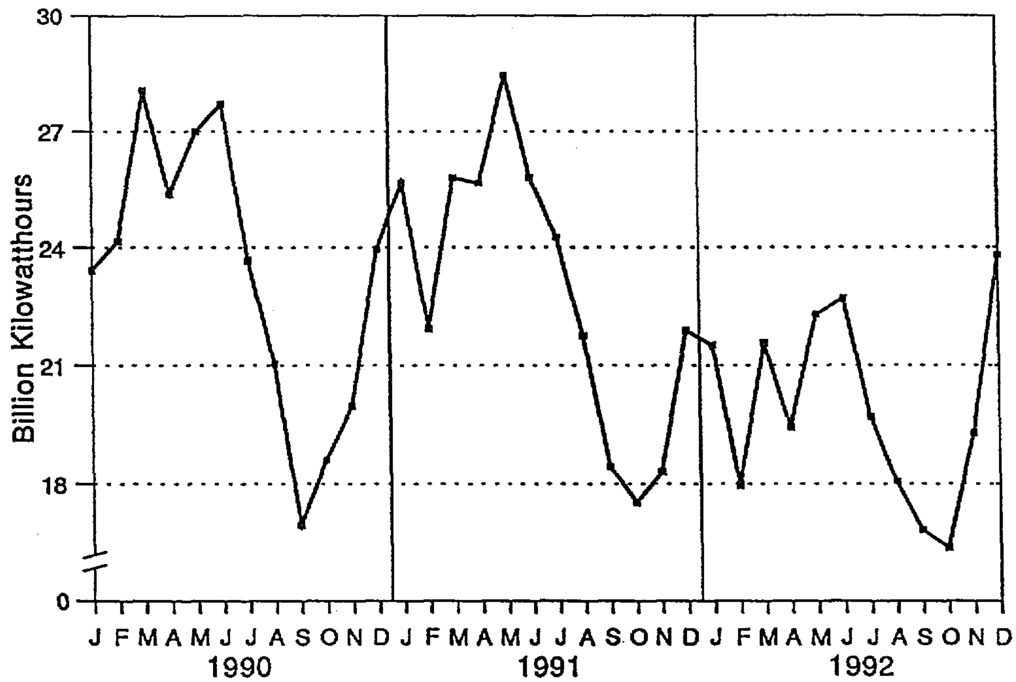
[1]George E. Box and Gwilym Jenkins, *Time Series Analysis, Forecasting and Control* (1976).

Figure 1
National Net Hydroelectric Generation by Year
1970 - 1992



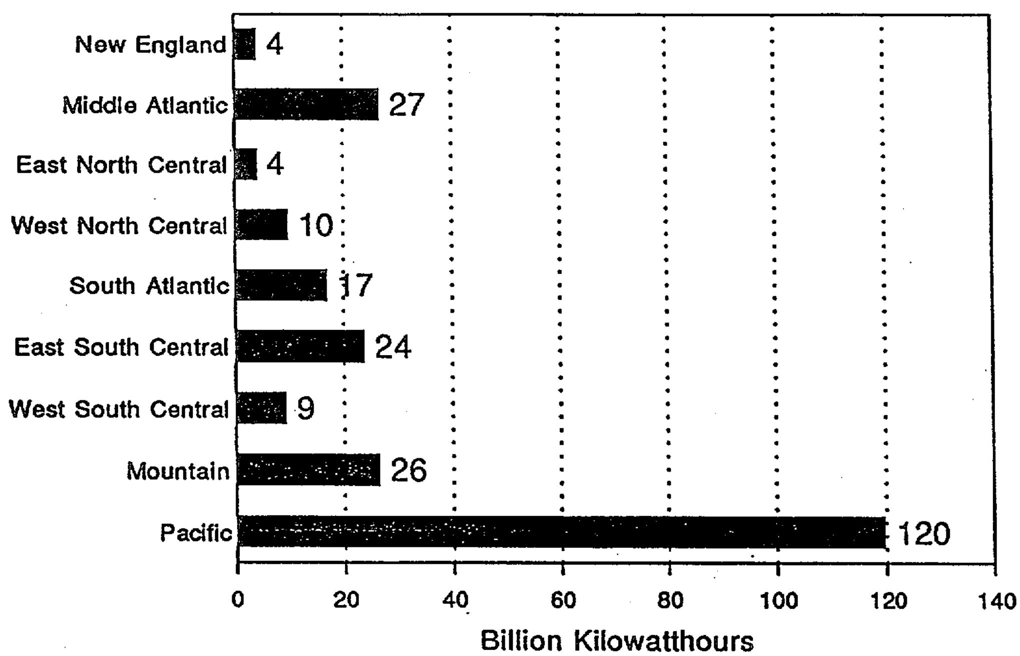
Source: Energy Information Administration, Form EIA-759, "Monthly Power Plant Report" and its predecessors.

Figure 2
National Net Monthly Generation
1990 - 1992



Source: Energy Information Administration, Form EIA-759, "Monthly Power Plant Report."

Figure 3
Net Hydroelectric Generation by Census Division
1992



Source: Energy Information Administration, Form EIA-759, "Monthly Power Plant Report."

Since historical monthly net hydroelectric generation data are correlated and the level of generation is affected by the availability of water, a time series model structure was selected. The model predicts future generation as a function of past generation, fixed seasonal components, and a lagged effect from precipitation levels. Precipitation levels determine the availability of water in reservoirs for electricity generation, as well as the effects of recent rain or snow on run-of-river hydroelectric facilities. It was found that lags of precipitation of 1, 2, 8, and 9 months were important in predicting generation.

The model is based on normal ARIMA modeling assumptions. That is, the data are assumed to follow a model which is linear and time invariant with constant coefficients. Fixed monthly values (dummy variables with estimated coefficients) are used to describe seasonality. For forecasting, precipitation is assumed to be "normal" from the last month of available data forward, where normal is defined to be the average monthly precipitation for a 57-year period.

Inputs to the model include two data series: (1) national totals of the plant-level net hydroelectric generation data reported monthly on the Form EIA-759, "Monthly Power Plant Report," and its predecessors; and (2) monthly precipitation data from the National Climatic Data Center.[2]

Specific model structure was selected from the family of ARIMA models based on standard diagnostics: the autocorrelation function, partial autocorrelation function, and inverse autocorrelation function of both the generation data and the model residuals, t-statistics for estimated parameters, the Portmanteau goodness of fit Chi square test for departures from model assumptions, the Akaike Information Criterion (AIC) to select the best from competing models, and comparisons of the model forecasts to actual data otherwise known as out-of-sample testing.[3]

The model was developed using the software SAS version 6. The initial step of the model reads in the historical net hydroelectric generation and precipitation data (Figure 4), and the data are divided by the number of days in each month in order to account for the different lengths of the months. Next, the first ARIMA model for generation is run to estimate coefficients for the monthly dummy variables and the appropriate lags of the precipitation data. The ARIMA procedure is a standard procedure in SAS and is based on the Box-Jenkins strategy for time series modeling. The procedure provides information for time series model identification, parameter estimation, and forecasting.[4]

As noted above, the model includes two functions that explain the time varying mean of the generation data, one due to seasonal variation and one due to the impact of precipitation. The seasonal effect on the generation is estimated using dummy variables each representing a month and estimated coefficients from the first ARIMA run. The impact of precipitation on the generation is estimated using the actual precipitation for past months and coefficients estimated in the first ARIMA run.

For purposes of forecasting, these two mean functions are subtracted from the historical generation data and the ARIMA procedure is run with this adjusted data, reestimating the autoregressive and moving average parts of the model. This ARIMA procedure provides the forecast for adjusted generation.

The seasonal effects and precipitation effects (using the assumed "normal precipitation") are added to the forecast of adjusted generation and the sum is multiplied by the number of days in each month. The output from this step is the final forecast.

After developing this model, procedures were established to evaluate the accuracy of the forecast and compare the results with our previous methodology. Until recently, the EIA had been using an informal methodology to produce the net hydroelectric generation forecasts. For each month during the current water year (which ends September 30) or a nine-month period from presently available data if longer, hydroelectric generation projections were based upon information obtained by phone from 10 utilities or organizations representing eight U.S. geographic regions. Total hydroelectric generation in each of the eight regions was projected to change by the same percentage as generation for the organizations contacted by phone.

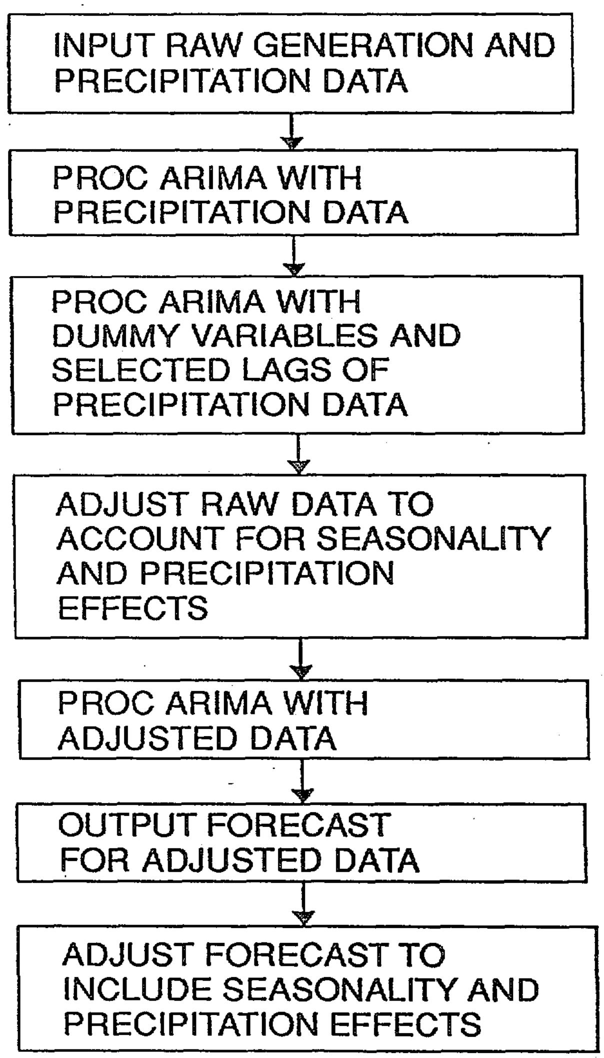
---

[2]**1970-1987:** National Climatic Data Center, *State, Regional, and National Monthly and Annual Precipitation Weighted by Area for the Contiguous United States January 1931 - December 1987* (Asheville, NC, August 1988), p. 66. **1988-1992:** National Climatic Data Center, *Monthly State, Regional and National Heating Degree Days Weighted by Population* (Asheville, NC, March, 1989 through 1993), Table 3.3, "Regional and National Average Precipitation."

[3]For more information, see: SAS Institute Inc., *SAS/ETS User's Guide*, Version 6 (Cary, NC, January 1989).

[4]For more information, see: SAS Institute Inc., *SAS/ETS User's Guide*, Version 6 (Cary, NC, January 1989), pp. 99-100.

Figure 4
Flow Diagram of the STHGM



Hydroelectric generation in the succeeding years was assumed to be normal. Normal generation was calculated by using hydroelectric capacity information from the Form EIA-860, "Annual Electric Generator Report," and historical monthly capacity factors averaged over 10 years.

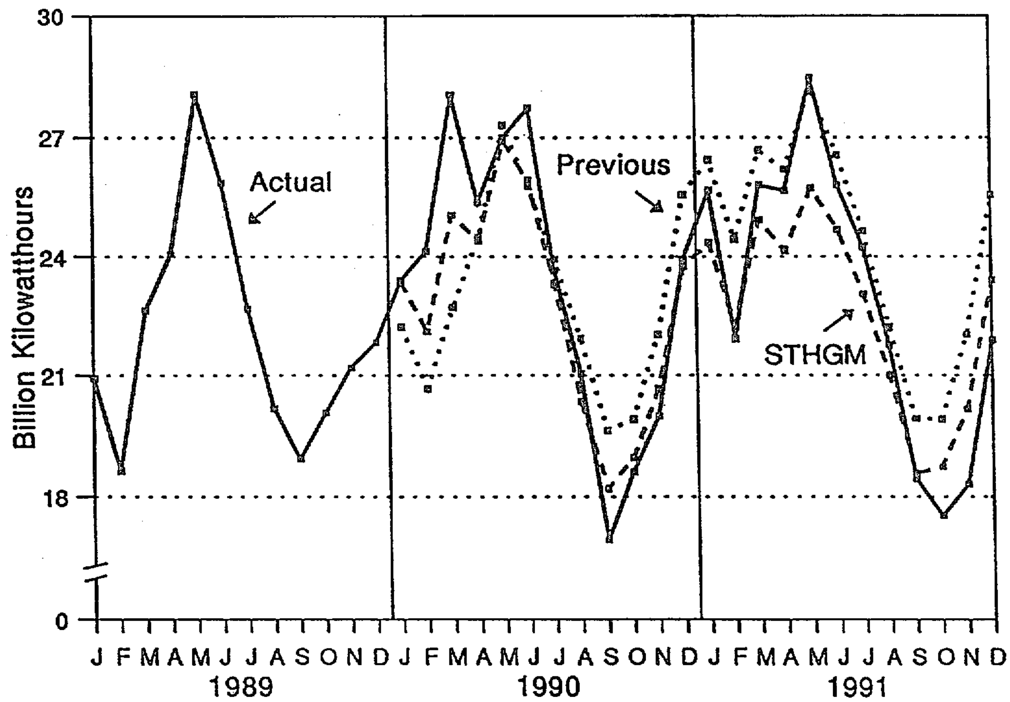
Evaluation of the model forecasts and forecasts from the previous methodology was performed using out-of-sample tests. That is, several different periods of actual data were deleted from the input data, the parameters in the model were reestimated, and forecasts made as described above for a 24-month period. These forecasts were compared to actual data and past forecasts using EIA's previous method obtained from memorandums produced when comparable historical data were available. Also, the Root Mean Square Errors (error) of the model estimates and the previous method were calculated and compared.

The first out-of-sample test was performed using historical data through 1989 (Figure 5). From the graph, the following were illustrated: (1) the model forecast was in general closer to the actual generation than the previous forecasts; and (2) the 1990 and 1991 actual data were close to the long run average hydroelectric generation and the model forecasts for these years was generally within 1 to 2 billion kilowatthours of the actual generation. The Root Mean Square Error (Table 1) of the forecasts using historical data through 1989 was 5 percent for 1990, the first year of the forecast, which was 45 percent less than the error for the previous method. For the 24 month forecast, the model had a 6 percent error, 38 percent less than the previous method.

The second out-of-sample test used historical data through 1991 (Figure 6). The graph illustrated the following points. First, for 1992, an unusually low year for hydroelectric generation, the forecast was not as good as for 1990 and 1991. This can be partially explained by the forecast assumption of normal precipitation while precipitation was below normal in 1992. Secondly, in 1992, with the below normal hydroelectric generation, the model still provided a closer forecast than the previous method. The Root Mean Square Error (Table 1) for 1992 for the model when using historical data through 1991 was 13 percent, 31 percent less than the previous method error. For both out-of-sample tests, the Root Mean Square Error in all cases was smaller for the model than for the forecasts from the previous methodology.

In conclusion, the Short-Term Hydroelectric Generation Model has been shown to improve upon EIA's prior procedure for projecting short-term hydroelectric generation. Of course, various uncertainties could affect the forecast, and, therefore, the validity of the forecasts rests on the legitimacy of the assumptions.

Figure 5
Comparison of the Forecast Methodologies with Actual Data, 1990 and 1991 (Forecasting with Actual Data through 1989)



Notes: • All forecasts are based on actual generation and precipitation data through the indicated year. • All forecasts assume normal precipitation during the forecast period.

Source: STHGM Forecasts: Energy Information Administration, STHGM run using HYDRO.TIME.SERIES.FINAL.D060193. Actual: Energy Information Administration, Form EIA-759, "Monthly Power Plant Report." Previous Forecasts: Energy Information Administration, Office of Coal, Nuclear, Electric, and Alternate Fuels memorandum to Office of Energy Markets and End Use dated March 13, 1990.

Table 1

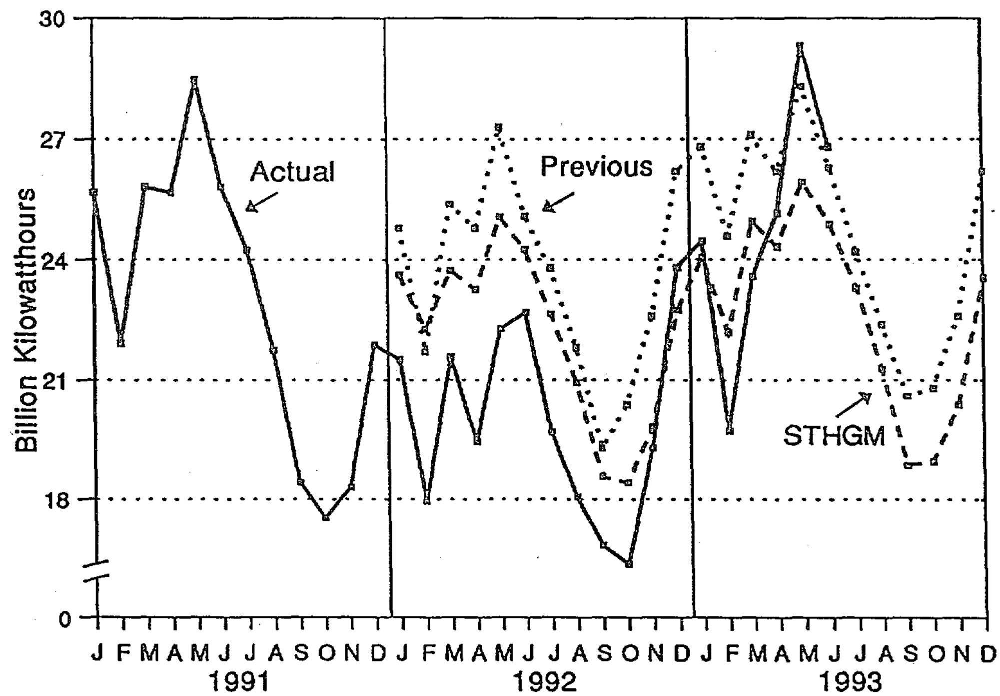
Comparison of Root Mean Square Errors

Using Historical Data through:	Year(s) of Forecast Included in the RMSE	Error (Million Kilowatthours)	
		STHGM	Previous
1989:	1990	1,267	2,289
		5.4%	9.8%
	1991	1,367	1,923
		6.0%	8.4%
	1990 and 1991	1,318	2,114
		5.7%	9.1%
1991:	1992	2,597	3,749
		13.0%	18.8%

Notes: • All forecasts are based on actual generation and precipitation data through the indicated year. • All forecasts assume normal precipitation during the forecast period. • Percents were calculated by dividing the error by the 12 (or 24) month average of the actual generation.

Source: STHGM Forecasts: Energy Information Administration, STHGM runs using HYDRO.TIME.SERIES.FINAL.D060193. Actual: Energy Information Administration, Form EIA-759, "Monthly Power Plant Report." Previous Forecasts: Energy Information Administration, Office of Coal, Nuclear, Electric, and Alternate Fuels memorandums to Office of Energy Markets and End Use dated March 13, 1990 and March 6, 1992.

Figure 6
Comparison of the Forecast Methodologies with Actual Data, 1992 and 1993 (Forecasting with Actual Data through 1991)



Notes: ● All forecasts are based on actual generation and precipitation data through the indicated year. ● All forecasts assume normal precipitation during the forecast period.
Source: STHGM Forecasts: Energy Information Administration, STHGM run using HYDRO.TIME.SERIES.FINAL.D060193. **Actual:** Energy Information Administration, Form EIA-759, "Monthly Power Plant Report." **Previous Forecasts:** Energy Information Administration, Office of Coal, Nuclear, Electric, and Alternate Fuels memorandum to Office of Energy Markets and End Use dated March 6, 1992.