Housing in the Macroeconomy

William Poole

I am very pleased to be here today to participate in this symposium sponsored by the Office of Federal Housing Enterprise Oversight. The topics are important, and the list of speakers impressive.

My purpose is to provide an overview of longer-run trends in housing and housing finance. The United States is well housed, and the housing finance system has been working efficiently in recent years. In the first two sections of my remarks, I'll discuss some of the history and report some measures showing how the housing stock has changed over time and how the housing finance system has developed. Our aim must be to sustain and extend this progress.

The third section of my remarks reflects my long-standing interest in issues of financial stability stemming from my study of monetary economics and financial history. Given the enormous importance of housing and housing finance to the U.S. economy, I think we do need to carefully examine the potential for financial instability and consider steps that could reduce the risk. In this context, I especially want to commend OFHEO for its recent report entitled, "Systemic Risk: Fannie Mae, Freddie Mac and the Role of OFHEO." This report displays an impressive depth of scholarship in reviewing a large body of professional literature on the subject. It deserves careful study by every economist interested in issues of financial stability and every policymaker with an interest in housing and housing finance.

Before proceeding, I want to emphasize that the views I express here are mine and do not necessarily reflect official positions of the Federal Reserve System. I thank my colleagues at the Federal Reserve Bank of St. Louis—especially Robert H. Rasche, senior vice president and director of research, and William R. Emmons, economist—for their assistance and comments, but I retain full responsibility for errors.

William Poole is the president of the Federal Reserve Bank of St. Louis.
This article was adapted from a speech of the same title presented at the Office of Federal Housing Enterprise Oversight Symposium, March 10, 2003. The author thanks colleagues at the Federal Reserve Bank of St. Louis for assistance and comments, especially Robert H. Rasche, director of research, and William R. Emmons, economist. The views expressed do not necessarily reflect official positions of the Federal Reserve System.

SOME FACTS ABOUT HOUSING

Housing, particularly owner-occupied housing, has long been a public policy issue in the United States. Over the years, these discussions developed in two different directions: one focusing on the availability of housing for lower-income families, which I will not address here, and the other on the development of housing in general and the efficiency of mortgage markets.

The discussion of policies toward housing and mortgage markets dates back to at least 1918.[1] During the Great Depression, the National Housing Act of 1934 created the Federal Housing Administration (FHA) with the mandate to insure private residential mortgages. In the aftermath of World War II, the Serviceman's Readjustment Act of 1944 created the Veterans Administration (VA) home-loan guarantee program.[2] Mortgages insured (or guaranteed) by the government gained considerable market share throughout the 1940s and 1950s, reaching a peak share of 44.3 percent of 1-4 family home mortgages in 1956. Since then, the share of government-insured mortgages has declined steadily; by the end of 2000 the share amounted to only 13.8 percent.[3]

The original Federal National Mortgage Association—Fannie Mae, as it came to be unofficially and affectionately called—was organized in February 1938 to increase the volume of residential construction and develop a secondary market in government-insured or -guaranteed mortgages.[4] To achieve the first objective, from its inception Fannie Mae purchased mortgages and issued its own debt. Initially,

[1] Harry S. Schwartz, "The Role of Government-Sponsored Intermediaries," *Housing and Monetary Policy*, Federal Reserve Bank of Boston, Conference Series 4, 1970, p. 68.

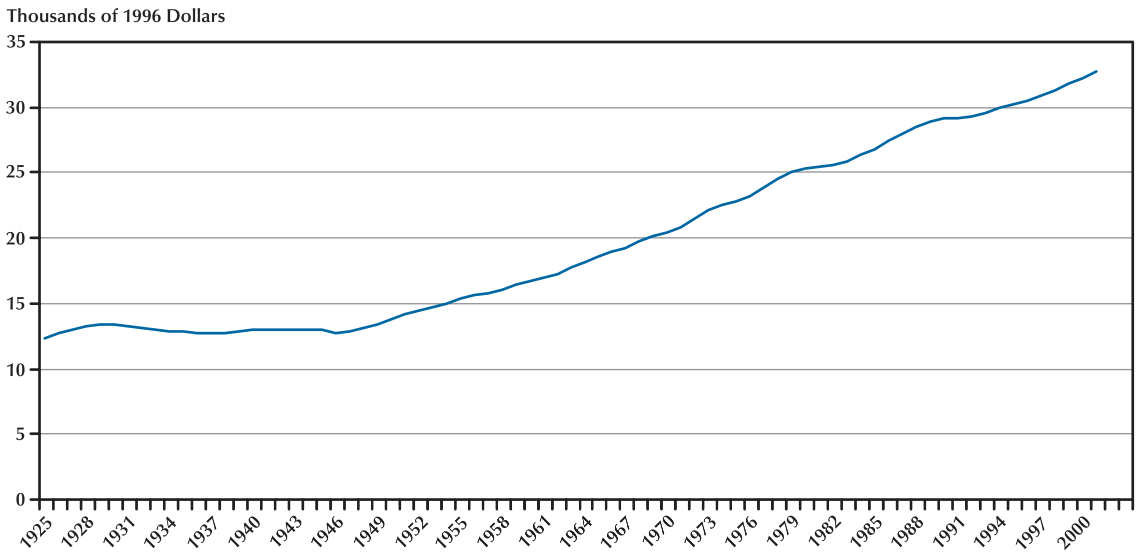
[2] George F. Break, "Federal Loan Insurance for Housing," *Housing and Monetary Policy, Federal Credit Agencies*, Commission on Money and Credit. Englewood Cliffs, NJ: Prentice-Hall, Inc. 1963, p. 2.

[3] Source: 1939-59: *Economic Report of the President*, February 1970, Table C-58; 1960-2000: *Economic Report of the President*, February 2002, Table B-75.

[4] Jack M. Guttentag, "The Federal National Mortgage Association," *Housing and Monetary Policy, Federal Credit Agencies*, Commission on Money and Credit, Englewood Cliffs, NJ: Prentice-Hall, Inc. 1963, p. 69.

Figure 1

U.S. Per Capita Real Net Residential Structures



Fannie Mae was funded through the sale of preferred stock to the Treasury. According to Jack M. Guttentag, writing in 1963, government support was regarded as transitory since it was "hoped that eventually the Treasury's investment can be retired with the proceeds of common stock along with retained earnings, and the function transferred to private ownership."[5] This objective was partially achieved in 1968 when the original Federal National Mortgage Corporation was split into two parts: Government National Mortgage Association, or Ginnie Mae, which remained a government agency, and a successor Fannie Mae (officially, still the Federal National Mortgage Association) that was spun off as a private corporation under a federal government charter. In 1970 Ginnie Mae started guaranteeing mortgage-backed pass-through securities representing shares in pools of FHA/VA guaranteed loans.[6] At the same time, the Federal Home Loan Mortgage Corporation—Freddie Mac—was created to promote the development of a secondary market in conventional mortgages.

Another important development in the 1930s was the creation in 1932 of the Federal Home Loan Bank System (FHLB), which was chartered to provide liquidity to thrift institutions. In 1934 the Federal Savings and Loan Insurance Corporation (FSLIC) was established to provide insurance on shares of depositors in thrift institutions.[7]

With these institutions in place, though not necessarily because of their creation, the net stock of real residential assets per capita began to grow after World War II. (See Figure 1.) The stock had been trendless between $12,500 and $13,000 1996 dollars from the mid-1920s until after World War II.[8] From 1948 to 1970 the net real per capita stock of residential structures grew at a 1.9 percent annual rate. From 1971 to 2001 the net stock grew at a somewhat lower average annual rate of 1.5 percent. By the end of 2001, the net per capita stock of real residential structures had grown to $32,700 1996 dollars.

As the stock of residential structures was growing, the quality of the housing stock was improving.

[5] Jack M. Guttentag, "The Federal National Mortgage Association," *Housing and Monetary Policy, Federal Credit Agencies*, Commission on Money and Credit. Englewood Cliffs, NJ: Prentice-Hall, Inc. 1963, p. 73.

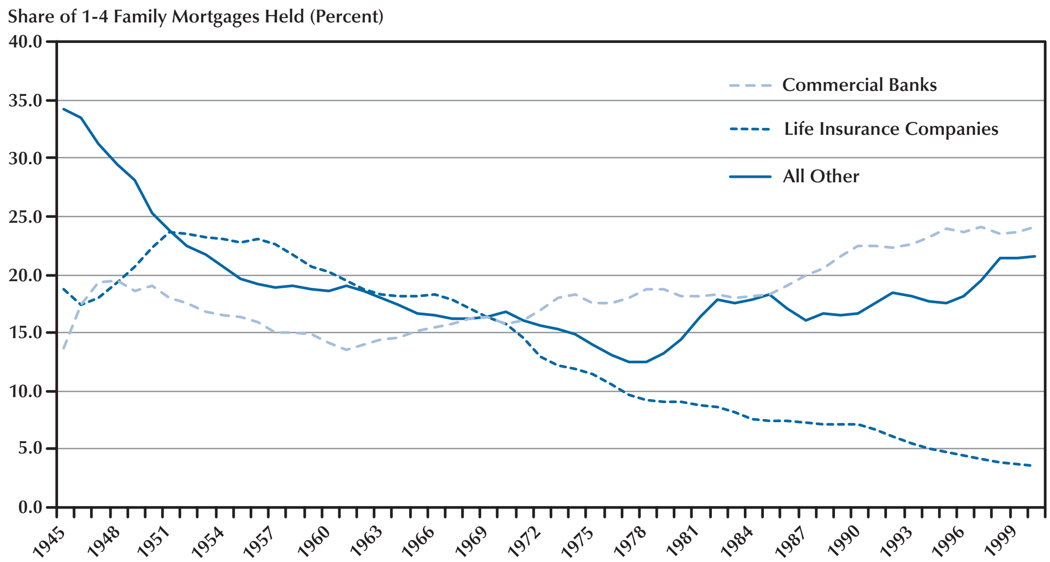
[6] P.H. Hendershott, "The Market for Home Mortgage Credit", in R.A. Gilbert, ed., *The Changing Market in Financial Services*, proceedings of the 15th Annual Economic Policy Conference of the Federal Reserve Bank of St. Louis. Norwell MA: Kluwer Academic Publishers, 1992, p. 100.

[7] Ernest Bloch, "The Federal Home Loan Bank System," *Housing and Monetary Policy, Federal Credit Agencies*, Commission on Money and Credit. Englewood Cliffs, NJ: Prentice-Hall, Inc. 1963, p. 168-72.

[8] Bureau of Economic Analysis, U.S. Department of Commerce, Chain-Type Quantity Indexes for Net Stock of Fixed Assets and Consumer Durable Goods. End-of-year quantity indexes are available from 1925-2001. The quantity indexes of net stocks of real residential and non-residential assets are converted into net stocks valued in 1996 dollars.

Figure 2

1-4 Family Mortgage Lending



According to the 1950 Census, 35.5 percent of houses lacked complete plumbing facilities. By 2000 the fraction of houses without complete plumbing had fallen to 0.6 percent. In the 1960 Census—the first census that included a question on telephones—21.5 percent of houses had no telephone. By 2000 only 2.4 percent of houses lacked a telephone. In the 1970 Census 4.4 percent of houses were recorded as lacking complete kitchen facilities. By 2000, only 1.3 percent of houses were recorded as without complete kitchen facilities. During this period the median size of houses also increased—from 4.6 rooms in 1950 to 5.3 rooms in 2000.[9]

As the quantity and quality of the residential housing stock increased, homeownership also became more widespread. In the 1950 Census the homeownership rate was reported at 55 percent—by the 2000 Census it had increased to 67.5 percent.

SOME FACTS ABOUT HOUSING FINANCE

Growth of the housing stock could not have occurred without a robust system of mortgage finance. There are several distinct sources of mortgage finance in the United States.[10] The importance of these sources has varied considerably over the years since World War II. The share of 1-4 family mortgage loans held by commercial banks increased in the immediate aftermath of World War II to a peak of 19.4 percent in 1948; it then trended down to 13.4 percent in 1961 at which point the trend reversed and the share trended up again, reaching almost 24 percent in 2000. (See Figure 2.) Life insurance companies were a significant player in the residential mortgage market immediately after World War II, but their share of lending peaked in 1951 at 23.5 percent and has trended down ever since. By 2000, the share of life insurance companies was only 3.4 percent, so these institutions have ceased to be a significant factor in the residential mortgage market. The share of "all other," which includes lending by individuals and private mortgage pools decreased from 34.1 percent at the end of World War II to 12.3 percent in 1977, after which it started trending up and reached 21.4 percent by 2000.

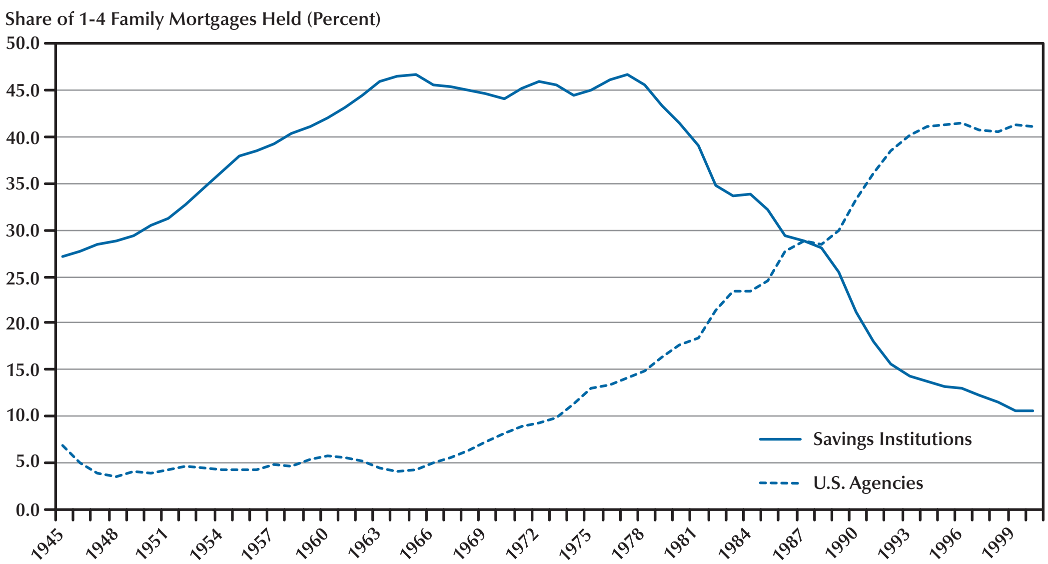
The two remaining types of institutions that at different times have been the most significant players in the residential mortgage lending market are savings institutions (including savings and loan associations and mutual savings banks) and U.S. agencies including Ginnie Mae, Fannie Mae, Freddie Mac, and mortgage pass-through securities guaranteed by federal agencies or government sponsored enterprises (GSEs). The share of savings institutions in

[9] Rooms exclude bathrooms.

[10] Source: 1939-59: *Economic Report of the President*, February 1970, Table C-59; 1960 to 2000: *Economic Report of the President*, February 2002, Table B-76.

Figure 3

1-4 Family Mortgage Lending



residential mortgage lending grew rapidly after World War II, reaching 46 percent in 1965. (See Figure 3.) These institutions maintained their market share until 1978, but then lost share dramatically.

The decline of the savings institutions was a consequence of rising nominal interest rates combined with duration mismatch, which together generated the savings and loan crisis of the 1980s. By 1990, when the S&L crisis was finally resolved, the share in the residential mortgage market of these institutions had shrunk to 21.1 percent, less than half of the peak market share twenty-five years earlier. In the subsequent decade the market share held by these institutions shrunk by half again, to only 10.4 percent at the end of 2000.

As the presence of savings institutions in the residential mortgage market receded, the financing void was filled by U.S. government agencies. In 1967, immediately before the Housing Act of 1968 and reorganization of the established Fannie Mae into Ginnie Mae and the new Fannie Mae, the share of the residential housing mortgage market for government agencies was 5.5 percent. By 1990, these institutions captured a third of the residential mortgage market, either through mortgages purchased for their own portfolios or through guaranteed mortgage-backed securities. Recent data indicate that their market share is 42.5 percent as of the end of the third quarter of 2002. Clearly, the efficiency and stability of these government agencies has become a critical factor in the financing of residential construction.

FINANCIAL STABILITY

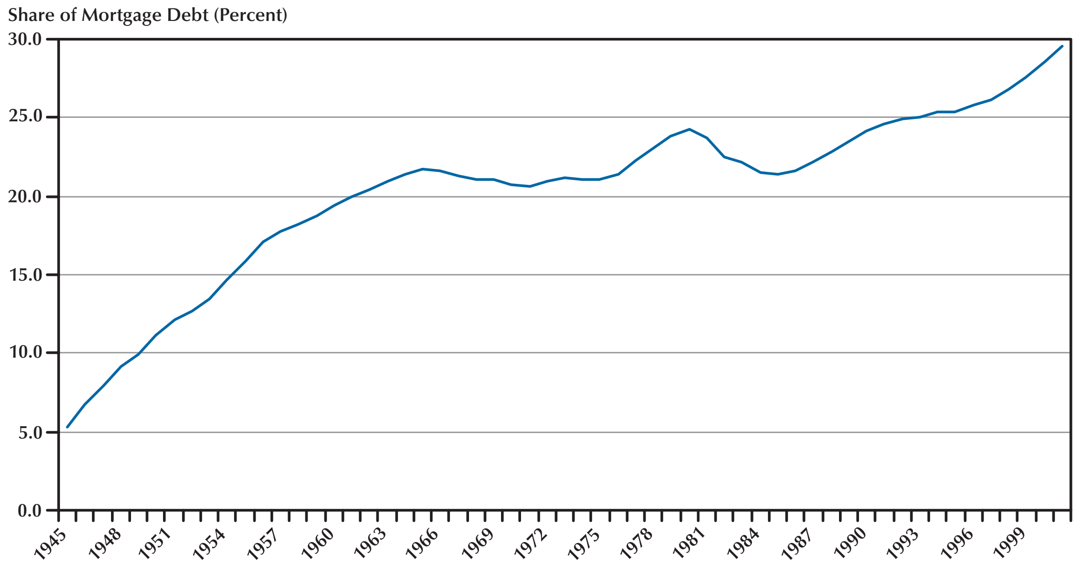
As Figure 4 shows, residential mortgage debt has grown enormously as a fraction of total nonfinancial debt in the United States. Starting at slightly more than 5 percent at the end of World War II, the share grew steadily until it exceeded 20 percent in the early 1960s. From then until the mid-1980s, the share fluctuated in the neighborhood of 20 percent or a bit more. In the past 15 years the share again grew steadily until it reached 30 percent at the end of 2001.[11] Given the current magnitude of mortgage debt outstanding relative to total credit market debt, any serious instability in the financing of the residential capital stock has the potential for significant effects not only on the housing industry but also on the entire economy.

The annual reports of Fannie Mae and Freddie Mac, and the recent OFHEO report on Systemic Risk, provide much useful information on risk management. It is insightful to divide this subject into two parts. One concerns management of credit, interest-rate, and operational risks that can be modeled with the assistance of financial theory and evidence from the behavior of financial markets. Risks that can

[11] Board of Governors of the Federal Reserve System, *Flow of Funds Accounts*, Table L.2, Credit Market Debt Owed by Nonfinancial Sectors.

Figure 4

Home Mortgage Debt as Share of Total Nonfinancial Debt



be studied and modeled can be termed "quantifiable risks." Nonquantifiable risks deserve separate attention.

There are certainly cases in which firms, and sometimes regulators, make mistakes in dealing with quantifiable risks. Over the years, many financial institutions have failed because of such mistakes. Savings and loan association failures, which ultimately led to the failure of the Federal Savings and Loan Insurance Corporation (FSLIC), were mostly of this type. Starting in the late 1960s, economists warned for years that the extreme maturity mismatch from S&L balance sheets with long-term, fixed-rate mortgages financed through short-term liabilities put the industry at great risk. As those risks were realized, many firms failed and the S&L industry declined to a shadow of its former self. The cost to taxpayers to make good on the insurance guarantee offered by FSLIC was in the neighborhood of $150 billion. As a consequence of this experience, managers of firms, regulators, and those active in financial markets are today well aware of the need for careful risk management.

The OFHEO report makes an extremely important point about nonquantifiable risks:

> A further obstacle to quantifying systemic risk is the inherent difficulty in using quantitative techniques to analyze catastrophic events such as wars and financial crises. Such events are rare, often involve significant departures from recent historical experience and can develop from a potentially infinite set of conditions. Analysts generally do not model, simulate, or predict the course and consequences of unconditional financial crises, making it difficult to obtain a precise estimate of the likelihood of a specific level of economic losses resulting from potential financial crises. As a result, government officials who seek to plan for such events cannot rely on the usual quantitative techniques to evaluate alternative strategies for addressing them. (p. 87)

In a previous speech I suggested that periods of great market instability arise when three conditions are met. First, something happens that has widespread significance—is large enough to matter to lots of people. Second, the triggering event is a surprise. Ordinarily, events long anticipated are not troublesome because corrective action occurs before problems arise. Third, substantial uncertainty clouds resolution of the problem. It is especially difficult for investors to know what to do when the government's response to an unfolding situation is highly uncertain.[12]

[12] "Financial Stability" presented before The Council of State Governments Southern Legislative Conference Annual Meeting, New Orleans, Louisiana, Aug. 4, 2002: < www.stlouisfed.org/news/speeches.html > .

Given the extensive discussion of quantifiable risks, I want to concentrate on the nonquantifiable risks. It helps to make this issue concrete by listing some examples. The failure or near failure of Penn-Central, Continental-Illinois, Long-Term Capital Management, Enron, and WorldCom may not have been complete surprises to knowledgeable insiders, but the shocks were certainly "news" to market participants, regulators, and the general public. No one predicted the timing of the stock market crash of 1987 or the peak of the equity markets in the spring of 2000. It is well known that even the great Yale economist Irving Fisher was caught completely off guard by the crash of 1929. Surprise legal decisions brought bankruptcy to 52 firms involved with asbestos, to Dow-Corning, and to Texaco. Finally, while experts in terrorism may have understood the risks of attacks on U.S. soil, their information was not sufficient to prevent the September 11 attacks; certainly no one else had any basis for predicting the attacks. All of these cases, with the possible exception of Continental-Illinois, reflected nonquantifiable risks.

The point here is not to fault the forecasting record of any person or any agency. Rather, it is to illustrate that major unforeseen events that can bring about a collapse in confidence or disruption to the normal function of financial markets without any warning can and do occur with some frequency. The history of the United States, as well as other countries, is replete with such examples.

A little-discussed but critically important dimension of systemic risk is the uncertainty about how the government and regulators will respond to a major unforeseen event.[13] Before the 1987 stock market crash there was considerable overconfidence that a break in equity prices such as occurred in 1929 was not possible given modern institutions. As a result, in the initial hours of the 1987 crash, the public did not know exactly how the Fed would react to a systemic liquidity crisis. The way the Fed handled that situation is, in my judgment, one of the high-water marks in the history of our central bank. Not only was a generalized liquidity crisis averted, but also considerable institutional credibility was created. The repercussions in financial markets on 9/11 might have been much worse had the Fed not demonstrated in 1987 that it could and would react immediately to major market disruptions.

There are historical cases where the reactions by government agencies and regulators to unpredicted crises, in my judgment, did not result in such institution building. A good example is the market perception that public policy has established a "too big to fail" doctrine. This perception grew over time and became more entrenched as a result of the Continental-Illinois situation. The net result is that market participants expect that, under ill-defined conditions, regulators and/or government agencies will in fact insure statutorily uninsured positions involving large financial institutions. Is the doctrine really "too big to fail" or "too big to liquidate quickly?" How big does a financial institution have to be, and does it have to be a depository institution, to be "too big to fail?" In this respect, there is tremendous ambiguity about the status of the GSEs. The market prices the GSEs' debt as if there were a federal guarantee, or a high probability of a guarantee, standing behind their entire outstanding obligations. Yet, there is no explicit guarantee in the law. Actual experience has left the markets with all of these important questions and ambiguities.

No one should underestimate the potential importance of the ambiguity over the financial status of the GSEs. Would "too big to fail" be extended to GSEs in a crisis, and if so how would it be effected in the absence of a federal insurance agency with an unlimited line of credit? How quickly could such a rescue be implemented?

It is not sufficient for any single GSE to argue that its own financial condition is sound. If one GSE comes under a cloud, others may also. That has been our experience with financial firms again and again. It is the process economists call "contagion," whereby uninvolved or innocent firms are affected because the market has difficulty distinguishing solid firms from those at risk.

In the case of the GSEs, the enormous scale of their liabilities could create a massive problem in the credit markets. If the market value of GSE debt were to fall sharply, because of ambiguity about the financial soundness of GSEs and about the willingness of the federal government to backstop the debt, what would happen? I do not know, and neither does anyone else.

Let me throw out for debate two steps the federal government might take to resolve the ambiguity that I see as a fundamental risk to the continuing stability of our financial system. First, various aspects of federal sponsorship that the market reads as providing an implied guarantee of GSE debt should

[13] I discussed this issue at some length in "Expectations," Federal Reserve Bank of St. Louis *Review*, March/April 2001, *83*(2), pp. 1-10.

be withdrawn.[14] The Secretary of the Treasury has the authority to buy GSE obligations; in the case of Fannie Mae and Freddie Mac, the authority is up to a maximum of $2.25 billion for each firm. The GSEs could easily replace this potential source of emergency financial support with credit lines at commercial banks, following the widespread practice among issuers of commercial paper. In any event, the amount available at the discretion of the Secretary of the Treasury is far too small to deal with a crisis in the GSE debt market. Eliminating the Treasury's authority to lend to the GSEs would provide a signal that the government is serious when it says that there is no government guarantee of GSE debt.

Second, over a transitional period of several years, the GSEs should add to the amount of capital they hold. Capital is critical because when there is a crisis in the securities markets, financially strong firms can stand the pressure without lasting damage. Capital provides a cushion against mistakes and unforeseeable circumstances. With adequate capital, a firm can almost always raise emergency loans to cover its liquidity problems. The importance of adequate capital became clear to policymakers as the S&L problems accumulated in the late 1980s. Tightening of capital standards for insured depository institutions and the administration of those requirements was a key part of the reforms put in place at that time.

Capital is especially important for the GSEs because their short-term obligations are large. Fannie Mae and Freddie Mac have debt obligations due within one year of about 45 percent of their debt liabilities. Any problem in the capital markets affecting these firms could become very large, very quickly. It is important to understand what "very quickly" means. Because of the scale of the short-term obligations of the GSEs, the GSEs are rolling over many billions of dollars of obligations each week. For this reason, a market crisis could become acute in a matter of days, or even hours.

Capital on the books of Fannie Mae and Freddie Mac is well below the levels required of regulated depository institutions. Let me quote a paragraph from the 2001 Annual Report of Fannie Mae, the largest single GSE. During 2001, Fannie Mae issued $5 billion of subordinated debt that received a rating of AA from Standard & Poor's and Aa2 from Moody's Investors Service.

> Fannie Mae's subordinated debt serves as a supplement to Fannie Mae's equity capital, although it is not a component of core capital. It provides a risk-absorbing layer to supplement core capital for the benefit of senior debt holders and serves as a consistent and early market signal of credit risk for investors. By the end of 2003, Fannie Mae intends to issue sufficient subordinated debt to bring the sum of total capital and outstanding subordinated debt to at least 4 percent of on-balance sheet assets, after providing adequate capital to support off-balance sheet MBS. Total capital and outstanding subordinated debt represented 3.4 percent of on-balance sheet assets at December 31, 2001. (pp. 44-5)

The capital situation at Freddie Mac is about the same as the one at Fannie Mae. The capital adequacy standards applying to these two GSEs were established by the Federal Housing Enterprises Financial Safety and Soundness Act of 1992. The core capital requirement is 2.5 percent of on-balance sheet assets and 0.45 percent of outstanding mortgage-backed securities and other off-balance sheet obligations. The off-balance sheet obligations have a capital requirement because they are guaranteed by Fannie and Freddie.

In the private sector, government securities dealers carry capital in the neighborhood of 5 percent, and other financial firms considerably more. For example, FDIC-insured commercial banks hold equity capital and subordinated debt of a bit under 11 percent of total assets.

The issue with Fannie Mae and Freddie Mac is not primarily one of disclosure. Their annual reports disclose quite well the high degree of complexity of their operations, and the small amount of capital they carry above what is required by law. My questions are these: Given the complexity of their operations, is the capital standard in the law adequate? Why is the standard so far below that required of federally regulated banks? What will happen to the housing market if Fannie and Freddie become unstable?

Reports issued by Fannie Mae and Freddie Mac, and the recent OFHEO report on Systemic Risk,

[14] Farmer Mac, another GSE, was much in the news in the recent past. An article in the *New York Times* noted that one of the advantages conferred by government sponsorship is "the ability to borrow almost as cheaply as the government does because of a perception of government backing that emanates from a single section in its charter. That provision allows the Treasury, in certain circumstances, to provide up to $1.5 billion in loans to Farmer Mac to support the guarantees the company extends on farm loans" (9 June 2002, p. 8, column 1).

indicate that the two firms employ state-of-the-art risk management. Nevertheless, my sense is that the firms are vulnerable to nonquantifiable risks, because their capital positions are so low.

In my judgment, the only way for financial institutions to ensure stability in the event of nonquantifiable shocks is for them to maintain a substantial extra capital cushion above that deemed necessary by analysis of quantifiable risks. One way of thinking about the appropriate size of that cushion might be to decide that a firm should be able to meet its maturing obligations without borrowing for a certain period of time. The length of the period would depend on an assessment of how long it would take to resolve whatever problem might arise. Under this criterion, the capital cushion would have to be invested in highly liquid, short-term assets not subject to depreciation due to interest rate changes or credit risks, so that maturing obligations could be met for a time without resort to issuing new obligations.

Dismissing the risks of nonquantifiable events on the grounds that they are too improbable to worry about is not a wise approach to public policy. For one thing, these events are not so rare as they might seem. For another, the costs of a rare event that has major consequences to the economy can easily outweigh a long stream of benefits that are orders of magnitude smaller.

SUMMING UP

The United States has enjoyed many years of a rising stock of residential capital. Moreover, dwellings have increased in average size and quality. The nation's housing finance system has been effective in making this growth possible.

The housing finance system historically has been highly diversified. As a group, the share of savings institutions in residential mortgage lending reached 46 percent in 1965, but hundreds of institutions were involved. The diversification of lending by different types of institutions and numerous firms within a class of institutions has been an important element of stability, because the failure of one or even many firms has not shaken the system. Competing firms have been able to enter the market to fill any voids left by failing firms.

Today, the housing finance system is heavily concentrated. Just three firms—Fannie Mae, Freddie Mac, and Ginnie Mae—account for over 40 percent of the residential mortgage market. Ginnie Mae is backed by the full faith and credit of the U.S. government. Fannie Mae and Freddie Mac are not so backed and hold capital far below that required of regulated banking institutions. Should either firm be rocked by a mistake or by an unforecastable shock, in the absence of robust contingency arrangements, the result could be a crisis in U.S. financial markets that would inflict considerable damage on the housing industry and the U.S. economy.

Congestion at Airports: The Economics of Airport Expansions

Jeffrey P. Cohen and Cletus C. Coughlin

Congestion has been and continues to be a problem at many airports throughout the United States. For example, in the first five months of 2001, over 25 percent of the flights arriving at the nation's 11 busiest airports were more than 15 minutes late.[1] Despite a decline in travelers and flights in 2001, which was associated with the recession that began that spring and the September 11 terrorist attacks, congestion remained a problem in some locations.[2] For example, 16.2 percent of the flights bound for Lambert–St. Louis International Airport from May 1, 2001, through June 30, 2001, arrived late, with an average delay of roughly 55 minutes. Using the same period one year later, 16.3 percent of the arriving flights were delayed, with an average delay time of roughly 56 minutes.[3]

Congestion imposes costs on both the users and providers of airline transportation services. A common response is to expand the capacity of airports in the most afflicted regions. Consequently, airport expansions have occurred and are occurring in many major cities, including Atlanta and St. Louis.[4] Figure 1 shows that the amount of federal, state, and local government spending on airports increased in all but two years between 1986 and 1999.[5] Federal, state, and local funds for U.S. airports in 1999 totaled over $20 billion, up from $11 billion in 1985 (using constant, 1996 dollars).

Expansions are costly, complex, and controversial. For example, the cost of "Phase 1" of the current expansion of Lambert–St. Louis International Airport is $1.1 billion. The key component of this project is the construction of a new runway.[6] To add this runway, the approved project entailed the acquisition of more than 1,500 acres of land, which ignited protests from affected homeowners and businesses; the reconfiguration of seven major roads; the movement of some airport support operations and the Missouri Air National Guard facility; and the construction of a new school.[7]

We begin our analysis by providing a discussion of how congestion arises and how it can be dealt with. Because the air transportation services provided by one airport are related to the services provided by many airports, delays at one airport have adverse effects on the movement of passengers and freight at other airports.[8] Thus, the expansion of one airport can assist the movement of passengers and freight at other airports. This interdependence provides an economic justification for a decisionmaking authority above the level of individual airports, such as a governmental body, to be involved in the approval as well as the financing of expansions. However, when both congestion and network externalities are present, the appropriate government actions may be to levy a tax, to provide a subsidy, or possibly to refrain from any intervention.

To justify a specific airport expansion, its benefits must exceed its costs. We examine how the benefits and costs of expansions are measured. We use the expansion of Lambert–St. Louis International Airport to illustrate many of the key points.

---

1 See Salant (2001).

2 The terrorist attacks also spurred increased screening of passengers and luggage, which generated other forms of congestion. See Coughlin, Cohen, and Khan (2002) for a discussion of aviation security and terrorism.

3 Data on average flight delays were found at <http://www.bts.gov/ntda/oai/SummaryStatistics?DEMO/SummaryStatistics>.

4 However, the current financial problems of airlines and airports have, at least temporarily, led to either the cancellation or delay of an estimated $16 billion in capital projects at numerous airports, including ones in Los Angeles, Boston, and Phoenix. See Bayles (2001).

5 Bureau of Transportation Statistics web site: <http://www.bts.gov/transtu/govfin/2001/tables/table_1b.html>.

6 Lambert International Airport web site: <http://www.lambert-pmo.org/about/phase1/default.asp>.

7 See Gilbert (2002) for details.

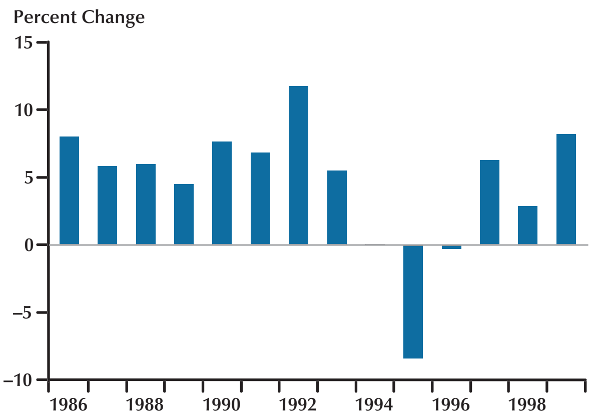
8 The relationships among airports create a network. One feature of this network is the "hub and spoke" system. Flights from various remote airports (the nodes on the spokes) converge on one airport (the hub). Flight schedules provide some time for passengers to change planes, before they depart for their final destinations.

Jeffrey P. Cohen is an assistant professor of economics at the Barney School of Business, University of Hartford, and Cletus C. Coughlin is deputy director of research at the Federal Reserve Bank of St. Louis.
The authors thank Chuck Reitter for useful information pertaining to the Lambert Airport expansion. Heidi L. Beyer and Molly J. Dunn provided research assistance.

Figure 1

Percent Change in Federal, State, and Local Spending on Airports

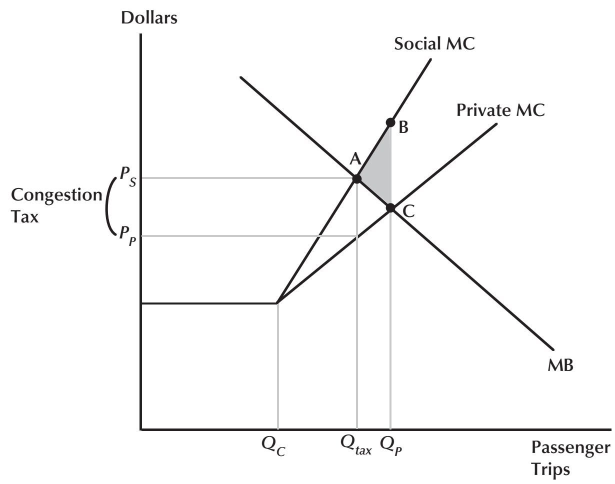
Chained 1996 Dollars, 1986-99



SOURCE: Bureau of Transportation Statistics web site.

Figure 2

Optimal Congestion Tax



We also examine two controversial aspects of expansions—the displacement of people and the environmental effects. The controversy as well as the cost of expansion projects has spurred the search for alternative ways to reduce congestion. One alternative that we examine, which reduces congestion by using existing capacity more efficiently, is congestion-based pricing of landing fees.

ANALYZING CONGESTION IN THE AIR TRANSPORTATION NETWORK

Expanding the capacity of an airport entails a multi-year capital expansion project to construct a runway and/or a terminal. The financing of expansions generally includes funding provided by a governmental body. An alternative in some cases to increasing an airport's infrastructure is to use its existing facilities more efficiently. This alternative approach to reduce the adverse effects of congestion can be implemented in the short run via the setting of appropriate prices or taxes.

The potential role of taxes/subsidies in responding to congestion as well as network externalities can be explained using marginal benefit/marginal cost curves.[9] A potential traveler who wishes to fly from St. Louis to Boston, for example, faces two costs—the airfare ("x" dollars per trip) plus the opportunity cost of travel time ("y" dollars per hour). In Figure 2, with the units of measurement being dollars on the vertical axis and passenger trips on the horizontal axis, the marginal benefit (or demand) curve has a negative slope. This curve indicates, for our example, that when the price of air travel from St. Louis to Boston is high, only the few people who value their trips highly would choose to travel. As the price of a trip is reduced, more travelers will take this trip.

The explanation for the shapes of the private and social marginal cost curves is slightly more complicated.[10] The shapes hinge on the impact of congestion costs. At air traffic levels less than Q_C in Figure 2, the marginal cost curve is flat. The flatness indicates that congestion has not set in yet, and the cost for passengers is the monetary cost of their tickets plus a "fixed" value related to the opportunity cost of travel.[11] Consequently, the "social" marginal

[9] Initially, we assume airport capacity is fixed. Later, we relax this assumption. In the long run, airport authorities can choose both the size of an airport as well as the congestion tax that generates the optimal traffic volume. For additional details on congestion taxes and traffic volume for highways, see O'Sullivan (2003, Chap. 11).

[10] Our exposition of the marginal cost curve is similar to O'Sullivan's (2003) analysis for highways. Neither exposition attempts to directly link the marginal cost curve with a supply curve; however, the monetary cost of airfare is related to a supply curve. Instead, we rely upon the principle that the socially optimal number of passenger trips occurs at a quantity where the marginal benefits curve intersects the social marginal cost curve.

[11] With no time delays stemming from congestion, travel time is fixed. Thus, the total value of travel time is also fixed.

cost curve coincides with the "private" marginal cost curve.

For air traffic levels greater than Q_C, however, the marginal cost curve is upward sloping. The positively sloped part of the private marginal cost curve can be understood by considering the notion that, beyond some threshold level, when a particular individual chooses to consume additional air travel, he adds additional traffic to the system, which slows down his own travel. Additional passenger trips translate into more flights, which is one source of congestion. In addition, more passengers cause more crowded airport terminals, creating delays at ticket counters and security checkpoints. When passengers anticipate these delays, they arrive at the airport earlier and increase their travel time. Moreover, additional passenger trips result in increased delays at the baggage claim carousel at the end of a trip. Thus, we are assuming that trip time varies directly with the number of passenger trips. This longer trip time increases the marginal passenger's own trip cost, due to the higher opportunity cost of the travel time, which is added to the monetary cost.

When accounting for congestion in the manner described above, the social marginal cost curve differs from the private marginal cost curve. Namely, when an additional passenger uses the airport more, in addition to increasing his own travel time by adding to congestion, this passenger increases the travel time of other airport users as well. This passenger does not take this additional cost into account, so existing travelers could be made better off if the would-be passenger did not travel. This additional cost causes the social marginal cost curve to lie above the private marginal cost curve.

This increase in total travel time can be calculated by multiplying the total number of airport users by the additional travel delay that the marginal user generates. At levels of passenger trips slightly greater than the congestion threshold, Q_C, this travel congestion externality is small, so the difference between the social and private marginal cost curves is relatively small. But as the number of passenger trips rises, an individual increasing his air travel adds greater cost to other travelers, due to the assumption that trip time is an increasing function of passenger volume. As a result, the difference between the social and private marginal cost curves is larger at higher levels of passenger traffic.

The socially optimal level of air traffic in this context occurs at Q_{tax} in Figure 2, where the marginal benefit curve intersects the social marginal cost curve. Note that this level of passenger trips is less than Q_P, which is the level of passenger trips associated with the intersection of the marginal benefit and private marginal cost curves.

One way to achieve the socially optimal level of passenger trips, which requires the marginal benefits of air travel to equal the social marginal costs, is through a congestion tax on air travel from St. Louis to Boston. In Figure 2, a tax per unit of air traffic volume equal to the distance from P_P to P_S will yield this socially optimal level of passenger trips. The cost associated with the tax forces each traveler to pay for the costs that his travel imposes on others. The net gain resulting from the tax is represented by the triangle *ABC*. The net gain reflects the fact that for a level of passenger trips Q_P, marginal social costs exceeds marginal benefits by the distance from *B* to *C*. The tax causes passenger trips to decrease from Q_P to Q_{tax} and eliminates the gap between marginal social costs and marginal benefits. Thus, the net gain is the sum of the differences between marginal social costs and marginal benefits as passenger trips decrease from Q_P to Q_{tax}.

A congestion tax, however, is not the only policy option. An alternative policy, shown in Figure 3, is to expand the airport at which the congestion is present. Airport expansion shifts the private and social marginal cost curves rightward. In addition, the congestion "threshold" occurs at a larger number of passenger trips, increasing from Q_T to Q_{TX}. Whether congestion persists depends not only on the size of the expansion, but also on the elasticity of air travel demand.[12] If travel demand is relatively inelastic, then it is likely that the expansion leads to a new equilibrium level of passenger trips, Q_I, below the congestion threshold. In this hypothetical example, the relatively inelastic demand implies that consumers of air travel are not very responsive to the lower cost of travel, so quantity demanded does not increase by enough to create additional congestion. If demand is relatively elastic, however, then it is possible that the new equilibrium, Q_E, will occur above the congestion threshold, and the social marginal cost will be greater than the private marginal cost. Because consumers are relatively responsive to changes in the trip cost, the expansion

[12] If a given percentage change in trip cost, say a 1 percent decline, leads to a relatively small change in air traffic volume, say a less than 1 percent increase, then the price elasticity of air traffic demand is inelastic. If, however, the 1 percent decline in trip cost is associated with a more than 1 percent increase in air traffic volume, the price elasticity of air traffic demand is elastic.

Figure 3

Social and Private Marginal Cost Before and After an Airport Expansion

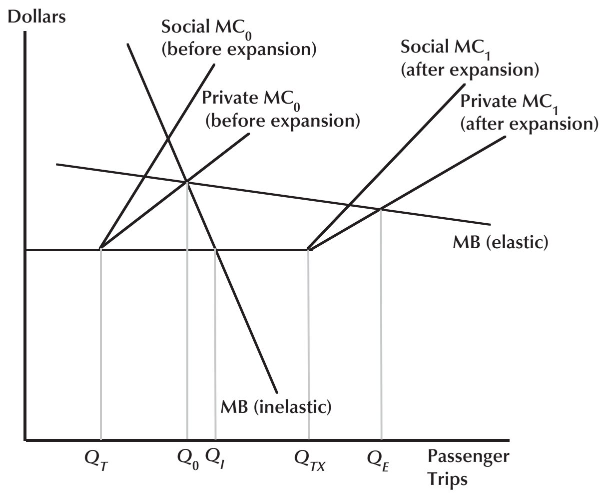
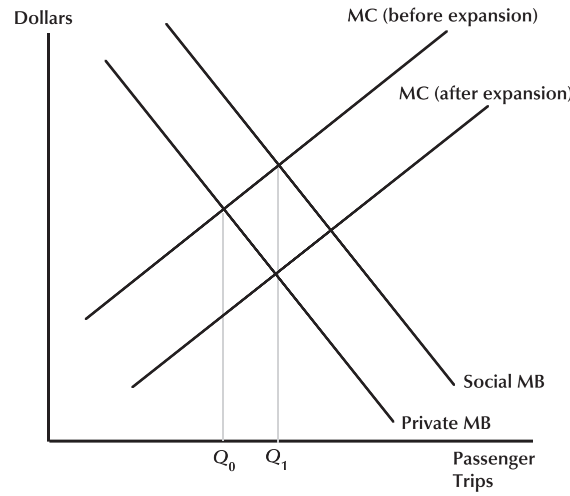


Figure 4

Airport Expansion to Account for Network Externalities



leads to a relatively large increase in the quantity of passenger trips demanded; thus, congestion persists.[13]

The preceding discussion ignores the crucial issue of whether a specific expansion generates benefits greater than its costs. In the example, depending on the elasticity of air traffic demand, travelers benefit by avoiding some or all congestion costs. The reduced congestion costs also induce more travel, which provides additional benefits. Reduced congestion also provides benefits for the providers of air transportation services. For example, airlines are likely to benefit in the form of reduced operating costs. Later, we examine how to calculate and compare the value of the benefits with the costs of an expansion.

The preceding discussion also focuses on congestion at a single airport and, as a result, ignores a key feature of the air transportation system. Accounting for the network structure of the hub and spoke system, while temporarily ignoring congestion issues, leads to a somewhat different picture concerning airport expansions. If the local airport planning authority makes a decision on the desired level of passenger capacity by equating the local (or private) marginal benefits to the marginal costs, then it is possible that there would be a suboptimal level of capacity. Namely, the hub and spoke network structure of the air transportation system implies that individuals who use a particular hub airport as a stopover point receive benefits in the form of reduced travel time from use of the hub airport in one leg of their travel.

In the specific case of St. Louis, if Lambert Airport does not expand, then it will cause congestion and slowdowns throughout the country, even on some flights that do not go through St. Louis. Thus, if a local authority chooses the level of airport capacity by considering only the local benefits, then equating the local marginal benefits to the marginal costs would lead to Q_O passenger trips (in Figure 4) and a passenger trip capacity that is too small. The reason for such a conclusion is that the social marginal benefits of air traffic capacity are likely to be greater than the local marginal benefits.

The socially optimal level of air capacity provision, however, occurs where the social marginal benefit curve intersects the marginal cost curve (before expansion). In Figure 4 this intersection yields a passenger trip volume of Q_1. One possibility for attaining point Q_1 is further expansion of local airports. This expansion, which lowers the marginal cost of passenger trips by shifting the marginal cost curve to the right, can be accomplished through subsidies (or intergovernmental grants) to the local

[13] According to O'Sullivan (2003), a similar phenomenon of "latent demand" occurs for highway expansions, contributing to the continuation of congestion following the expansion.

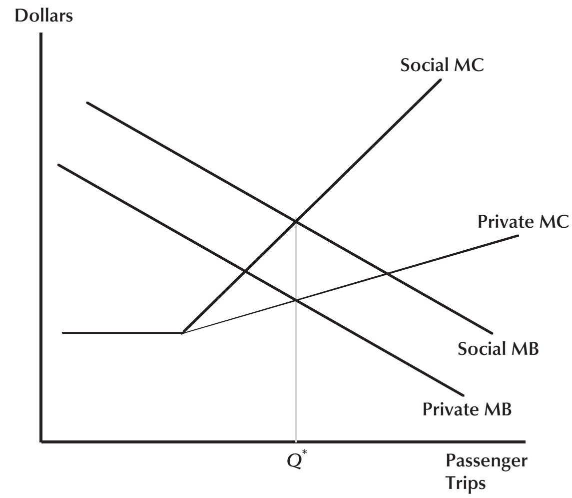
airport authorities from the federal government.[14] The resulting equilibrium of passenger trips occurs at the socially optimal Q_1, where the new marginal cost curve intersects the local (private) marginal benefit curve.

If we consider both types of spillovers simultaneously—the congestion externality and network externality—we obtain an ambiguous policy recommendation. Namely, in Figure 5, the social marginal cost curve lies above the private marginal cost curve because of the congestion externality, and the social marginal benefit curve lies above the private marginal benefit curve because of the network externality. The socially optimal level of airport traffic, Q^*, occurs at the intersection of the social marginal benefit curve and the social marginal cost curve. In general, this occurrence may be at a passenger trip level that is greater than, less than, or equal to the level that would be attained privately.

In Figure 5, the social marginal benefit and social marginal cost curves are drawn so that government intervention would not be necessary. In this case, the congestion effect, which can be mitigated by a tax, is exactly offset by the network effect, which can be accommodated by a subsidy. If the congestion effect were relatively more pronounced than the network effect, then governmental intervention by means of a tax would be justified in theory. On the other hand, if the network effect were relatively more pronounced than the congestion effect, then governmental intervention by means of a subsidy would be justified in theory. Thus, when accounting for both types of externalities, the policy prescription is ambiguous unless one knows the sizes of the effects. To accurately determine the amount of the tax or subsidy to address both externalities simultaneously, one must know the sizes of both the network and congestion effects.

Figure 5

Congestion and Network Externalities



MEASURING THE BENEFITS AND COSTS OF AIRPORTS AND AIRPORT EXPANSIONS

Nearly all studies attempting to measure the benefits and costs of airports and airport expansions are encompassed by the following categories of analysis:

- input-output models,
- production function/growth studies, or
- benefit-cost analyses.[15]

For the analysis of individual expansion projects, a benefit-cost analysis is a crucial component for determining whether an expansion is justified. We will examine this approach after highlighting some results from the two other approaches that are used to analyze the economic impact of airports.

Input-Output Models and Local Economic Impacts of Airports and Airport Expansions

A standard tool for calculating the economic impact of airports in general, as well as the economic impact of a specific airport, is an input-output model.[16] The impact of an expansion can be estimated as well. What this tool cannot provide, how-

[14] The suggestion that subsidies can lead to the social optimum ignores the practical problem of financing the subsidy, which may entail the use of distortionary taxes.

[15] Pfähler (2001) notes that computational general equilibrium analysis is the state-of-the-art method to estimate the microeconomic effects of large publicly funded projects. This approach, because of its complexity and cost, is not common in airport studies. Thus, we do not examine this approach.

[16] An input-output model is a mathematical description of how all the sectors of an economy are related. For an introduction to input-output models and their limitations, see Coughlin and Mandelbaum (1991), Pfähler (2001), and O'Sullivan (2003). For advanced discussions, see Miller (1998) and Lahr and Dietzenbacher (2001).

ever, is a rigorous determination of whether a specific project should be undertaken.[17]

A study of the economic impact of Lambert–St. Louis International Airport by the Mumphrey Group (1996) estimated that for 1994 the economic impact of Lambert Airport was $5.1 billion, supporting 53,900 jobs, which is 4.4 percent of the total jobs in the St. Louis area. Based on forecasts of future aircraft operations, passenger enplanements, and cargo, for 2015 the economic impact of Lambert was estimated to be $7.8 billion, with employment of 84,500. This estimate of the future economic impact did not depend on the proposed expansion program; however, separate from the ongoing operations of Lambert, estimates of the impact of total construction spending at Lambert (including the runway expansion) were generated. Over a ten-year period, the total estimated impact of new construction at Lambert was $4.7 billion, or $470 million per year. This impact translates into annual household payroll of $108 million and 4,300 jobs.

Frequently, airport expansions have occurred simultaneously with the growth of an area. According to former Atlanta mayor Bill Campbell, as reported by McRainey (2002), Atlanta's Hartsfield Airport "anchors a regional economy that in 1998 outpaced the U.S. economic growth by more than two to one. And, the benefits of this output, including the performance of our airport, extend well beyond the boundaries of our region." The 1980 terminal addition and the 1994 new international concourse have "helped Atlanta become a business magnet." What is difficult to measure, however, is the extent to which the airport expansions contributed to economic growth.

Cooper (1990) notes that the purpose of economic impact studies is "more political than analytical."[18] Thus, one might have more confidence in studies of the economic impact of airport expansions in which the authors are less likely to have a political or economic stake in the results. One academic study of Amsterdam Schiphol Airport's expansions by Hakfoort, Poot, and Rietveld (2001) finds relatively modest economic impacts. Namely, they find that if the Schiphol airport had stopped growing in 1987, there would have been 42,000 fewer jobs in the Greater Amsterdam region in 1998.[19]

In summarizing their usefulness, Pfähler (2001) notes that input-output studies capture only some of the economic effects of a public expenditure program and that these effects may not be the relevant ones. Such studies are focused on the short-run demand effects and ignore the economic reasons for these projects. These projects should deliver services either to be used as intermediate inputs to produce goods/services by private firms or as a final consumption good/service. In the former case, the projects should stimulate productivity and growth; in the latter case, the projects should expand the consumption opportunities of consumers. Thus, an input-output approach will necessarily fail to provide an adequate analysis of a specific expansion project.

Production Function/Growth Studies

A common assertion is that an airport expansion will stimulate growth. This assertion is part of a view that public capital is a key ingredient for economic growth in a region. A large literature, propelled by Aschauer (1989), has addressed this issue. Aschauer (2000a) summarized this literature by noting that the estimates concerning the effect of public capital investment on private output ranged from very high to very low (low being not only zero, but negative).

In the present context, the issue is the impact of airport infrastructure on growth. Generally speaking, transportation infrastructure has been found to matter for growth. For example, Easterly and Rebelo (1993) find that infrastructure capital, especially that associated with transportation and communications, is related positively to economic growth.[20] However, few studies have focused on airport infrastructure. One exception is Cohen and Paul (2001), who estimate that a doubling of airport infrastructure capital in a particular state would lead to a 10 percent fall in manufacturing costs in

[17] See Niemeier (2001) for a discussion of the (mis)use of input-output analysis by the airline industry and for a statement that such an approach cannot answer whether the benefits of an expansion exceed its costs.

[18] A similar conclusion was reached by Sanders (2002) in a review of convention center expansion studies. In a similar vein, Pfähler (2001) notes that input-output studies are more useful for generating political support for an expansion project than in assessing its merits.

[19] Montalvo (1998) reported results of a 1991 study by the Air Transport Action Group that found the worldwide average of between 2,500 and 7,500 total jobs generated per million passengers. The total economic impact per million passengers is between $120 million and $1.5 billion. See Pfähler (2001) for a number of impact studies.

[20] Other studies have found that a given percentage increase in core capital (i.e., streets, highways, mass transit, sewers, and water systems) tends to have larger effects on output than other types of public capital. See Aschauer (1989, 2000b), Munnell (1990), and Lighart (2000).

that state. Thus, significant cost savings may accrue to the manufacturing sector from airport investments.[21] A second exception is Brueckner (2003), who examines the impact of changes in passenger enplanements in metropolitan areas on employment in the given area. He finds that employment in a metropolitan area increases by 1 percent for every 10 percent increase in passenger enplanements. To reiterate, however, the issue of whether a specific expansion is justified is not answered by such studies.

Benefit-Cost Analysis

Plans to expand the capacity of airports are subjected to a required benefit-cost analysis (BCA) prior to approval by the Federal Aviation Administration (FAA) (1999). During a BCA, the benefits realized and the costs borne by the producer and the consumers of the output, as well as the benefits realized and the costs borne by anyone else (so-called third parties), are incorporated in the analysis. A BCA, especially for a large project, can be very complicated. We examine key aspects of the BCA process using information from a BCA of Lambert–St. Louis International Airport.

Background on the Lambert–St. Louis International Airport Expansion. During 2001, passenger traffic at Lambert totaled 26.7 million, making Lambert the 17th busiest airport in the United States.[22] Passenger traffic during 2001 declined sharply, falling nearly 13 percent from the level of 30.6 million in 2000. Similar to airports throughout the United States, the sharp decline in passengers can be tied directly to the national economic recession that began during the spring of 2001 and the terrorist attacks of 9/11. The airport serves the 2.6 million residents and the businesses in the St. Louis metropolitan statistical area. In addition, especially because it is a hub for American Airlines, the airport plays a key role in the nation's air transportation system.[23] The airport has two parallel air carrier runways that are separated by 1300 feet. When adverse weather conditions occur, the airport is reduced to one precision instrument approach, which causes the capacity of the airport to decline 40 percent and congestion costs to increase substantially.

Step 1: Identify Project Objectives. Possible objectives for an airport expansion project include the following:

- Reduce delays associated with airport congestion,
- improve the efficiency of airport operations,
- increase the number of aircraft and passengers the airport can serve, and
- permit new service by accommodating larger and more efficient aircraft at the airport.

The Lambert infrastructure project was motivated by an assessment that the existing airport was already severely constrained (especially during bad weather) and that projected demand could not be met without unacceptable operational delays. A related assessment was that Lambert, especially because of its role as a hub, should not be allowed to become a bottleneck that would have detrimental effects throughout the national system.[24] Thus, the proposed project would allow the airport to effectively and safely accommodate projected levels of aviation activity at an acceptable level of delay by increasing airfield capacity for all types of weather conditions and decreasing delays at Lambert as well as nationwide.

Step 2: Specify Assumptions About Future Airport Conditions. Generally, to formulate reasonable plans, policymakers and researchers make assumptions about the projected growth in demand for airport services, changes in airport facilities and capacity that are likely to occur independently of the proposed project, and likely technological advancements in air traffic control. Forecasts of projected growth in airport activity are both difficult to make and very important. Incorrect forecasts will likely lead to a poor timing of airport investments, with overly optimistic (i.e., too high) forecasts causing underused facilities and with overly pessimistic (i.e., too low) forecasts imposing costs in terms of delays and inconvenience.[25] Similarly, ignoring

---

[21] Two other related issues—how public capital is financed and how it is used—have drawn attention. Barro (1990) showed that the benefits from productive government spending should be compared with the costs of the distortionary taxes required to fund the spending. See Aschauer (2000b) for a simultaneous estimate of these two effects. See the boxed insert for a discussion of the financing of airport expansions.

[22] Airports Council International web site: <http://www.airports.org/traffic>.

[23] During the BCA, Lambert was the hub for TWA. Subsequently, American Airlines acquired TWA and uses Lambert as one of its hubs.

[24] It was also recognized that Lambert provides economic benefits throughout the region and that those benefits would be threatened by an inadequate airport.

[25] Estimates of passenger enplanements and aircraft operations for 1997 through 2010 for the Lambert expansion were generated by the FAA. For 2011 through 2021, national estimates of the growth of passenger enplanements and aircraft operations were assumed to hold for Lambert. See FAA (1997).

FINANCING OF EXPANSIONS

A financial plan is a key component of any airport expansion project.[1] Three major sources are used for financing the expansions of U.S. airports. One of these is grants from the federal government through the FAA's Airport Improvement Program (AIP), which is discussed in more detail below. AIP grants are funded with ticket taxes and fees that are deposited into an airport's trust fund. The second source of funds, passenger facility charges (PFCs), are additional fees charged by airlines designated for capital improvements. Approximately $1.5 billion in PFCs are collected each year.[2] The third major source for airport financing, also discussed by Truitt and Esler (1996), consists of tax-exempt bonds issued by airports, which are based upon the expected future revenue streams of the airports.[3] Nearly 80 percent of all funding from these airport revenue bonds are allocated to the top 71 airports, while the more than 3,000 remaining airports share what is left. Funds generated by the new facility are used to repay the principal and interest on these bonds.

Economic Issues Associated with AIP Grants

An important issue with respect to AIP grants is their effect on the recipient's spending. One hypothesis is that, when a state or local government receives grant funds from a higher level of government, the lower level of government acts in the interests of the average constituent—the median voter. The lower level of government treats the grant funds as if they were an increase in the median voter's disposable income, a certain portion of which would be spent on the designated public services—in this case, airport usage. (That portion depends on the voter's marginal propensity to consume on that category of public services.) So, if the lower level of government acts in the interest of the median voter, one would expect it to spend a share of the grant equal to the median voter's marginal propensity to consume that public good or service out of his disposable income.

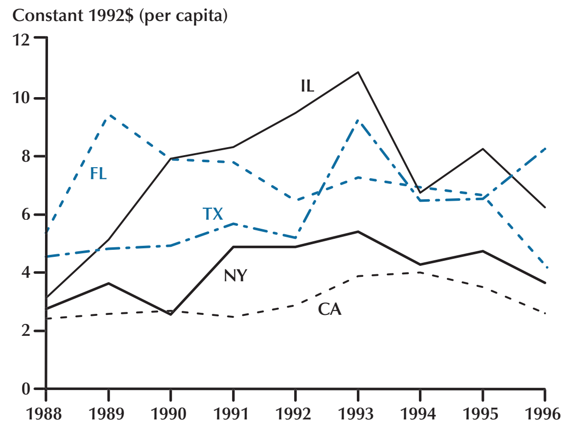
The empirical evidence for most types of public services does not support this theory. Namely, many researchers, such as Case, Rosen, and Hines (1993), have found that the impact of intergovernmental grants on state and local public expenditures exceeds the marginal propensity to consume out of disposable income. This has been called the "flypaper effect" because grant money tends to stick where it hits.[4] Cohen (2002) finds significant evidence of a flypaper effect for AIP grants. That is, he finds that the effect of AIP grants on state and local airport spending is greater than the median voter's marginal propensity to spend on airports, implying the presence of a flypaper effect for airports.

The figure shows that AIP grants per capita have varied over time in many states that have

(*Continued on p. 17*)

Figure B1

Per Capita AIP Grants to Selected States
1988-96



SOURCE: Cohen (2002).

[1] The financial condition of airlines and airports has deteriorated in recent years. See Golaszewski (2003) and Gillen and Lall (2003) for analyses of the financial shockwaves stemming from the events of 9/11. One possibility is that financial plans and the benefit-cost analyses supporting major capital improvements developed prior to 9/11 may no longer be justified.

[2] Information on PFCs and revenue bonds was obtained from the Air Transport Association's *Airline Handbook* (2001, Chap. 7).

[3] The Air Transport Association's *Airline Handbook* refers to these as general airport revenue bonds, which amounted to $53.6 billion since 1982 and have comprised over 95 percent of all airport debt. The bonds are serviced by airline revenues (e.g., landing fees and terminal building rentals), nonairline revenues (e.g., parking lot revenues, rental car concessions, terminal concessions, and general aviation fees), and interest income from various funds and accounts.

[4] See Gramlich (1977).

(*Continued from p. 16*)
large hub airports (i.e., an airport with greater than 1 percent of the nation's passenger enplanements). The U.S. Congress has debated reauthorization of the AIP on more than one occasion in recent years, which leads one to question how states and localities respond to cuts compared with increases in AIP grants. Specifically, when federal government grants are reduced, do states and localities reduce their own spending in the same way they increase their spending when grants are increased? States and localities may also try to pick up the slack when grants from the federal government are cut. Cohen (2002) looks at the question of symmetric state and local airport expenditure responses to cuts and increases in AIP grants over the period 1988-96, and he cannot reject the hypothesis that the state and local spending response is symmetric in the cases of cuts and increases. This is important because it implies that states and localities do not replace funds for airport expenditures with funds from other areas.

One other possible application of intergovernmental grants discussed in the public finance literature in Oates (1972) is to induce states to internalize spillovers. In other words, a higher level of government (such as the FAA) may use grants to subsidize states and localities for the benefits that they confer upon other states through their airport spending choices, as proposed by Cohen (1997). This subsidy would result in a shift in the marginal benefit curve that would further result in a higher equilibrium level of airport spending.[5] Thus, if these grants were used to subsidize all states that confer benefits upon other states, the subsidies could effectively induce all states to spend the socially optimal amount on airports. Ideally, however, this type of subsidy would be done at the individual airport level. The practical implementation of such a policy would be very challenging because the complexity of the air transportation network would make it extremely difficult to assess the magnitude of the spillover benefits.

[5] Coughlin, Cohen, and Khan (2002) present a similar marginal benefit/marginal cost framework in more detail for the optimal provision of aviation security.

future changes in airport facilities and capacity can lead to ill-timed investments. For example, a project to reduce runway congestion may be unwise if some other project affects the benefits of the proposed project. Finally, a runway expansion project may be unwise if scheduled improvements in air traffic equipment will accomplish the same result.

Step 3: Identify the Base Case. To assess one course of action, it is necessary to compare it with an alternative. One alternative, the base case, is not a "do nothing" alternative, but rather a reasonable course of action, absent the proposal being scrutinized. The base case assumes that airport managers will make optimal use of the existing and planned (i.e., currently underway and/or funded) airport infrastructure. In addition, as problems arise due to increases in airport traffic, airport managers, users, and air traffic managers are assumed to respond as best they can. Without question, the assumptions underlying this alternative can affect the decision of whether or not to undertake an expansion project.[26] With respect to the Lambert expansion, the assessment of the no-action alternative was that it would neither solve the capacity shortcomings nor mitigate the delays.

Step 4: Specify Alternatives To Meet the Objectives. With respect to Lambert, numerous alternatives were examined and eliminated from detailed analysis. These alternatives included the increased use of other modes of transportation, such as railroads, buses, and automobiles; the construction of a new airport to replace Lambert; a multiple airport system including Lambert and a supplemental airport; numerous airfield alignments; and other on-airport alternatives, such as the use of advanced navigational aids.

Various reasons led to the rejection of many alternatives. The increased use of other modes of transportation was rejected because it did not meet either local or national aviation needs and it would not have enhanced Lambert's contribution to the region. The construction of a new airport was rejected because of time and cost requirements. Moreover, other airports in the region could not fulfill the need for a new runway for various reasons, including Lambert's role as a hub, the lack of facilities at other airports, and airspace constraints. The airfield alignment alternatives were rejected because they were infeasible or because of environmental

[26] For details, see Appendix J of the FAA's *Record of Decision* that was signed on September 30, 1998.

reasons. Several on-airport alternatives were eliminated because they did not provide sufficient capacity or they did not allow for operations considered critical for a hub airport.

Based on discussions between airport officials and the FAA, an agreement was reached in February 1997 that the only feasible alternative to the airport's system of operation at that time was the airport sponsor's preferred plan for expansion. Thus, the comparison in the BCA was between the baseline and the alternative known as W-1W. The baseline alternative was to maintain Lambert in its present configuration with no additional capacity, whereas alternative W-1W was the construction of a new runway and the extension of an existing runway. The planned completion date was late 2001, with 2002 being the first full year of operation.

Step 5: Determine the Evaluation Period. The evaluation period is the number of years over which the benefits and costs of an investment are considered. To determine an appropriate evaluation period, one must first determine the economic life of the investment, which is the length of time the asset is expected to achieve its goals in a cost-effective manner. Generally speaking, the FAA uses an economic life of 20 years beyond the completion of construction for major airport infrastructure projects. For the Lambert expansion, the evaluation period began with the expected start of construction in 1998. Benefits were first expected from W-1W in 2002 and were expected to continue for 20 years through 2021.

Step 6: Measure and Compare the Benefits and Costs of Alternatives. The benefits of capacity-related airport projects primarily stem from the cost savings to airport users. A new runway, for example, should reduce aircraft, passenger, and cargo delay during normal airport operations. The costs of delay are incurred primarily by aircraft operators via more aircraft operating hours, passengers via longer trip times, and cargo shippers via longer shipping times.[27] Estimates of the time savings associated with an expansion project are generated with capacity simulation models.

The time savings must be translated into a single overall monetary measure. All values for benefits as well as costs, regardless of the year in which the benefits or costs arise, are expressed in constant (real) dollars. The FAA provides instructions as to the appropriate source for converting the following: (i) the reduction in aircraft delays into operating cost savings, (ii) the reduction in passenger travel delays into passenger travel time values, and (iii) the reduction of air cargo delay hours into opportunity cost savings.

Obviously, the benefits of any alternative will be at least partially offset by its costs. Costs include all capital, labor, and natural resources required for each alternative, no matter who—a governmental unit, the flying public, the general public, or some other group—bears the cost. Most airport projects involve relatively large expenditures at the beginning of a project in return for a flow of future benefits. Because a dollar expended today is more costly than a dollar expended later and a dollar of benefits received today is more valuable than a dollar of benefits received later, future cash flows must be converted into present value to allow for a meaningful comparison of a project's benefits with its costs.

The formula for the present value (PV) of a future benefit or cost to be received/incurred n years from the present is

$$PV = V_n/(1+r)^n,$$

where V is future value in year n in constant dollars, r is the annual real discount rate, and n is the number of years from the base year.

The PV formula allows a comparison of the benefits and costs for each alternative to see which, if any, is the most attractive. Various comparison methods are used, suggesting that no one method is superior to the others. Here, we will restrict our attention to two methods that are recommended when federal grants are involved—net present value (NPV) and benefit-cost ratio (B/C).

The formula for the calculation of NPV is

$$NPV = \sum_{t=0}^{k} (B-C)_t/(1+r)^t,$$

where B is future annual benefits in constant dollars, C is future annual costs in constant dollars, r is the annual real discount rate, k is the number of years from the base year over which the project is to be evaluated, and t is an index from 0 to k representing the number of years from the base year.

To be selected, an alternative must meet two conditions. First, it must have a positive NPV—that is, its benefits must exceed its costs. Second, the

[27] One complication arises from the fact that an airport expansion project, because it improves service, will cause greater use of the facility than would have occurred without the investment. These additional users will place new demands on the facility, eroding to some extent the per operation delay savings to preexisting airport users.

selected alternative must have an NPV exceeding all other alternatives.[28]

One problem with ranking projects using NPV involves comparisons of mutually exclusive alternatives of greatly different sizes. For example, assume two mutually exclusive alternatives—one with a $1 billion cost and a $100 million NPV and the other with a $100 million cost and a $30 million NPV. Based on the NPV, the former alternative would be chosen over the latter. If the airport is able to borrow as much as required to undertake all worthwhile projects, the NPV criterion would yield the correct choice.

However, if the airport is constrained in the amount of debt it can issue, say to $1 billion, and has several other independent potential projects with positive NPVs, then it may be best to select the $100 million alternative and use the remaining $900 million to fund the other projects. In this way, the $1 billion of borrowed funds might yield a sum of NPVs in excess of $100 million.

A second method of comparing benefits with costs is the B/C, which is simply the PV of benefits divided by the PV of costs. The formula for this calculation is

$$B/C=\left\{\sum_{t=0}^{k} B_t/(1+r)^t\right\}/\left\{\sum_{t=0}^{k} C_t/(1+r)^t\right\}.$$

Obviously, projects with positive NPVs will have B/Cs greater than 1. However, the B/C method might fail to answer correctly the question of how to accomplish a project's objectives most effectively. Analogous to our preceding discussion, problems with using this method can arise when comparing mutually exclusive projects of different sizes. Note that the larger alternative in the preceding example has a B/C ratio of 1.1, while the smaller alternative has a B/C ratio of 1.3. Recall that, depending on funding options and other considerations, the smaller project may or may not be the better alternative.

With respect to the BCA for the Lambert expansion, a number of economic and programmatic assumptions were made. Benefits and costs were denominated in 1997 dollars and, following FAA guidelines, discounted at 7 percent. Benefits consist of cost savings to passengers and to airlines. Passengers avoid delays. As a result, cost savings to passengers are the value of the time saved. The value of passenger time was assumed to be $27.90 per hour. Similarly, airlines avoid delays, so the cost savings are in the form of reduced variable operating costs. Variable operating costs per hour were assumed to be $1,189 in 2002, $1,333 in 2010, and $1,581 in 2021.

Alternative W-1W was assumed to be completed in 2001, with benefits beginning in 2002 and extending through 2021. Table 1 shows the calculated benefits in (discounted) present value dollars. With passenger delay savings of $661 million and operation delay savings of $908 million, total benefits were estimated to be $1,569 million.

Turning to the cost side, investment costs were assumed to occur so that 10 percent were incurred in 1998, 35 percent in both 1999 and 2000, and 20 percent in 2001. Annual recurring costs were assumed to be 3 percent of the initial capital cost per year. Table 2 shows the calculated costs. In present value terms, investment costs were estimated to be $537 million and recurring costs were estimated to be $142 million. These costs are offset to some extent by the disposal value of $76 million.[29] Thus, total costs were estimated to be $603 million.

Given the estimated benefits of $1,569 million and costs of $603 million, the NPV of alternative W-1W was $966 million. Thus, alternative W-1W was viewed as economically viable. The B/C was 2.6.

Step 7: Evaluate the Variability of Benefit-Cost Estimates. Underlying every BCA are numerous assumptions, estimates, and forecasts, which, if inaccurate, may introduce error in the NPV results. A thorough BCA will include a sensitivity analysis, which is a standard tool that varies key assumptions, estimates, and forecasts and provides alternative results for examination.

The evaluation of alternative W-1W examined three areas of risk: (i) investment cost overruns of 10, 25, and 50 percent; (ii) slippage in the construction schedule by 1, 2, and 5 years; and (iii) three cases in which traffic deviates from forecasted levels after 2005—traffic is held constant at 2005 levels, traffic is at 90 percent of the forecast, and traffic is at 110 percent of forecast.

The NPV and B/C results are shown in Table 3. For each case in which only one variable is altered, the NPV remains positive. A slippage in the construction schedule by five years causes the lowest NPV

[28] Note that this criterion is consistent with the notion discussed earlier that decisionmakers should choose the level of airport capacity such that the marginal benefits of an expansion equal the marginal costs. Such an action would maximize the net benefits of an expansion.

[29] The disposal value was determined by depreciating the initial capital investment at 3 percent per year until 2021 and then discounting that result to its present value.

Table 1

Benefit Estimate for W-1W* (1997 dollars in millions)

Benefit	Present value
Passenger delay savings	$661
Operational delay savings	$908
Total delay savings	$1,569

NOTE: *See FAA (1997, Table 4).

Table 2

Cost Estimate for W-1W* (1997 dollars in millions)

Cost	Present value
Investment	$537
Recurring	$142
Disposal value	($76)
Total cost less disposal value	$603

NOTE: *See FAA (1997, Table 3).

Table 3

Sensitivity Analysis for W-1W*

Scenario	NPV (1997 dollars in millions)	B/C
W-1W as planned	$966	2.6
Investment cost overrun		
10%	$912	2.4
25%	$832	2.1
50%	$698	1.8
Slippage in construction schedule		
1 year	$795	2.3
2 years	$639	2.1
5 years	$248	1.4
Traffic relative to forecast after 2005		
Constant	$551	1.9
90% of forecast	$782	2.2
110% of forecast	$1,095	2.7
Two-variable analysis		
10% overrun and 1-year slippage	$741	2.1
25% overrun and 2-year slippage	$505	1.7
50% overrun and 5-year slippage	($21)	1.0

NOTE: *See FAA (1997, Table 5).

of $248 million and the lowest B/C of 1.4. Not surprisingly, the NPV rises to $1095 million and the B/C rises to 2.7 when traffic exceeds its forecasted levels by 10 percent. Only when the overrun in investment costs is 50 percent and the slippage in the construction schedule is 5 years does the NPV become negative.

Step 8: Recommend a Best Course of Action. The final step of a BCA is to recommend a specific alternative. The FAA recommends that the alternative with the largest, positive NPV be given primary consideration. However, the project sponsor may select another alternative if it provides important hard-to-quantify benefits or is substantially less risky than the alternative with the largest NPV. For the BCA of the Lambert expansion performed by the FAA, alternative W-1W is the preferred alternative. Its NPV is positive and, because it was the only alternative evaluated, there is no option that yields a larger NPV. In addition, the sensitivity analysis suggests that the net effects of W-1W are likely to be positive under a wide range of scenarios.

AIRPORT EXPANSION IMPACTS THAT GENERATE CONTROVERSY

Airport expansion projects have impacts far beyond those identified in economic impact studies. Numerous social and environmental impacts that generate much displeasure and contentiousness are the norm.[30] Lawsuits are filed in opposition to virtually every expansion of a major airport. Opponents generally challenge the right of airport officials to override local zoning rules or increase noise or air pollution. Almost without exception, the legal challenges are unsuccessful. Thus, the projects, albeit delayed, do proceed.

Prior to FAA approval of an expansion project, actions to mitigate adverse effects of expansions are reviewed. Below we take a closer look at two common impacts—one that can be characterized as a social impact and one as an environmental impact.

Displacement and Compensation of People and Businesses

According to the FAA (1998), the Lambert expansion would necessitate the relocation of 5,680 people, including 2,324 households. Over 1900 residential parcels of property were to be acquired. In addition, the relocation of 75 businesses would be required. The city of Bridgeton was to bear the brunt of the relocation.[31] To mitigate the substantial disruptive effects, various services designed to educate, inform, and respond to those affected were and are being provided. Nonetheless, controversy abounds.

Not surprisingly, many of the homeowners forced to sell their property and relocate were less than pleased with the airport expansion. Litigation between them and the City of St. Louis over condemnation of their homes has been extensive. As reported by McClellan (2001), the litigation has occurred despite the fact that the city has the authority to seize homes and has made "an effort to negotiate with each of the homeowners." Despite attempts to use independent citizens to determine a "fair and reasonable price," some think the city's offers were unfair. Often, residents accept them anyway since they do not want to have to continue fighting, and they want to escape "neighborhoods where nobody dares put money into their homes."

Market forces underlie these compensation prices. Although some might consider the buyout price to be too low, because the homes are already located near the airport and, thus, affected by associated noise, it is likely that the value of these homes and those in surrounding communities already reflects the proximity to the airport. These circumstances might make them less desirable in general in the local real estate market, which we discuss further.

Environmental Issues: Noise

The expansion of an airport likely leads to additional noise in the area surrounding an airport. Some of the affected area is likely to be acquired by the airport as part of the expansion. To the extent that remaining residential areas must endure increased noise, one can argue that these residents and homeowners should be compensated. Not surprisingly, there is often opposition from local community groups when an airport in an urban area plans to expand. Healy (2002) notes one example of this resistance in the Boston area, which has led to the stipulation that the new runway be used only on "days when northwest winds blow at 11.5 miles per hour or more."

Housing prices are likely affected by airport noise.[32] Because noise becomes capitalized into property values, it is an example of a pecuniary

[30] A lengthy, but not exhaustive, list of the social and environmental impacts includes the disruption/destruction of communities/neighborhoods, historic areas, biotic communities, floodplains, farmland, and parks as well as adverse effects on noise levels, air quality, and water quality.

[31] Higgins and Chase (2002) state that the cities of Bridgeton and St. Charles brought lawsuits in state court challenging Lambert's right to override local zoning laws and in federal court arguing that the FAA failed to consider more cost-effective ways to improve the airport. The lawsuits failed, and ground on the new runway was broken in July 2001.

[32] For example, Espey and Lopez (2000) find for the Reno-Sparks, Nevada, area a $2400 difference between the prices of homes subject to at least 65 decibels of noise and other equivalent homes. Tompkins et al. (1998) find mixed evidence of capitalization of noise for an airport in Manchester, England. They also found, however, that the benefits of easy access to the airport outweighed other costs associated with proximity, such as noise.

externality.[33] Consequently, an airport expansion leading to additional noise should be countered with a one-time compensation of property owners to account for this capitalization, or property owners should be helped to take defensive action against the noise. Those who choose to live near airports should not be compensated further, since they receive relatively inexpensive housing and they choose to live there. Local government in Chicago provides an example of city assistance to homeowners: In 1996, mayor Richard M. Daley began a program that has led to the soundproofing of nearly 4,000 homes and 75 schools in communities surrounding O'Hare Airport. In the case of the Lambert expansion, homeowners in a designated area likely to be adversely affected by increased noise were offered three options: sales assistance, sound insulation, or a cash payment. In exchange the airport received the right, known as an avigation easement, to generate increased noise.[34]

OTHER POSSIBLE SOLUTIONS TO CONGESTION

Airport expansions increase the capacity of the air transportation system by increasing the amount of capital. As discussed previously, an alternative solution might entail using the existing capital more efficiently. In this section we examine one such solution, congestion-based pricing of landing fees.[35]

Airport congestion is an example of a true (or technological) externality because the resulting delays imposed from too many aircraft operators choosing to land at peak hours causes disutility to others who are waiting to use the runways. Some researchers, most notably Morrison and Winston (1989), have argued in favor of congestion-based pricing of landing fees to solve the airport congestion problem. They claim that this type of pricing system may make airport use more efficient. At most airports, landing fees are structured according to aircraft size or weight, and the runway capacity is allocated to aircraft on a "first come, first served" basis (Bailey, 2002). Under the current system, there is a market failure due to the fact that aircraft owners fail to internalize (i.e., fully bear) the costs to society that arise when they choose to land at an already busy airport. These costs come in the form of travel time delay due to longer taxi lines, circling of aircraft before being cleared for landing, etc.

With congestion-based pricing, the owners of aircraft would pay landing fees based on the marginal damage in terms of runway delay that is caused by their aircraft rather than pay fees based on aircraft size or weight. These fees might vary at a particular airport depending on the time of day. This way, when deciding when and where to land, an owner of an aircraft would be forced to consider the marginal social costs of landing at a particular airport at a particular time, instead of merely looking at the marginal private costs. Thus, congestion-based landing fees can correct for the externality that causes congestion at certain airports at certain times of the day. But it also has implications for equity, since small planes would pay as much as large ones at any given time of day. The FAA is considering imposing congestion-based landing fees at some airports, including New York's La Guardia airport and Boston's Logan airport.[36]

The arguments for congestion pricing made by Morrison and Winston are similar to studies from the road-pricing literature discussed by Small (1992), where there are many users of the public infrastructure. But in the case of most airports, many of the airport users (i.e., airlines) have market power, which differs from the road-pricing case. Brueckner (2002a) shows that when an airport is dominated by a single monopolist airline, the congestion effects are completely internalized by the airline, thus providing no role for congestion pricing. In competitive situations, even when the number of competitors is two, the airlines do not internalize the congestion that they impose on each other. Thus, there would be a role for congestion pricing, where the toll varies inversely with the carrier's flight share.

Another possible solution to airport congestion is airline route restructuring. As reported by McCartney (2002), American Airlines recently

[33] Some argue that because airport noise is an externality, government intervention is necessary. Baumol and Oates (1988, p. 17) provide a required condition for a (technological) externality to be present: "An externality is present whenever some individual's (say A's) utility or production relationships include real (that is, nonmonetary) variables, whose values are chosen by others (persons, corporations, governments) without particular attention to the effects on A's welfare." However, they note that a pecuniary externality (as opposed to a technological externality) occurs when one person's actions affect the financial circumstances of another. A pecuniary externality occurs when one person's actions affect only the prices that another party faces, and thus it is not an externality that necessitates governmental intervention.

[34] See FAA (1998) for details.

[35] The inefficiencies that result in frequent delays at airports have convinced some that airports should be privatized. See Morrison and Winston (2000) for a discussion of airport privatization.

[36] See Brueckner (2002b) and *Boston Globe* (2002).

announced that they would begin to move away from the traditional "hub and spoke" route structure at some of their hubs and move toward a "rolling hub" structure where passengers wait for planes to arrive at hubs instead of planes waiting for passengers to arrive. Southwest Airlines has been successful at implementing such a route structure, and if enough airlines were to adopt such a structure there could be noticeable amelioration of congestion.

CONCLUSIONS

Air transportation services play an important role in every metropolitan area. Congestion is a problem at many of these airports. Because of the network features of the air transportation system, congestion at one airport can have adverse effects on the operations at other airports. The large cost associated with congestion provides the incentive for mitigating congestion. But, finding satisfactory solutions is complicated by the interconnectedness of the air transportation system. In this paper, we reviewed the economic analyses and issues associated with one common solution, expanding the capacity of airports. While an expansion may reduce congestion, we show that congestion may persist.

Expanding an airport is a complex undertaking that requires a large capital expenditure. A fundamental question is whether the benefits of the project exceed its costs. Ideally, the proposed expansion maximizes the net present value of the difference between the benefits and the costs. By examining the steps in a benefit-cost analysis for Lambert–St. Louis International Airport, we illustrate the economic analysis that underlies a typical expansion project.

The benefits of an expansion project to users of air transportation services extend throughout the local economy and likely beyond, whereas some significant social and environmental impacts are concentrated near the airport. Not surprisingly, the typical expansion project is controversial, with lawsuits being quite common. We attempt to illustrate these social and environmental effects and examine ways to mitigate their adverse effects. Airport expansions frequently disrupt neighborhoods and nearby communities, an example of which is the destruction of homes in Bridgeton that was judged to be a necessary part of the Lambert expansion. Homeowners received compensation for their property, but entire neighborhoods were destroyed. Another disruption is the additional noise imposed on surrounding communities due to larger airports. The new runway project in Boston and the associated noise led to agreements on limitations of its future use.

Finally, we examined an alternative to expansion—congestion pricing—that attempts to use the existing airport capital more efficiently. We demonstrate that a tax on airport users would improve airport efficiency. This is consistent to some extent with the notion of congestion pricing, which posits that users should be charged separate landing fees, based on the marginal damage in terms of travel delays that given users impose on others.

REFERENCES

Air Transport Association. *Airline Handbook*, 2001. <http://www.airlines.org/public/publications/display1.asp?nid=961>.

Aschauer, David A. "Is Public Expenditure Productive?" *Journal of Monetary Economics*, March 1989, *23*(2), pp. 177-200.

___________. "Do States Optimize? Public Capital and Economic Growth." *Annals of Regional Science*, August 2000a, *34*(3), pp. 343-63.

___________. "Public Capital and Economic Growth: Issues of Quantity, Finance, and Efficiency." *Economic Development and Cultural Change*, January 2000b, *48*(2), pp. 391-406.

Bailey, Elizabeth E. "Aviation Policy: Past and Present." *Southern Economic Journal*, July 2002, *69*(1), pp. 12-20.

Barro, Robert J. "Government Spending in a Simple Model of Endogenous Growth." *Journal of Political Economy*, October 1990, *98*(5, Part 2), pp. 103-26.

Baumol, William J. and Oates, Wallace E. *The Theory of Environmental Policy*. Second Edition. Cambridge: Cambridge University Press, 1988.

Bayles, Fred. "Airport Expansion Plans Come to Screeching Halt." *USA Today*, 12 November 2001, p. 6B.

Boston Globe. "A Runway If..." Editorial, 25 June 2002, p. A14.

Brueckner, Jan K. "Airport Congestion When Carriers Have Market Power." *American Economic Review*, December 2002a, *92*(5), pp. 1357-75.

___________. "Internalization of Airport Congestion." *Journal of Air Transport Management*, May 2002b, *8*(3), pp. 141-47.

___________. "Airline Traffic and Urban Economic Development." *Urban Studies*, 2003 (forthcoming).

Case, Anne C.; Rosen, Harvey S. and Hines, James R. Jr. "Budget Spillovers and Fiscal Policy Interdependence: Evidence from the States." *Journal of Public Economics*, October 1993, *52*(3), pp. 285-307.

Cohen, Jeffrey P. "Some Issues in Benefit Cost Analysis for Airport Development." *Transportation Research Record*, July 1997, No. 1567, pp. 1-7.

___________. "Reciprocal State and Local Airport Spending Spillovers and Symmetric Responses to Cuts and Increases in Federal Airport Grants." *Public Finance Review*, January 2002, *30*(1), pp. 41-55.

___________ and Morrison Paul, Catherine J. "Airport Infrastructure Spillovers in a Hub and Spoke System." Working Paper 01-011, University of California-Davis, Department of Agricultural and Resource Economics, December 2001.

Cooper, Ronald. "Airports and Economic Development: An Overview." *Transportation Research Record*, 1990, No. 1274, pp. 125-33.

Coughlin, Cletus C.; Cohen, Jeffrey P. and Khan, Sarosh. "Aviation Security and Terrorism: A Review of the Economic Issues." Federal Reserve Bank of St. Louis *Review*, September/October 2002, *84*(5), pp. 9-24.

___________ and Mandelbaum, Thomas B. "A Consumer's Guide to Regional Economic Multipliers." Federal Reserve Bank of St. Louis *Review*, January/February 1991, *73*(1), pp. 19-32.

Easterly, William and Rebelo, Sergio. "Fiscal Policy and Economic Growth: An Empirical Investigation." *Journal of Monetary Economics*, December 1993, *32*(3), pp. 417-58.

Espey, Molly and Lopez, Hilary. "The Impact of Airport Noise and Proximity on Residential Property Values." *Growth and Change*, Summer 2000, *31*(3), pp. 408-19.

Federal Aviation Administration. *Benefit-Cost Analysis for Lambert–St. Louis International Airport Capacity Enhancement Project*, July 1997.

___________. *Record of Decision on Lambert–St. Louis International Airport*, September 1998.

___________. Office of Aviation Policy and Plans. *FAA Airport Benefit-Cost Analysis Guidance*, December 1999.

Gilbert, Virginia Baldwin. "Lambert Field Tunnel Will Be First of Its Kind." *St. Louis Post-Dispatch*, 17 November 2002, pp. F1,8.

Gillen, David and Lall, Ashish. "International Transmission of Shocks in the Airline Industry." *Journal of Air Transport Management*, January 2003, *9*(1), pp. 37-49.

Golaszewski, Richard. "Network Industries in Collision: Aviation Infrastructure Capacity, Financing and the Exposure to Traffic Declines." *Journal of Air Transport Management*, January 2003, *9*(1), pp. 57-65.

Gramlich, Edward M. "Intergovernmental Grants: A Review of the Empirical Literature," in Wallace E. Oates, ed., *The Political Economy of Fiscal Federalism*. Lexington, MA: Heath-Lexington, 1977, pp. 219-39.

Hakfoort, Jacco; Poot, Tom and Rietveld, Piet. "The Regional Economic Impact of an Airport: The Case of Amsterdam Schiphol Airport." *Regional Studies*, October 2001, *35*(7), pp. 595-604.

Healy, Jack. "FAA Seen Giving Boost to New Logan Runway: Foes Greet Plan with Skepticism." *Boston Globe*, 22 June 2002, pp. A1,5.

Higgins, Michael and Chase, John. "Suits Over Airport Projects Rarely Win but Often Delay." *Chicago Tribune*, 4 June 2002, news section p. 1.

Lahr, Michael L. and Dietzenbacher, Erik eds. *Input-Output Analysis: Frontiers and Extensions*. New York: Palgrave, 2001.

Lighart, Jenny E. "Public Capital and Output Growth in Portugal: An Empirical Analysis." Working Paper No. 00/11, International Monetary Fund, January 2000.

McCartney, Scott. "Clipped Wings: American Airlines to Retrench in Bid to Beat Discount Carriers." *Wall Street Journal*, 13 August 2002, p. A1.

McClellan, Bill. "Bridgeton Die-Hards Live To Fight Over Airport Expansion." *St. Louis Post-Dispatch*, 2 September 2001, p. C1.

McRainey, Megan. "#1: Hartsfield Soars with Atlanta." *Atlanta Business Chronicle*, 27 May 2002. <http://atlanta.bizjournals.com/atlanta/stories/2002/05/27/focus1.html>.

Miller, Ronald E. "Regional and Interregional Input-Output Analysis," in Walter Isard et al., eds., *Methods of Interregional and Regional Analysis*. Aldershot, England: Ashgate, 1998, pp. 41-134.

Montavlo, Jose G. "A Methodological Proposal to Analyze the Economic Impacts of Airports." *International Journal of Transport Economics*, June 1998, *25*(2), pp. 181-203.

Morrison, Steven A. and Winston, Clifford. "Enhancing the Performance of the Deregulated Air Transportation System." *Brookings Papers on Economic Activity: Microeconomics*, Special Issue 1989, pp. 61-112.

___________ and ___________. "The Remaining Role of Government Policy in the Deregulated Airline Industry," in Sam Peltzman and Clifford Winston, eds., *Deregulation of Network Industries: What's Next?* AEI-Brookings Joint Center for Regulatory Studies, 2000, pp. 1-40.

Mumphrey Group. *Economic Impact Study*, prepared for Lambert–St. Louis International Airport, City of St. Louis Airport Authority, January 1996.

Munnell, Alicia H. "How Does Public Infrastructure Affect Regional Economic Performance?" *New England Economic Review*, September/October 1990, pp. 11-32.

Niemeier, Hans-Martin. "On the Use and Abuse of Impact Analysis for Airports: A Critical View from the Perspective of Regional Policy," in Wilhem Pfähler, ed., *Regional Input-Output Analysis: Conceptual Issues, Airport Case Studies, and Extensions*. Baden-Baden, Germany: Nomos, 2001, pp. 201-20.

Oates, Wallace. *Fiscal Federalism*. 5th Edition. New York: Harcourt, Brace, Jovanovich, 1972.

O'Sullivan, Arthur. *Urban Economics*. New York: McGraw-Hill, 2003.

Pfähler, Wilhelm, ed., "Input-Output Analysis: A User's Guide and Call for Standardization." *Regional Input-Output Analysis: Conceptual Issues, Airport Case Studies, and Extensions*. Baden-Baden, Germany: Nomos, 2001, pp. 11-45.

Salant, Jonathan D. "Airports Across the Country Tardy." Associated Press, 22 July 2001.

Sanders, Heywood T. "Convention Myths and Markets: A Critical Review of Convention Center Feasibility Studies." *Economic Development Quarterly*, August 2002, *16*(3), pp. 195-210.

Small, Kenneth A. *Urban Transportation Economics*. Chur, Switzerland: Harwood Academic Publishers, 1992.

Tomkins, J.; Topham, N.; Twomey, J. and Ward, R. "Noise versus Access: The Impact of an Airport in an Urban Property Market." *Urban Studie*s, February 1998, *35*(2), pp. 243-58.

Truitt, Lawrence J. and Esler, Michael. "Airport Privatization: Full Divestiture and Its Alternatives." *Policy Studies Journal*, 1996, *24*(1), pp. 100-10.

FOMC Forecasts: Is All the Information in the Central Tendency?

William T. Gavin

Since 1980, there has been a substantial improvement in the performance of monetary policy among most of the industrialized countries. Bernanke et al. (1998) attribute this success to the adoption of inflation targeting, but they are somewhat vague in defining what it means to have an inflation target. Even countries that do not have numerical inflation targets are said to be inflation targeting if they focus on a long-term objective for price stability (low inflation) and publish inflation forecasts. By this broad definition, even the Federal Reserve may be considered to have an inflation target.

This article describes and documents 23 years of FOMC forecasts. Congress required the Fed to begin preparing and reporting forecasts in 1979. Section 108 of the Humphrey-Hawkins Act explicitly required the Fed to submit "written reports setting forth (1) a review and analysis of recent developments affecting economic trends in the nation; (2) the objectives and plans…with respect to the monetary and credit aggregates…; and (3) the relationship of the aforesaid objectives and plans to the short-term goals set forth in the most recent *Economic Report of the President*." In order to satisfy the third item, the Federal Reserve Chairman began reporting a summary of Fed policymakers' forecasts to Congress in July 1979.[1] Since then, similar summaries of forecasts have been reported every February and July. Forecasts are made of annual, fourth-quarter-over-fourth-quarter growth rates for nominal gross domestic product (GDP), real GDP, and inflation.[2] Fed policymakers also forecast the average level of unemployment for the fourth quarter of the year. In February, the forecasts pertain to the current calendar year (referred to below as the 12-month-ahead forecasts). In July, forecasts are updated for the current calendar year (6-month-ahead forecasts) and preliminary projections are made for the next calendar year (18-month-ahead forecasts).

Separate forecasts are made by each of the FOMC members (and nonvoting Federal Reserve Bank presidents), but the individual forecasts are not published.[3] Rather, the Fed reports two summary statistics: the low and the high forecast among all the policymakers and the *central tendency*, which has been reported since February 1983 and is a smaller range that omits extreme forecasts.

This paper evaluates the information content of the two summary statistics, the full range and central tendency, and defines two FOMC forecasts based on those statistics: the midpoint of the full range (FR) and the midpoint of the central tendency range (CT). The paper compares forecast accuracy, tests for bias and efficiency, and uses encompassing tests to see whether either of these summary statistics encompasses the other. Admittedly, the two forecasts calculated from the FR and CT are quite close. In many of the tests, we cannot distinguish between the two. However, where statistically significant differences exist, they generally favor the use of the FR.

THE HISTORICAL RECORD

If the outlying forecasts for the FR and CT are symmetric, then the two will be the same. The FOMC has generated a large number of forecasts since 1979. There are two alternative ranges for each of four forecasted economic variables made at three different horizons (24 separate forecast series rang-

[1] The reporting requirements of the Humphrey-Hawkins Act expired in May 2000. Congress amended and continued the reporting requirements in the American Homeownership and Economic Opportunity Act of 2000 (Section 1003).

[2] The Fed followed the Bureau of Economic Analysis, switching from GNP to GDP in 1992.

William T. Gavin is a vice president and economist at the Federal Reserve Bank of St. Louis. Athena Theodorou provided research assistance.

[3] The FOMC is the policymaking committee of the Federal Reserve System. When the Board is full, the Committee consists of the 7 governors of the Board, the president of the Federal Reserve Bank of New York, and 4 of the remaining 11 Federal Reserve Bank presidents who serve on a rotating basis. All 12 presidents attend every meeting, contribute to the discussion, and provide forecasts that are summarized in testimony to the Congress. In this article, when we refer to the FOMC forecasts we mean the summary of forecasts that are collected from FOMC members and nonvoting Federal Reserve Bank presidents.

ing from 18 to 23 years each). To compare the forecast with the outcome, we measure actual output, inflation, and unemployment using real-time data rather than the latest vintage data.

There is always an issue about how to define the "actual" data because the data are regularly revised as we get more information about the past. The analysis in this paper is based on the real-time data that are maintained at the Federal Reserve Bank of Philadelphia. They created vintages of data that were available in the middle of the following months: February, May, August, and November.[4] Except where explicitly noted, the actual values are taken from the first report for a calendar year that occurs at the end of January and is stored in the "February" files by the Philadelphia Fed.

In preliminary work we examined the root-mean-squared errors (RMSE) calculated using the different versions of the data that became available during the year following the forecast period. We found that the FOMC's nominal GDP and real GDP forecasts were most accurate for the first released figures reported in February. The FR forecast for inflation was about equally accurate using the February or November data, and the CT inflation forecast was most accurate using the November data. In all cases the inflation forecasts were closer to the February version than they were to the May version. The unemployment data were rarely revised within the first year, and there were no revisions to the fourth-quarter average in our sample.

The 12-month forecasts are used to illustrate the FOMC forecasts in Figure 1. The top panel of Figure 1 shows nominal GDP growth. Through 1995, the FOMC, on average, overpredicted nominal GDP growth. The nominal GDP forecast errors were particularly large in years that included recessions: 1981-82, 1990-91, and 2001. On average the real GDP forecasts shown in the second panel appear to be unbiased, but the forecast errors were large and negative in recessions and large and positive in the years 1995 through 1999. There was an unexpected increase in productivity growth during the latter half of the 1990s.

The third panel shows the inflation forecasts made by the FOMC. The FOMC began forecasting the implicit price deflator for GNP in 1979. Much of the bias in the nominal GDP forecast came from a tendency to overpredict inflation in the 1980s. The FOMC continued reporting forecasts for the deflator until 1989, when it began making inflation forecasts in terms of the consumer price index (CPI). CPI inflation accelerated to 7 percent in the second two quarters of 1990, exceeding the FOMC forecasts. Inflation decelerated rapidly in 1991 and stayed, on average, below the forecast through the 1990s. In 2000, the FOMC switched once again, this time to the chain price index for personal consumption expenditures (PCE).

The fourth-quarter average unemployment rate is shown in the bottom panel of Figure 1. Here, the interesting features are the trend decline of unemployment that appears to mirror the decline of inflation, the tendency to overpredict unemployment in the 1990s (another facet of the surprisingly high productivity growth), and the relatively accurate prediction of unemployment in the 1990-91 recession.

Descriptive statistics and tests comparing the alternative forecasts are shown in Table 1. Our analysis uses 1983-2001 data for the 6- and 12-month forecasts and 1984-2001 data for the 18-month forecasts so that we can compare the forecasts derived from the FR with those computed from the CT.[5] Five statistics are reported in Table 1: the mean error, the RMSE, p-values from a Wilcoxon signed-rank statistic that tests whether the RMSEs from the FR and CT forecasts are equal, the widths of the intervals for FR and CT, and the percentage of times that the actual value fell outside the FR and CT.

As we saw with the 12-month forecasts in Figure 1, the FOMC tended to overpredict inflation, nominal GDP growth, and the unemployment rate. In each of these cases, the negative bias in the mean error was largest for the 18-month forecasts. There was no bias in the 6-month real GDP forecast and a positive bias in the 12- and 18-month real GDP forecasts.

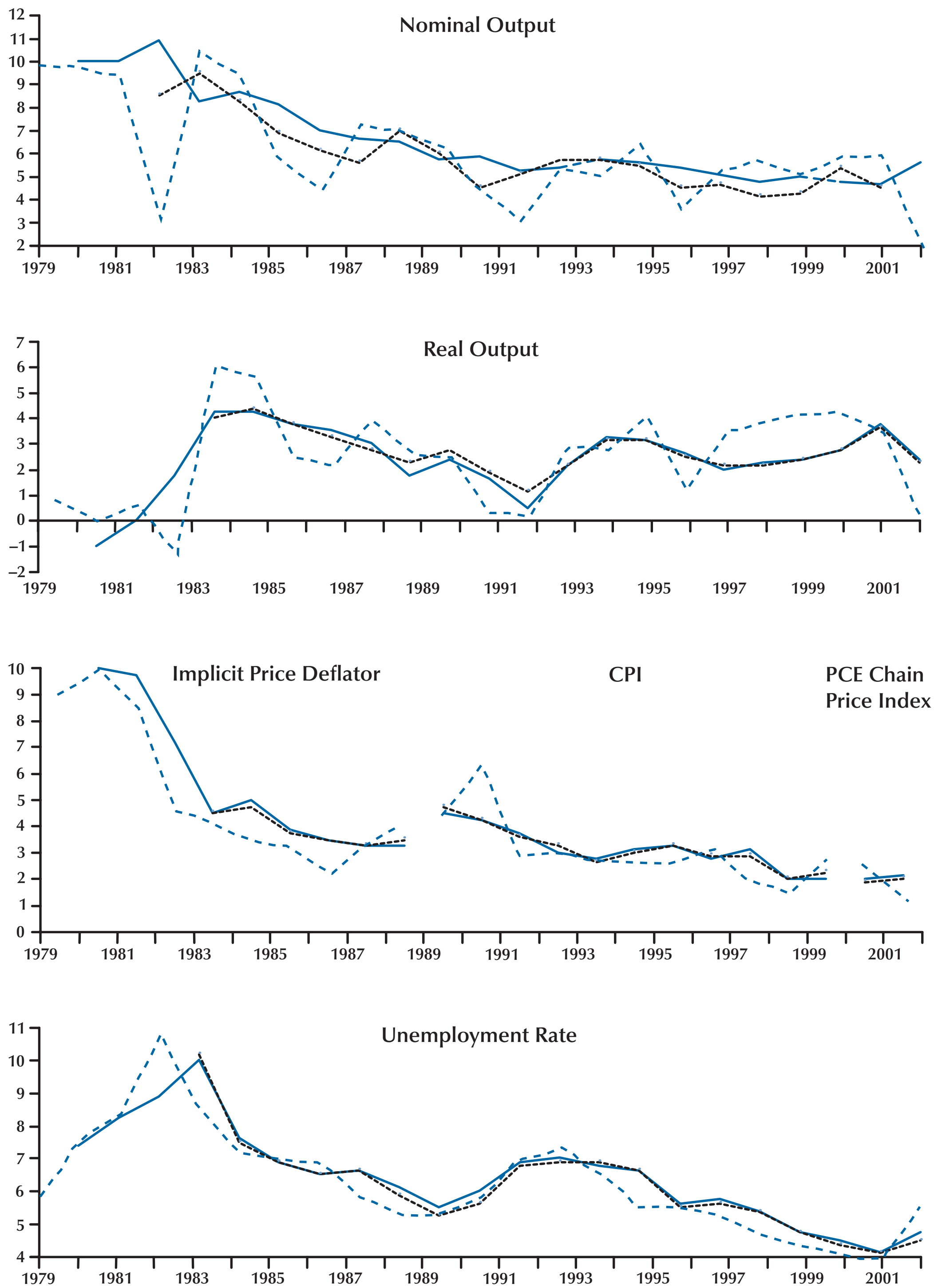
The RMSEs are shown in the second row of each panel. The nominal GDP RMSEs were largest, 10 to 20 percent larger than those for real GDP and almost twice as large as those for inflation. Generally, the real GDP growth and inflation errors were weakly negatively correlated. The probability values of a Wilcoxon signed-rank test are reported in parentheses in the third row of each panel for the FR forecasts. This is a test of the null hypothesis that

[4] The forecasts and real-time data used in this article are available on the Bank's web site: <research.stlouisfed.org>.

[5] See Gavin and Mandal (2002) for an analysis of the FR forecasts for output and inflation, including a comparison with the Blue Chip consensus and Green Book forecasts. They find that the FOMC's output forecasts were as good as the Blue Chip and Green Book forecasts. The FOMC's inflation forecasts were better than the Blue Chip, but not as good as the Green Book.

Figure 1

FOMC Forecasts: 1979 to 2001



NOTE: Actual values (blue dashed lines) are those first available in February of the following year. The FR is shown by a solid line and the CT by the black dotted line.

Table 1

FOMC Forecasts by Forecast Horizon

	FR			CT		
	6-month	12-month	18-month	6-month	12-month	18-month
Nominal GDP						
Mean error	–0.31	–0.17	–0.43	–0.35	–0.18	–0.59
RMSE	0.97	1.34	1.52	0.93	1.45	1.67
Wilcoxon signed-rank test (p-value)	(0.98)	(0.14)	**(0.05)**			
Width of range in percentage points	1.61	2.09	2.37	0.66	0.72	0.98
Percent of time outside the range	53	58	44	79	89	72
Real GDP						
Mean error	–0.01	0.25	0.19	0.00	0.20	0.09
RMSE	0.78	1.26	1.33	0.79	1.27	1.41
Wilcoxon signed-rank test (p-value)	(0.86)	(0.86)	(0.56)			
Width of range in percentage points	1.11	1.43	1.55	0.42	0.53	0.51
Percent of time outside the range	68	68	50	95	100	78
Inflation						
Mean error	–0.29	–0.41	–0.60	–0.34	–0.39	–0.65
RMSE	0.51	0.69	0.85	0.52	0.69	1.01
Wilcoxon signed-rank test (p-value)	(0.42)	(0.92)	**(0.07)**			
Width of range in percentage points	0.95	1.32	1.94	0.39	0.49	0.63
Percent of time outside the range	47	53	39	89	74	67
Unemployment						
Mean error	–0.15	–0.29	–0.31	–0.12	–0.20	–0.23
RMSE	0.38	0.60	0.81	0.37	0.63	0.80
Wilcoxon signed-rank test (p-value)	(0.70)	(0.73)	(0.76)			
Width of range in percentage points	0.56	0.64	0.88	0.19	0.22	0.35
Percent of time outside the range	37	68	67	68	89	78

NOTE: Wilcoxon's signed-rank test is used to test the null hypothesis that the forecasts are equally accurate (that is, the RMSEs are equal). The significance levels were computed from a program provided by Rob van Son found at <http://www.fon.hum.uva.nl/Service/Statistics/Signed_Rank_Algorihms.html>.

Bold cells indicate that the probability value is less than 10 percent.

The data span the period 1983-2001 for 6- and 12-month horizons and 1984-2001 for the 18-month horizon.

the RMSEs from the FR and CT forecasts are equal. We use a Wilcoxon signed-rank statistic because Diebold and Mariano (1995) show that this test statistic is well-sized in cases where the sample is small and the alternative forecast errors are highly correlated and possibly serially correlated—as we expect for the 18-month forecasts.[6] In only 2 of the 24 cases is the probability value less than 10 percent—the 18-month forecasts for nominal GDP growth and inflation. In both cases the forecast derived from the FR was more accurate than the forecast derived from the CT.

The fourth row of each panel reports the average

[6] The 18-month forecast for year t is made before the actual value for year t–1 is known. Any surprise in the actual value for year t–1 will be part of the forecast error for year t–1. This surprise is also likely to be reflected in the actual value (and forecast error) for year t. Thus, we can expect first-order serial correlation in the 18-month forecast errors.

width of the FR or CT intervals. The width of the range is the distance in percentage points between the high- and low-point forecasts. We define the FOMC consensus forecast as the midpoint of the range between the high and low forecasts. A way to gauge whether the range is a good measure of the uncertainty about the forecast is to see how often the actual value falls outside the range. The bottom row in each panel reports the percentage of time that the outcome fell outside the reported FR or CT. The actual values were outside the CT most of the time, indeed, all of the time for the 12-month output forecasts.

The average width of the FR for the 18-month output forecasts is 1.55 percentage points. The actual value fell outside this range 50 percent of the time. The range of 12-month output forecasts was almost as wide (1.43 percentage points), but the actual value fell outside this range 68 percent of the time. At the shortest horizon, the width of the range fell to 1.11 percentage points, and the actual value also fell outside this range 68 percent of the time.

The economic future is always uncertain, and a wider forecast range reflects greater disagreement among the policymakers about economic trends and policy objectives. However, even if the forecasts of all the policymakers were identical (perfect consensus), they would likely be wrong because the future is inherently unpredictable. Table 2 attempts to gauge the degree of consensus about the outlook, relative to the economy's unpredictability. The table shows the ratios of the RMSEs of the consensus forecasts to one-half the width of the forecast ranges. (We use one-half the range to make the scale comparable to the RMSE, which is an index of unpredictability.) The width of the range is an inverse measure of consensus. As such, the ratios reveal, in a simple way, the degree of consensus among the policymakers about the outlook for a variable, relative to the difficulty of predicting that variable. A high ratio, for example, indicates a strong consensus regarding the outlook, relative to the unpredictability of the variable.

In every case except for the 18-month inflation forecasts, the RMSE is larger than half the width of the range. Paradoxically, the highest figures (where the consensus is strong relative to the degree of predictability) are for "real" variables, such as GDP and the unemployment rate, over which the Fed has little control. At the same time, the smallest figures are observed for inflation, over which the Fed is widely regarded as having considerable influence, although perhaps not over horizons as short as 18 months. The individual inflation forecasts reflect, to a greater or lesser extent, the policymaker's beliefs about the desired long-run inflation objective. Among inflation forecast horizons, the least consensus emerges for the longest horizon—18 months—where the Fed's control presumably is the strongest. Since the formation of the European Central Bank, the United States and Japan are the only major central banks in the world that do not announce a numerical objective for inflation. The statistics in Table 2 suggest one reason why U.S. policymakers may have been reluctant to adopt a specific objective for inflation—namely, there is little consensus about what the numerical objective should be.

Table 2

A Relative Measure of FOMC Consensus

	Forecast horizon		
	6-month	12-month	18-month
Nominal GDP	1.21	1.28	1.28
Real GDP	1.40	1.76	1.72
Inflation	1.08	1.05	0.87
Unemployment	1.35	1.86	1.83

NOTE: The values in this table are calculated as the ratio of the RMSE to one-half the width of the range as reported in Table 1. It is a measure of the predictive uncertainty (RMSE) divided by an inverse measure of policy consensus. Therefore, a high number indicates a relatively high degree of consensus.

ARE THE FOMC FORECASTS UNBIASED?

There is a problem in evaluating the FOMC inflation forecasts because, as noted above, the FOMC switched among price indices over our sample period. Our analysis in Figure 1 and Tables 1, 2, and A1 used the inflation forecasts for the different price indices as they were reported. In the remaining part of the paper, we use an implied forecast that is calculated by subtracting the midpoint of the range of real GDP forecasts from the midpoint of the range of nominal GDP forecasts (Tables 3 through 5). We do this because the FOMC has consistently forecasted nominal and real output since 1979. Even when the Fed was basing the inflation forecast on the CPI (from 1989 through 1999), they were also reporting forecasts for both nominal and real output, so there was always an implied forecast for the output deflator.

In this section, we check the alternative forecasts for unbiasedness and estimate the following regression in the first part of our test:

$$(1) \qquad x_t = \alpha + \beta\, {}_{t-i}x_t^f + {}_{t-i}\varepsilon_t,$$

where x is the variable being forecast (the fourth-quarter-over-fourth-quarter growth rate of output or the price deflator for output). The forecast (x^f) is indexed according to the time when the forecast was made ($t-i$, where i refers to three forecast horizons) and the year to which it applies (t). If the estimates of (α, β) are equal to (0,1), the forecasts are unbiased. We use an F-statistic to test for unbiasedness. Holden and Peel (1990) show that even if we can reject the null hypothesis that (α, β) is equal to (0,1), it is still possible that the forecasts may be unbiased. The intuition for their result can be understood by thinking about equation (1) as a mechanism for combining unbiased forecasts where the constant is an unbiased forecast of the series. If we cannot reject the null hypothesis, we can conclude that the forecast is unbiased. If we reject the null hypothesis, it is necessary to examine the properties of the forecast error, $x_t - {}_{t-i}x_t^f$. To make a complete test of unbiasedness, we also compute the regression

$$(2) \qquad x_t - {}_{t-i}x_t^f = \gamma + {}_{t-i}\varepsilon_t^\gamma$$

and test whether γ is equal to zero. In addition, we must take account of possible serial correlation in the error for the July next-year forecast. Because the forecast horizon is longer than the interval over which output growth is measured, the forecast error for year t is not available when the forecasts for year $t+1$ are made. Therefore, information that arrives in the second half of year t may be reflected in forecast errors for both years t and $t+1$. If it does, the errors will display first-order serial correlation. For this case, Hansen (1982) has shown that ordinary least-squares (OLS) estimates will be unbiased, but the standard errors will be too small, leading to too many rejections of the null hypothesis. Therefore we use the correction for serial correlation suggested by Hansen (1982) when we report test statistics for the July next-year forecasts.

OLS estimates of equations (1) and (2) for the FOMC forecasts are listed in Table 3, which includes the estimates of α and β, with standard errors in parentheses. The third column of results shows the probability values of the F-statistic for the null joint hypothesis that (α, β) = (0,1). For the 12 cases with nominal and real output, the probability values are all above 10 percent, so we conclude that the output forecasts are unbiased. Table 2 also reports the t-statistics for estimates of γ in equation (2). It is not necessary to estimate equation (2) in the case of the output forecasts, however, because we could not reject unbiasedness using the F-statistic.

In contrast, the probability values of the F-statistic for the inflation and unemployment forecasts are quite low. In the case of inflation, the F-statistic always rejects unbiasedness at a 10 percent level. The fifth column of Table 3 reports the probability values for the t-statistic for testing whether γ is equal to zero. Here we find strong evidence that the inflation forecasts are biased, which is not too surprising since the inflation forecasts were above actual inflation most of the time (see Figure 1). In the case of the unemployment rate, we could reject the joint hypothesis that (α, β) = (0,1) in five of the six cases. In two of those, however, we could not reject that the mean of the forecast error was zero. Here, the FR forecasts for unemployment were found to be significantly biased, but the CT forecasts were not.

ARE THE FOMC FORECASTS EFFICIENT?

Forecasts are considered efficient if the forecaster takes account of all relevant information. This means that forecasters continue to invest in using more information as long as benefits of using more information exceed the costs of acquiring it. In practice, we check for informational efficiency by testing whether the forecast errors are systematically related to information that was readily available when the forecasts were made. Because the data set is small, we check for a bivariate relationship between the forecast errors and information from the past. To conduct this test for efficiency, we run the following regression:

$$(3) \qquad {}_{t-i}e_t = \alpha + \beta\, {}_{t-i}x_{t-j}^k + {}_{t-i}u_t,$$

where ${}_{t-i}e_t$ is the forecast error for year t (for each of the four variables) made at the $t-i$ horizon. So there will be a total of 24 different dependent variables: The FR and the CT each have three horizons for each of the four forecasted variables. The term ${}_{t-i}x_{t-j}^k$ is the $t-j$ information variable that was available when the forecast was made. Thus, $j=2$ when the dependent variable is an 18-month forecast error, and $j=1$ when it is a 6- or 12-month forecast error. The superscript k refers to the information variable included in our efficiency test.

In principle, we could test efficiency against

Table 3

Tests for Unbiasedness

Equation (1) $x_t = \alpha + \beta\, {}_{t-i}x_t^f + {}_{t-i}\varepsilon_t^a$

Equation (2) $(x_t - {}_{t-i}x_t^f) = \gamma + {}_{t-i}\varepsilon_t^b$

	Forecast	α (SE)	β (SE)	F-statistic (p-value)	γ (p-value)
Nominal GDP	6-month FR	–0.477 (0.86)	1.028 (0.14)	0.396	0.172
	6-month CT	–0.560 (0.80)	1.035 (0.13)	0.268	0.104
	12-month FR	0.269 (1.32)	0.925 (0.22)	0.823	0.596
	12-month CT	0.674 (1.41)	0.855 (0.23)	0.720	0.599
	18-month FR	1.444 (1.94)	0.681 (0.33)	0.326	0.245
	18-month CT	2.182 (1.72)	0.541 (0.28)	0.101	0.141
Real GDP	6-month FR	–0.301 (0.47)	1.098 (0.15)	0.799	0.960
	6-month CT	–0.312 (0.48)	1.105 (0.15)	0.781	1.000
	12-month FR	–0.068 (0.92)	1.115 (0.32)	0.677	0.412
	12-month CT	–0.697 (1.08)	1.330 (0.38)	0.542	0.484
	18-month FR	–0.124 (1.49)	1.121 (0.56)	0.829	0.560
	18-month CT	0.644 (1.67)	0.794 (0.61)	0.915	0.800
Inflation	6-month FR	–0.075 (0.35)	0.926 (0.11)	**0.028**	**0.009**
	6-month CT	–0.198 (0.35)	0.952 (0.11)	**0.009**	**0.002**
	12-month FR	0.233 (0.41)	0.790 (0.13)	**0.006**	**0.005**
	12-month CT	0.256 (0.43)	0.788 (0.13)	**0.012**	**0.009**
	18-month FR	0.115 (0.54)	0.775 (0.16)	**0.002**	**0.001**
	18-month CT	0.592 (0.59)	0.615 (0.17)	**0.001**	**0.003**
Unemployment	6-month FR	0.395 (0.40)	0.909 (0.06)	**0.080**	**0.072**
	6-month CT	0.671 (0.35)	0.868 (0.06)	**0.036**	0.152
	12-month FR	0.798 (0.56)	0.824 (0.09)	**0.019**	**0.033**
	12-month CT	1.144 (0.57)	0.779 (0.09)	**0.032**	0.160
	18-month FR	1.425 (0.96)	0.713 (0.16)	**0.058**	**0.099**
	18-month CT	1.480 (1.01)	0.713 (0.17)	0.122	0.218

NOTE: FR represents the forecasts calculated as the midpoint of the full range and CT represents the forecasts calculated as the midpoint of the CT. The tests use the sample period for which the CT is available: 1983-2001 for 6- and 12-month forecasts and 1984-2001 for 18-month forecasts. Bold cells indicate that the probability value is less than 10 percent.

any information that was available at the time the forecasts were made. In Table 4A, we include the most recently available information on the variables being forecasted—nominal GDP growth, real GDP growth, inflation, and the unemployment rate. The table includes the p-values for the t-statistic testing the null hypothesis that $\beta = 0$. We also derive Q-statistics (from the regression of the forecast error on the last observed value of the forecasted variable) to test for serial correlation in the forecast errors. For example, in Table 4A, the first row of results are calculated using the 18-month nominal GDP forecast error. The Q-statistic is taken from the equation that includes past nominal GDP growth on the right-hand side. In the second row of each panel, the forecast error for real GDP growth is the dependent variable and the Q-statistic is taken from the equation that includes past GDP growth as the information variable.

The top panel of Table 4A includes tests using the 18-month forecasts. The July next-year forecast is made before the current-year results are known,

Table 4A

Are the FOMC Forecasts Efficient? History of Forecast Variables

(Using May vintage for July forecasts, values in the table are probability values for the t-statistic on the coefficient β in equation (3): ${}_{t-i}e_t = \alpha + \beta\, {}_{t-i}x^k_{t-j} + {}_{t-i}u_t$.)

	Nominal GDP	Real GDP	Inflation	Unemployment rate	Q*
6-month forecast errors					
FR					
Nominal GDP	0.25	0.64	0.23	0.67	**0.00**
Real GDP	0.10	0.30	0.28	0.49	**0.01**
Inflation	0.58	0.38	0.58	0.70	0.11
Unemployment rate	0.22	0.20	0.93	0.08	**0.04**
CT					
Nominal GDP	0.14	0.56	0.11	0.64	**0.00**
Real GDP	**0.03**	0.21	0.15	0.41	**0.03**
Inflation	0.34	0.22	0.66	0.55	0.16
Unemployment rate	0.12	**0.08**	0.75	**0.01**	0.39
12-month forecast errors					
FR					
Nominal GDP	0.21	0.58	0.25	0.70	**0.03**
Real GDP	0.27	0.46	0.57	0.28	**0.06**
Inflation	0.67	0.71	0.14	0.12	0.11
Unemployment rate	0.52	0.48	0.90	0.05	**0.05**
CT					
Nominal GDP	**0.10**	0.38	0.23	0.60	**0.01**
Real GDP	0.21	0.48	0.37	0.21	**0.09**
Inflation	0.24	0.57	0.32	0.15	**0.01**
Unemployment rate	0.42	0.44	0.94	0.02	0.22
18-month forecast errors					
FR					
Nominal GDP	**0.01**	**0.08**	0.30	0.82	0.38
Real GDP	**0.08**	0.17	0.51	0.34	0.23
Inflation	**0.10**	0.28	0.34	0.11	0.16
Unemployment rate	0.17	0.23	0.73	0.03	**0.01**
CT					
Nominal GDP	**0.02**	0.17	**0.10**	0.86	0.47
Real GDP	0.10	0.22	0.48	0.46	0.18
Inflation	0.11	0.62	**0.05**	**0.07**	**0.04**
Unemployment rate	0.19	0.31	0.60	0.05	**0.01**

*Q is a test for randomness based on the first four autocorrelations of the residual in the regression with the lag of the forecast variable. For example, in the first row of each panel we use the residuals from the equation including lagged nominal GDP growth. Bold cells indicate that the probability value is less than 10 percent.

Table 4B

Are the FOMC Forecasts Efficient? Interest Rates and Money

(Using June interest rates for July forecasts and January rates for February forecasts, values in the table are probability values for the t-statistic on the coefficient β in equation (3): ${}_{t-i}e_t = \alpha + \beta\, {}_{t-i}x^k_{t-j} + {}_{t-i}u_t$.)

	10-year Treasury bond	1-year Treasury bill	Overnight federal funds rate	Monetary aggregates		
				M1	M2	M3
6-month forecast errors						
FR						
Nominal GDP	0.78	0.96	0.86	0.40	0.86	0.73
Real GDP	0.70	0.66	0.49	0.54	0.82	0.62
Inflation	0.89	0.47	0.37	0.54	0.94	0.84
Unemployment rate	0.66	0.47	0.57	0.15	0.96	0.70
CT						
Nominal GDP	0.53	0.65	0.53	0.46	0.89	0.68
Real GDP	0.48	0.42	0.26	0.42	1.00	0.89
Inflation	0.91	0.51	0.36	0.96	0.76	0.54
Unemployment rate	0.57	0.57	0.76	**0.07**	0.98	0.70
12-month forecast errors						
FR						
Nominal GDP	0.35	0.32	0.34	0.99	0.54	0.42
Real GDP	0.67	0.57	0.59	0.94	0.46	0.36
Inflation	0.20	0.29	0.30	0.87	0.83	0.90
Unemployment rate	0.67	0.83	0.99	0.42	0.65	0.67
CT						
Nominal GDP	0.27	0.25	0.26	0.99	0.59	0.45
Real GDP	0.59	0.46	0.46	0.92	0.34	0.26
Inflation	0.12	0.24	0.25	0.83	0.46	0.55
Unemployment rate	0.68	0.99	0.79	0.27	0.56	0.63
18-month forecast errors						
FR						
Nominal GDP	0.17	0.13	0.11	0.83	0.99	0.96
Real GDP	0.54	0.37	0.26	0.87	0.91	0.70
Inflation	**0.06**	0.11	0.23	0.40	0.77	0.46
Unemployment rate	0.94	0.54	0.38	0.13	0.81	0.70
CT						
Nominal GDP	**0.07**	**0.05**	**0.05**	0.93	0.87	0.96
Real GDP	0.45	0.30	0.21	0.60	0.95	0.72
Inflation	**0.03**	**0.06**	0.14	0.26	0.83	0.58
Unemployment rate	0.79	0.44	0.30	0.14	0.92	0.76

*Q is a test for randomness based on the first four autocorrelations of the residual in the regression with the lag of the forecast variable. For example, in the first row we use the residuals from the equation including lagged nominal GDP growth. Bold cells indicate that the probability value is less than 10 percent.

so in this case we use the results from the previous year. We are testing whether the 18-month forecast errors from year t are related to the variables dated t–2. We report results using both the FR and the CT forecasts. In the case of the FR, coefficients on nominal GDP growth, real GDP growth, and inflation are statistically significant at a 10 percent critical level. In the case of the CT forecasts, the nominal GDP forecast errors are related to past nominal GDP growth and past inflation. Inflation forecast errors are related to past inflation and to the past unemployment rate.

The 12-month forecasts are shown in the middle panel. There is only one significant coefficient among the 24 coefficients reported. The CT forecast for nominal GDP growth is correlated with last year's nominal GDP growth. However, six of the eight Q-statistics have probability values less than 10 percent. In the case of the 6-month forecasts, none of the FR forecasts were related to past values of the information variables, but both the nominal and real GDP forecasts errors were serially correlated.

Since the FOMC was also setting targets for monetary growth and the federal funds rate for much of this period, the forecasts were checked for efficiency with respect to information about monetary growth and interest rates. The FOMC set targets for M1, M2, and M3; all were included. Also included were the currently observed interest rates for the overnight federal funds rate, 1-year Treasury bills, and 10-year Treasury bonds. Although the FOMC does not set targets for Treasury rates, they embody expectations about future inflation and future overnight interest rates.

The results including money and interest rates are shown in Table 4B. The monetary data were those reported to the FOMC at the first policy meeting of the year. They are fourth-quarter-over-fourth-quarter growth rates. The interest rates are averages for the months of January (used with the February forecasts) and June (used with the July forecasts). Again, there is almost no evidence that this information was ignored. An exception may be the 18-month CT forecasts for nominal GDP and inflation. Here, the forecast errors for nominal GDP growth and inflation are significantly related to interest rates. Only in one case, the 6-month forecast error for the unemployment rate, was past money growth significant.

With some exceptions, the forecast errors appear to be serially correlated, but unrelated to our information variables. When the forecasts fail tests for efficiency, it is usually a forecast calculated from the CT rather than the FR. We estimated 240 coefficients on past information. We found that 19 of these were significant at a 10 percent level; of those 19, only 4 applied to cases using an FR forecast. Overall, there were fewer significant coefficients than one would expect to observe by chance, even if the forecasts were fully efficient.

DO EITHER THE RANGE OR CENTRAL TENDENCY FORECASTS ENCOMPASS THE OTHER?

In this section, we test whether the CT or the FR encompasses the other. We say that one forecast encompasses another if the former incorporates all of the relevant information of the latter. If it does, then adding information from the latter will not help predict the actual value. We run the following regression:

$$x_t - {}_{t-i}x_t^{fr} = c + \beta_{ct}\, {}_{t-i}x_t^{ct} + {}_{t-i}\varepsilon_t^{fr}, \tag{4}$$

where we compare the FR and CT forecasts of x, x^{fr}, and x^{ct}. Forecast x^{fr} encompasses x^{ct} if $\beta_{ct} = 0$. We test the alternative, that is, forecast x^{ct} encompasses x^{fr}, by switching the roles of the FR and CT forecasts. The results of the encompassing tests are shown in Table 5. The results are p-values for the t-statistic testing that the coefficient β is equal to zero. In all cases except the unemployment forecasts, the probability values are greater than 10 percent and we cannot reject that they both encompass the other. In practical terms, this means that if you have one forecast, you do not need the other. In the case of the unemployment rate, we can reject that the 18- and 12-month FR forecasts encompass the CT tendency forecasts, and that the 12- and 6-month CT forecasts encompass the respective FR forecasts. In practical terms, this means that there is useful and independent information in both forecasts.

CONCLUSIONS

This article reviews two decades of experience with FOMC forecasts. As one would expect, forecasts calculated from the FR are very similar to those calculated from the CT. In the two cases where there is a significant difference, the forecasts constructed from the FR were more accurate than those calculated from the CT. In two cases where the FOMC was forecasting the next calendar year in July, we found that the nominal GDP and inflation forecasts that

Table 5

Encompassing Tests for the FOMC Forecasts: FR and CT

(Probability values for the t-statistic on β in equations (4): $x_t - {}_{t-i}x_t^{fr} = c + \beta_{ct}\, {}_{t-i}x_t^{ct} + {}_{t-i}\varepsilon_t^{fr}$ and $x_t - {}_{t-i}x_t^{ct} = c + \beta_{fr}\, {}_{t-i}x_t^{fr} + {}_{t-i}\varepsilon_t^{ct}$.)

	Nominal GDP forecasts		Real GDP forecasts	
Forecast horizon	FR encompasses CT	CT encompasses FR	FR encompasses CT	CT encompasses FR
6-month	0.68	0.84	0.45	0.41
12-month	0.57	0.80	0.64	0.37
18-month	0.31	0.17	0.83	0.69

	Inflation forecasts		Unemployment forecasts	
	FR encompasses CT	CT encompasses FR	FR encompasses CT	CT encompasses FR
6-month	0.73	0.78	0.23	**0.03**
12-month	0.18	0.28	**0.05**	**0.04**
18-month	0.11	0.26	**0.07**	0.14

NOTE: Bold cells indicate the null hypothesis (that the FR encompasses the CT or vice versa) is rejected with a 10 percent critical region.

included the outliers were significantly more accurate than those taken from the CT. We also found that the sizes of the ranges reported were not very closely related to the inherent unpredictability in a variable. Nominal and real output growth rates were more variable and less predictable than inflation, but the range for the high and low 18-month inflation forecasts were larger than those for the output measures. One inference is that there was a wide difference of opinion among FOMC members regarding the medium-term outlook for inflation, relative to the underlying predictability of inflation. This may explain, in part, the FOMC's reluctance to publicly commit to an inflation target.

We could not reject the hypothesis that the FOMC forecast was unbiased in any case involving nominal or real output growth. In contrast, all six forecasts of inflation were significantly biased. In the case of the unemployment rate, we could reject unbiasedness for all of the forecasts calculated from the FR, but for none calculated from the CT.

In general, the FOMC passed tests for efficiency when the tests involved past values of nominal and real output growth, inflation, the unemployment rate, interest rates, and monetary growth. To the extent that the CT and the FR were different, the CT failed the tests more often, especially in the cases using the 18-month forecasts. The encompassing tests did not distinguish between the CT and the FR forecasts.

Overall, the FR forecasts produced smaller RMSEs and were more efficient than the CT forecasts. For the purposes of doing research using the FOMC's GDP growth and inflation forecasts, the FR is almost certainly preferred, since it did at least as well on every count as the CT. Furthermore, the Fed published the FR from July 1979 on, whereas the CT did not become available until 1983.

REFERENCES

Bernanke, Ben S.; Laubach, Thomas; Mishkin, Frederic S. and Posen, Adam S. *Inflation Targeting: Lessons from the International Experience*. Princeton: Princeton University Press, 1998.

Diebold, Francis X. and Mariano, Roberto S. "Comparing Predictive Accuracy." *Journal of Business and Economic Statistics*, July 1995, *13*(3), pp. 253-65.

Faust, Jon; Rogers, John H. and Wright, Jonathan H. "News and Noise in G-7 GDP Announcements." *International Finance Discussion Papers*, Board of Governors of the Federal Reserve System, No. 690, December 2000.

Gavin, William T. and Mandal, Rachel J. "Evaluating FOMC Forecasts." Working Paper 2001-005C, Federal Reserve Bank of St. Louis, August 2002 (forthcoming in *International Journal of Forecasting*).

Hansen, Lars Peter. "Large Sample Properties of Generalized Method of Moments Estimators." *Econometrica*, July 1982, *50*(4), pp. 1029-54.

Holden, Kenneth and Peel, David A. "On Testing for Unbiasedness and Efficiency of Forecasts." *Manchester School of Economics and Social Studies*, June 1990, *58*(2), pp. 120-27.

Appendix

The FOMC forecasts are taken directly from the monetary policy reports to Congress. The first report was made in February 1979, but forecasts were not included with the initial release of the report.

Tables A1 through A4 include the nominal GDP growth, real GDP growth, inflation, and fourth-quarter average unemployment rate forecasts. These forecasts come from the semiannual monetary policy reports made by the Fed. The inflation forecasts in Table A3 include forecasts for a mix of price indices, the output deflator, the CPI, and the chain price index for PCE.

Tables A5 and A6 include the real-time data from the Federal Reserve Bank of Philadelphia web site. The data are constructed from the quarterly data files that include vintages of data that were available in the middle of each indicated month—February, May, August, and November. For example, the February 1980 file contains the data that were available to the FOMC in 1980, when the Fed made its February 1980 monetary policy report. The exception was for 1995. In early 1996, the Bureau of Economic Analysis switched to the chain weighted indices for GDP and delayed publishing the initial release until March of 1996. Therefore, the February data for the fourth-quarter-over-fourth-quarter value for 1995 comes from the March issue of the *Survey of Current Business* (Table 7.1).

Table A7 includes the fourth-quarter-over-fourth-quarter growth rates of the monetary aggregates as they were reported following the annual benchmarking in February each year. This is the real-time monetary data that the FOMC would have had at the first FOMC meeting each year. This table also includes the real-time CPI data and PCE chain price data. The monetary growth rates and CPI are taken from the monetary policy reports. The real-time price data for PCE are from the February release of the *Survey of Current Business*. Table A8 includes the January and June monthly average interest rates for overnight federal funds, 1-year Treasury bills, and 10-year Treasury bonds. These are from the Fed's H.15 release.

Table A1

FOMC Nominal GNP/GDP Forecasts by Forecast Horizon

	6-month		12-month		18-month	
	FR	CT	FR	CT	FR	CT
1979	8 to 10					
1980	5 to 7.5		7.5 to 11		8.5 to 11.5	
1981	10 to 11.5		9 to 12		8.5 to 11.5	
1982	5.5 to 7.5		8 to 10.5		9.5 to 12.25	
1983	9.25 to 10.75	9.75 to 10	7.25 to 11.25	8 to 9	7 to 9.5	
1984	9.5 to 11.5	10.5 to 11	8 to 10.5	9 to 10	7 to 10.25	9 to 10
1985	6.25 to 7.75	6.5 to 7	7 to 8.5	8.5 to 8	6.75 to 9.5	8 to 9
1986	3.75 to 6.5	4.75 to 5.75	5 to 8.5	6.5 to 7.25	5.5 to 8.5	7 to 7.5
1987	5.75 to 7.25	6.25 to 7	4.5 to 7.5	5.75 to 6.5	5 to 8.25	6 to 7.5
1988	4 to 7	5.75 to 6.75	4 to 6.5	5.25 to 6	5 to 8	5.75 to 7
1989	5 to 7.75	6 to 7	5.5 to 8.5	6.5 to 7.5	4 to 7.5	5 to 7
1990	5 to 6.5	5.5 to 6.5	4 to 7	5.5 to 6.5	4.25 to 7.5	5.5 to 6.75
1991	3.75 to 5.75	4.5 to 5.25	3.5 to 5.5	3.75 to 5.25	3.5 to 7	5.25 to 6.5
1992	5 to 6.25	5.25 to 6	4.5 to 5.75	5 to 6.25	4 to 6.75	5.5 to 6.5
1993	4.75 to 6.25	5 to 5.75	5.25 to 6.25	5.5 to 6	4.5 to 7	5.5 to 6.25
1994	5.25 to 6.5	5.5 to 6	4.75 to 7.5	5.5 to 6	4.5 to 6.75	5 to 6.5
1995	3.75 to 5.25	4.25 to 4.75	4.75 to 6.5	5 to 6	4.5 to 6.25	5 to 5.5
1996	4.75 to 5.75	5 to 5.5	4 to 5	4.25 to 4.75	4.625 to 5.5	4.75 to 5.375
1997	5 to 6	5 to 5.5	4.25 to 5.25	4.5 to 4.75	4 to 5.5	4.25 to 5
1998	4.25 to 5	4.5 to 5	3.5 to 5	3.75 to 4.5	4.25 to 5.75	4.5 to 5
1999	4.75 to 5.5	5 to 5.5	3.75 to 5	4 to 4.5	4 to 5.5	4.25 to 5
2000	6 to 7.25	6.25 to 6.75	5 to 6	5.25 to 5.5	4 to 5.25	4 to 5
2001	3.25 to 5	3.5 to 4.25	3.75 to 5.25	4 to 5	5 to 6.25	5.5 to 6

NOTE: The date refers to the calendar year being forecast. For example, the 18-month FR forecast for 1980 (8.5 to 11.5) was made in July 1979.

Table A2

FOMC Real GNP/GDP Forecasts by Forecast Horizon

	6-month		12-month		18-month	
	FR	CT	FR	CT	FR	CT
1979	–2 to –0.5					
1980	–5 to –2.5		–2.5 to 0.5		–0.5 to 2	
1981	1 to 3.5		–1.5 to 1.5		0.5 to 3	
1982	0.5 to 1.5		0.5 to 3		1 to 4	
1983	4.75 to 6	5 to 5.75	3 to 5.5	3.5 to 4.5	2.5 to 4	
1984	6 to 7	6.25 to 6.75	3.5 to 5	4 to 4.75	3 to 5	4 to 4.5
1985	2.25 to 3.25	2.75 to 3	3.25 to 4.25	3.5 to 4	2 to 4	3 to 3.25
1986	2.25 to 3.5	2.5 to 3	2.75 to 4.25	3 to 3.5	2 to 4	2.5 to 3.25
1987	2 to 3.75	2.5 to 3	2 to 4	2.5 to 3	2 to 4.25	3 to 3.5
1988	1 to 3.25	2.75 to 3	0.5 to 3	2 to 2.5	1 to 3	2.5 to 3
1989	1.5 to 2.75	2 to 2.5	1.5 to 3.25	2.5 to 3	1 to 3	2 to 2.5
1990	1 to 2	1.5 to 2	1 to 2.25	1.75 to 2	1 to 2.5	1.5 to 2
1991	0.5 to 1.5	0.75 to 1	–0.5 to 1.5	0.75 to 1.5	0 to 3	1.75 to 2.5
1992	2 to 3.25	2.25 to 2.75	1.5 to 2.75	1.75 to 2.5	2 to 3.5	2.25 to 3
1993	2 to 3.5	2.25 to 2.75	2.5 to 4	3 to 3.25	2.5 to 3.5	2.75 to 3
1994	3 to 3.5	3 to 3.25	2.5 to 3.75	3 to 3.25	2 to 3.25	2.5 to 3.25
1995	1.375 to 3	1.5 to 2	2 to 3.25	2 to 3	2.25 to 2.75	2.5 to 2.75
1996	2.5 to 3	2.5 to 2.75	1.5 to 2.5	2 to 2.25	2.125 to 3	2.25 to 2.75
1997	3 to 3.5	3 to 3.25	2 to 2.5	2 to 2.25	1.5 to 2.5	1.75 to 2.25
1998	2.75 to 3.25	3 to 3.25	1.75 to 3	2 to 2.75	2 to 3	2 to 2.5
1999	3.25 to 4	3.5 to 3.75	2 to 3.5	2.5 to 3	2 to 3	2 to 2.5
2000	3.75 to 5	4 to 4.5	3.25 to 4.25	3.5 to 3.75	2 to 3.5	2.5 to 3
2001	1 to 2	1.25 to 2	2 to 2.75	2 to 2.5	2.5 to 4	3.25 to 3.75

NOTE: See note for Table A1.

Table A3

FOMC Inflation Forecasts by Forecast Horizon

	6-month		12-month		18-month	
	FR	CT	FR	CT	FR	CT
1979	9.5 to 11					
1980	9 to 10		9 to 11		8.5 to 10.5	
1981	7.5 to 9		9 to 10.5		7.75 to 9.5	
1982	4.75 to 6		6.5 to 7.75		6.5 to 8.5	
1983	4 to 5.25	4.25 to 4.75	3.5 to 5.5	4 to 5	4 to 5.75	
1984	3.25 to 4.5	4 to 4.5	4 to 6	4.5 to 5	3.25 to 6.5	4.25 to 5
1985	3.5 to 4.25	3.75 to 4	3 to 4.75	3.5 to 4	3.5 to 6.5	5.25 to 5.5
1986	1.5 to 3.25	2.25 to 2.75	2.5 to 4.5	3 to 4	3 to 5.5	3.75 to 4.75
1987	3 to 4.25	3.5 to 4	2.5 to 4	3 to 3.5	1.5 to 4.25	3 to 4
1988	2.75 to 4	3 to 3.75	2.5 to 4	3.25 to 3.75	2.5 to 5	3.75 to 4.25
1989	4.5 to 5.75	5 to 5.5	3.5 to 5.5	4.5 to 5	2 to 5	3 to 4.5
1990	4 to 5	4.5 to 5	3.5 to 5	4 to 4.5	3 to 5.75	4.5 to 5
1991	3 to 4.5	3.25 to 3.75	3 to 4.5	3.25 to 4	3.5 to 5	3.75 to 4.5
1992	3 to 3.5	3 to 3.5	2.5 to 3.5	3 to 3.5	2.5 to 4.25	3 to 4
1993	3 to 3.5	3 to 3.25	2.5 to 3	2.5 to 2.75	2.5 to 4	2.75 to 3.25
1994	2.5 to 3.5	2.75 to 3	2.25 to 4	3 to 3	2 to 4.25	3 to 3.5
1995	3 to 3.5	3.125 to 3.375	2.75 to 3.75	3 to 3.5	2 to 4.5	2.75 to 3.5
1996	3 to 3.25	3 to 3.25	2.5 to 3	2.75 to 3	2.5 to 3.5	2.875 to 3.25
1997	2 to 2.75	2.25 to 2.5	2.75 to 3.5	2.75 to 3	2.5 to 3.25	2.75 to 3
1998	1.25 to 2.25	1.75 to 2	1.5 to 2.5	1.75 to 2.25	2.5 to 3	2.5 to 3
1999	1.75 to 2.5	2.25 to 2.5	1.5 to 2.5	2 to 2.5	1.5 to 3	2 to 2.5
2000	2 to 2.75	2.5 to 2.75	1.5 to 2.5	1.75 to 2	1.75 to 2.5	2.25 to 2.5
2001	2 to 2.75	2 to 2.5	1.75 to 2.5	1.75 to 2.25	1.75 to 3	2 to 2.5

NOTE: From 1979 through 1988, the FOMC reported forecasts for the GNP/GDP deflator; from 1989 through 1999, the FOMC reported forecasts for the CPI; and in 2000 and 2001, the FOMC reported forecasts for the chain price index for PCE. See, also, note for Table A1.

Table A4

FOMC Fourth-Quarter Average Unemployment Rate Forecasts

	6-month		12-month		18-month	
	FR	CT	FR	CT	FR	CT
1979	6.25 to 7					
1980	8.5 to 9.25		6.75 to 8		6.75 to 8.25	
1981	7.5 to 8.25		8 to 8.5		8 to 9.25	
1982	9 to 9.75		8.25 to 9.5		7 to 8.5	
1983	9 to 9.75	9.5 to 9.5	9.5 to 10.5	9.9 to 10.4	8.5 to 9.5	
1984	6.5 to 7.25	6.75 to 7	7.25 to 8	7.25 to 7.75	8.25 to 9.25	8.25 to 8.75
1985	6.75 to 7.25	7 to 7.25	6.5 to 7.25	6.75 to 7	6.25 to 7.25	6.5 to 7
1986	6.9 to 7.2	7 to 7	6.25 to 6.75	6.5 to 6.5	6.75 to 7.5	6.75 to 7.25
1987	6.1 to 6.5	6.2 to 6.4	6.5 to 6.75	6.5 to 6.75	6.5 to 7	6.75 to 6.75
1988	5.25 to 6.5	5.25 to 5.75	5.5 to 6.75	5.75 to 6	5.9 to 6.8	6 to 6.5
1989	5 to 6	5.5 to 5.5	5 to 6	5.25 to 5.25	5 to 7	5.5 to 6
1990	5.5 to 6.5	5.5 to 5.75	5.5 to 6.5	5.5 to 5.75	5 to 6.5	5.5 to 6
1991	6.5 to 7	6.75 to 7	6.25 to 7.5	6.5 to 7	5.25 to 7	5.5 to 6
1992	7 to 7.5	7.25 to 7.5	6.75 to 7.25	6.75 to 7	6 to 6.75	6.25 to 6.5
1993	6.5 to 7	6.75 to 6.75	6.5 to 7	6.75 to 7	6.5 to 7.25	6.5 to 7
1994	6 to 6.25	6 to 6.25	6.5 to 6.75	6.5 to 6.75	6.25 to 7	6.5 to 6.75
1995	5.5 to 6.25	5.75 to 6.125	5.25 to 6	5.5 to 5.5	5.75 to 6.5	6 to 6.25
1996	5.25 to 5.75	5.5 to 5.5	5.5 to 6	5.5 to 5.75	5.5 to 6.25	5.75 to 6.125
1997	4.75 to 5.25	4.75 to 5	5.25 to 5.5	5.25 to 5.5	5.5 to 6	5.5 to 5.75
1998	4.25 to 4.5	4.25 to 4.5	4.5 to 5	4.75 to 4.75	4.5 to 5.25	4.75 to 5
1999	4 to 4.5	4 to 4.25	4.25 to 4.75	4.25 to 4.5	4.25 to 4.75	4.5 to 4.75
2000	4 to 4.25	4 to 4	4 to 4.25	4 to 4.25	4 to 4.5	4 to 4.25
2001	4.75 to 5	4.75 to 5	4.5 to 5	4.5 to 4.5	4 to 4.5	4 to 4.25

NOTE: See note for Table A1.

Table A5

Real-Time Data Available in February and May

	February				May			
	Nominal output growth	Real output growth	Inflation	Unemployment rate	Nominal output growth	Real output growth	Inflation	Unemployment rate
1978	12.91	4.26	8.29	5.83	13.11	4.44	8.30	5.83
1979	9.87	0.83	8.97	5.87	9.92	0.96	8.87	5.87
1980	9.82	–0.03	9.86	7.50	9.39	–0.34	9.75	7.50
1981	9.31	0.67	8.58	8.37	9.80	0.86	8.87	8.37
1982	3.27	–1.23	4.56	10.67	3.50	–0.87	4.40	10.67
1983	10.37	6.06	4.06	8.47	10.50	6.20	4.05	8.47
1984	9.35	5.62	3.53	7.20	9.53	5.70	3.62	7.20
1985	5.78	2.54	3.15	7.00	5.37	2.14	3.16	7.00
1986	4.42	2.21	2.16	6.83	4.18	2.04	2.10	6.83
1987	7.23	3.85	3.25	5.90	7.45	4.00	3.31	5.90
1988	7.02	2.71	4.19	5.33	7.23	2.81	4.29	5.33
1989	6.37	2.43	3.85	5.30	6.44	2.57	3.77	5.30
1990	4.34	0.35	3.98	5.90	4.50	0.49	3.99	5.90
1991	3.22	0.23	2.99	6.97	3.28	0.27	3.00	6.97
1992	5.36	2.92	2.37	7.33	5.71	3.15	2.48	7.33
1993	5.11	2.84	2.21	6.53	5.36	3.10	2.19	6.53
1994	6.38	4.00	2.28	5.57	6.47	4.14	2.24	5.57
1995	3.68*	1.27*	2.37*	5.57	3.68	1.27	2.37	5.57
1996	5.18	3.36	1.76	5.27	4.97	3.14	1.77	5.27
1997	5.76	3.89	1.80	4.70	5.58	3.74	1.77	4.70
1998	5.07	4.15	0.88	4.40	5.17	4.25	0.88	4.40
1999	5.92	4.25	1.60	4.10	6.26	4.60	1.59	4.10
2000	5.92	3.49	2.34	3.97	5.80	3.41	2.31	3.97
2001	1.94	0.13	1.81	5.60	2.34	0.48	1.86	5.60

NOTE: Fourth-quarter-over-fourth-quarter growth rates.

*February real-time values for 1995 were not actually released until March. The values here are taken from the *Survey of Current Business*, March 1996, Table 7.1.

Table A6

Real-Time Data Available in August and November

	August				November			
	Nominal output growth	Real output growth	Inflation	Unemployment rate	Nominal output growth	Real output growth	Inflation	Unemployment rate
1978	13.39	4.80	8.20	5.83	13.39	4.80	8.20	5.83
1979	9.92	0.96	8.87	5.87	9.92	0.96	8.87	5.87
1980	9.39	–0.34	9.75	7.50	9.39	–0.34	9.75	7.50
1981	9.63	0.72	8.84	8.37	9.63	0.72	8.84	8.37
1982	2.55	–1.74	4.37	10.67	2.55	–1.74	4.37	10.67
1983	10.36	6.35	3.77	8.47	10.36	6.35	3.77	8.47
1984	9.53	5.70	3.62	7.20	9.53	5.70	3.62	7.20
1985	6.30	2.89	3.31	7.00	6.30	2.89	3.31	7.00
1986	4.48	2.21	2.22	6.83	4.48	2.21	2.22	6.83
1987	8.32	5.04	3.12	5.90	8.32	5.04	3.12	5.90
1988	7.53	3.40	4.00	5.33	7.53	3.40	4.00	5.33
1989	5.58	1.82	3.69	5.30	5.58	1.82	3.69	5.30
1990	4.50	0.49	3.99	5.90	4.50	0.49	3.99	5.90
1991	3.45	0.10	3.35	6.97	3.45	0.10	3.35	6.97
1992	5.71	3.15	2.48	7.33	6.73	3.87	2.76	7.33
1993	5.01	3.11	1.84	6.53	5.01	3.11	1.84	6.53
1994	6.47	4.14	2.24	5.57	6.47	4.14	2.24	5.57
1995	3.78	1.30	2.45	5.57	3.78	1.30	2.45	5.57
1996	5.57	3.25	2.24	5.27	5.57	3.25	2.24	5.27
1997	5.58	3.83	1.69	4.70	5.58	3.83	1.69	4.70
1998	5.17	4.25	0.88	4.40	5.85	4.61	1.18	4.40
1999	6.52	4.96	1.48	4.10	6.52	4.96	1.48	4.10
2000	5.31	2.81	2.43	3.97	5.31	2.81	2.43	3.97
2001	2.00	0.05	1.95	5.60				

NOTE: Fourth-quarter-over-fourth-quarter growth rates.

Table A7

Real-Time Data for Monetary Aggregates and Other Price Indices

	Output forecasts			CPI: Monetary policy reports	PCE chain price index: Bureau of Economic Analysis
	M1 growth	M2 growth	M3 growth		
1978					
1979					
1980	7.3	9.8	9.9		
1981	2.3	9.4	11.4		
1982	8.5	9.2	10.1		
1983	7.2	8.3	9.7		
1984	5.2	7.7	10.5		
1985	12.7	8.6	7.4		
1986	15.2	8.9	8.8		
1987	6.2	4	5.4		
1988	4.3	5.3	6.2		
1989	0.6	4.6	3.3	4.5	
1990	4.2	3.9	1.8	6.3	
1991	8	3.1	1.3	2.9	
1992	14.3	1.9	0.5	3.0	
1993	10.5	1.4	0.6	2.7	
1994	2.3	1	1.4	2.6	
1995	–1.8	4.2	6.1	2.6	
1996	–4.5	4.6	6.8	3.1	
1997	–1.2	5.7	8.8	1.9	
1998	1.8	8.5	10.9	1.5	
1999	1.9	6.2	7.5	2.6	
2000	–1.8	5.6	8.9	3.4	2.4
2001	6.9	10.3	12.9		1.3

NOTE: Fourth-quarter-over-fourth-quarter growth rates.

Table A8

H.15 Interest Rate Data

	January average			June average		
	10-year Treasury bond	1-year Treasury bill	Overnight federal funds rate	10-year Treasury bond	1-year Treasury bill	Overnight federal funds rate
1978	7.96	7.28	6.70	8.46	8.09	7.60
1979	9.10	10.41	10.07	8.91	9.57	10.29
1980	10.80	12.06	13.82	9.78	8.16	9.47
1981	12.57	14.08	19.08	13.47	14.86	19.10
1982	14.59	14.32	13.22	14.30	14.07	14.15
1983	10.46	8.62	8.68	10.85	9.66	8.98
1984	11.67	9.90	9.56	13.56	12.08	11.06
1985	11.38	9.02	8.35	10.16	7.80	7.53
1986	9.19	7.73	8.14	7.80	6.73	6.92
1987	7.08	5.78	6.43	8.40	6.80	6.73
1988	8.67	6.99	6.83	8.92	7.49	7.51
1989	9.10	9.05	9.12	8.28	8.44	9.53
1990	8.21	7.92	8.23	8.48	8.10	8.29
1991	8.09	6.64	6.91	8.28	6.36	5.90
1992	7.03	4.15	4.03	7.26	4.17	3.76
1993	6.60	3.50	3.02	5.96	3.54	3.04
1994	5.75	3.54	3.05	7.10	5.27	4.25
1995*	7.78	7.05	5.53	6.17	5.64	6.00
1996	5.65	5.09	5.56	6.91	5.81	5.27
1997	6.58	5.61	5.25	6.49	5.69	5.56
1998	5.54	5.24	5.56	5.50	5.41	5.56
1999	4.72	4.51	4.63	5.90	5.10	4.76
2000	6.66	6.12	5.46	6.10	6.17	6.53
2001	5.16	4.81	5.98	5.28	3.58	3.97

Strategic Interaction in Tax Policies Among States

Rubén Hernández-Murillo

In this paper we test for strategic interaction among U.S. states in the determination of tax rates on capital income. We find that states have a positively sloped reaction function to the tax policies of rival states when tax rates are chosen simultaneously. We also identify that a state's size has a positive effect on local tax rates.

A large literature in public economics is devoted to the study of competition among governments. The problem is interesting because competition among public agents—governments—differs importantly from many aspects of what we know about competition among private agents—consumers or firms—both in terms of equilibrium allocations and welfare outcomes. Underlying this problem are the effects of competition on governments' behavior and its normative implications, in particular how it affects the size of governments and the allocation of public funds for the provision of public goods.

Examples of competition in tax policies among governments include the following: competition for mobile capital, or *tax competition*; competition for mobile shoppers, or *commodity tax competition*; and competition for mobile firms. These examples are characterized by the mobility of the tax base in response to tax differentials among jurisdictions. In the presence of this phenomenon, a state government's choice of tax instruments is crucial for the determination of stable sources of revenues.

Competition among governmental bodies can also take many forms. For example, it can be *horizontal*: among states within a federation, among independent countries, or among regions within an economic union. It can also be *vertical*: between local and state governments or between state and federal governments.

As motivation for the empirical analysis, we focus on the literature on tax competition, reviewed recently by Wilson (1999). In this framework, governments determine tax rates on capital. The equilibrium net rate of return on capital income determines the configuration of the mobile resource, the capital stock, among the jurisdictions. Revenues from capital taxation are used to finance local public goods. In the standard model there is a large number of jurisdictions, which are unable to affect the net rate of return on capital. Thus there is no strategic interaction among governments. Mintz and Tulkens (1986) and Wildasin (1988) first analyzed competition among a small number of jurisdictions.[1] Allowing for a small number of jurisdictions gives individual governments the ability to affect the net rate of return by exercising market power, and strategic considerations are important—as in oligopolistic competition among firms.

In all these cases, the prevailing equilibrium in tax rates results in an underprovision of public goods because jurisdictions set tax rates that are inefficiently low in an attempt to retain their tax base. This occurs because, when one jurisdiction raises its tax rate, it generates a positive *fiscal* externality on its rivals (by expanding their tax base with businesses or individuals fleeing the high-tax jurisdiction). Early studies examine the features of the equilibrium among identical jurisdictions. When asymmetries are introduced, usually in terms of jurisdiction size (e.g., Bucovetsky, 1991, and Wilson, 1991), larger jurisdictions are found to set higher taxes in equilibrium (e.g., Kanbur and Keen, 1993), but smaller jurisdictions are found to enjoy higher welfare. This effect is interpreted as reflecting the market power of large jurisdictions because they control a larger share of the aggregate tax base (e.g., Hoyt, 1992).

In this paper we test for strategic competition among state governments. That is, we examine interaction among a relatively small number of

Rubén Hernández-Murillo is an economist at the Federal Reserve Bank of St. Louis. Deborah Roisman provided research assistance.

[1] General conditions guaranteeing uniqueness and existence of the equilibrium in these models were established only recently, for example, by Laussel and LeBreton (1998) in the symmetric two-region game and by Peralta and van Ypersele (2002) in the asymmetric case with an arbitrary number of locations. The equilibrium of the model with specific functional forms is, however, explicitly derived in Bucovetsky (1991).

competing states, which are presumed to be able to affect the determination of the net rate of return on capital income.

The empirical literature, in its approach to the models of strategic interaction, estimates the reaction function of tax policies—the rule that individual states follow in choosing their tax rate, taking as given the configuration of tax rates chosen by its relevant competitors. The goal of such an approach is, as in this paper, to account for the effects from the interaction of *neighboring* governments, or relevant rivals, in the determination of policies. We identify neighboring states in terms of geographic and economic distance.

Identifying strategic interaction, or the ability of individual governments to affect the net rate of return on capital income, has important implications for the properties of the equilibrium configuration of tax rates among rival states: In particular, changes in the tax rate of one of the states, triggered by exogenous factors in that state, will imply cascading ramifications into other states' tax rates if there is strategic interaction.

This literature borrows from the spatial econometrics literature (see Anselin, 1988, for an introduction). One of the first studies is that of Case, Rosen, and Hines (1993), in which the authors test for spatial interaction in the determination of state expenditures. Besley and Case (1995) study strategic interaction in tax policies among states in a model where voters evaluate the performance of incumbent politicians using the taxing behavior of other jurisdictions as a benchmark, or *yardstick*. Brueckner (1998) analyzes the determination of growth control policies for cities in California. More recently, Saavedra (2000) applies the standard model of tax competition to the case of welfare benefits, Brueckner and Saavedra (2001) investigate property-tax competition among local governments in the Boston metropolitan area, and Fredriksson and Millimet (2002) examine the determination of environmental policies in the United States (see also Levinson, 1999a,b).[2]

A study closely related to ours is Buettner (2001), which analyzes the strategic determination of tax rates on capital income using a panel of municipalities in Germany. Buettner finds strong evidence of strategic interaction. Other studies that investigate strategic interaction among the continental United States include Case, Rosen, and Hines (1993), Besley and Case (1995), and Fredriksson, List, and Millimet (2002).

We estimate the reaction function for the determination of tax rates on capital income with yearly panel data using the methodology of instrumental variables, which is computationally simpler than alternative methods.[3]

In the following section we present the econometric model and the reaction function that is estimated. We then examine the data and comment on our estimation results. Finally, we provide some concluding remarks and discuss avenues for future research.

ECONOMETRIC MODEL

Tax-competition models predict the following: When tax rates are chosen strategically, the reaction function reflects the response of the representative jurisdiction to the policies set by its neighbors, because the mobility of the tax base—the capital stock, for example—depends on the policies of all the rival jurisdictions.

The model we estimate posits that a state's tax rate depends on the neighboring states' tax rates and on a set of local socioeconomic variables:

$$(1) \qquad t_{it} = \rho \Sigma_{j \neq i} \omega_{ij} t_{jt} + x'_{it} \beta + u_{it}.$$

Since we use a cross-section of states over time, the subscripts i and t represent states and time periods, respectively; ρ is a scalar parameter measuring the slope of the reaction function; ω_{ij} are spatial weights used to compute the effect of the tax rates of the relevant competitors of state i; $\omega_{ij} \neq 0$ if states i and j are rivals, and, by convention, $\omega_{ii} = 0$; finally, x_{it} is a vector of state i's socioeconomic conditions, and β is the corresponding vector of coefficients. We assume that the parameters ρ and β are constant across time periods and cross-section units. In vector form, the model is given by

$$(2) \qquad t_t = \rho W t_t + X_t \beta + u_t,$$

where $t_t = (t_{1t}, \ldots, t_{Nt})'$ is the $(N \times 1)$ vector of state policies for the cross-section of N states at time t, W is an $(N \times N)$ matrix of spatial weights, and X_t is an $N \times K$ matrix with rows given by the set of vectors x'_{it}; u_t is the corresponding $(N \times 1)$ error term vector. This functional form can be derived, for example, from the theoretical model of Peralta and van Ypersele (2002) with asymmetric jurisdictions.

2 For a comprehensive review of recent empirical studies, see Brueckner (2003).

3 We make the computer code and data available to the reader at our web site, < http://research.stlouisfed.org/publications/review/ >.

The spatial weights matrix, W, is constructed from exogenous factors; it determines the set of relevant competitors and captures the notion of proximity among the states in the tax competition game. When the spatial weights ω_{ij} are appropriately normalized to add up to 1, the *spatial lag* term, Wt_t, represents the average tax rate of the rivals and the *spatial autoregressive parameter* (or the slope of the reaction function), ρ, measures the individual response of a state's own tax rate to changes in the configuration of the tax rates among its neighbors.

Many alternative weighting schemes have been used in the literature. Weights are usually assumed to be known and constant over time; but, more generally, they are chosen to reflect some notion of geographic or socioeconomic distance between jurisdictions. For example, the simplest measure of proximity gives a weight $\omega_{ij} = 1$ if states i and j are contiguous and 0 if they are not. In this case, the weights are equal for each of the state's competitors. Another measure is based on gravity models where the weights are inversely proportional to geographic or economic distance, for example, $\omega_{ij} = 1/d_{ij}^2$, where d_{ij} is the distance in miles between state i and state j. An example of time-variant weights assigns a weight of zero to noncontiguous states and weights each contiguous state by its population. In our analysis we examine a variety of alternative weighting schemes, both with constant and time-variant measures.

The test for strategic interaction consists of testing for the significance of the spatial autoregressive parameter, ρ. Alternative models of interaction among governments yield different predictions on the sign of ρ. A common feature of models of horizontal competition, such as tax competition for capital, is the prediction of a positive slope. (See Brueckner, 2003.) Analogous to oligopolistic models of competition among firms, the interpretation of a positive sign is that tax rates are *strategic complements*, as in the differentiated products case of price competition. Models of vertical interaction among different levels of government often yield a negatively sloped reaction function. This indicates that policies are *strategic substitutes*, as in the case of quantity competition among firms. In our case, we expect the estimated sign of ρ to be positive.

Tax rates are determined endogenously in equilibrium; thus, the spatial lag term, Wt_t, is correlated with the error term u_t, and ordinary least-squares (OLS) yields inconsistent estimates of the parameters. Removing the simultaneity in model (2), the estimation can be carried out with alternative methods. The reduced form equation

$$t_t = (I_N - \rho W)^{-1} X_t \beta + (I_N - \rho W)^{-1} u_t, \tag{3}$$

where I_N represents an identity matrix of size N, is essentially the solution of the Nash equilibrium of the game.

Equation (3) can be estimated by maximum likelihood under normality assumptions. The study of Case, Rosen, and Hines (1993) was among the first to use this method. Other studies that use maximum likelihood methods are Besley and Case (1995), Brueckner (1998, 2000, 2003), Saavedra (2000), Brueckner and Saavedra (2001), and Garrett and Marsh (2002). Another alternative is the use of instrumental variables, as suggested by Kelejian and Robinson (1993) and Kelejian and Prucha (1998, 1999). Empirical studies that use this technique are Buettner (2001) and Fredriksson and Millimet (2002). These two methodologies yield consistent estimates of the parameters.

Even in the absence of spatial autoregression ($\rho = 0$), the estimation of model (2) can lead us to conclude erroneously that there is strategic interaction if the error term itself is subject to spatial autocorrelation, for example, in the form of

$$u_t = \lambda W u_t + \varepsilon_t = (I_N - \lambda W)^{-1} \varepsilon_t, \tag{4}$$

where ε_t is distributed with mean zero and covariance matrix $\sigma_\varepsilon^2 I_N$. In this case, spatial dependence in the error—for example, resulting from similar geographical conditions—can induce correlation in tax rates even though states may have no strategic considerations. Uncorrected spatial correlation in the error term would not affect the unbiasedness of the estimated parameter β, but it would reduce its efficiency. If there is strategic interaction ($\rho \neq 0$), ignoring the spatial lag term, Wt_t, in the estimation is more serious, since it yields inconsistent estimates of β. (See Case, Rosen, and Hines, 1993.) It is important, therefore, to test for both kinds of spatial dependence (in the dependent variable and in the error term).

Maximum likelihood estimation is complicated when we account for spatial correlation in the error term by possible identification problems. (See Anselin, 1988.) In our application we follow the instrumental variables approach because it avoids this issue, it is computationally easier to implement, and it does not require distributional assumptions on the error term ε.

Since we assumed that the parameters $\gamma = (\beta', \rho)'$ and $(\sigma_\varepsilon^2, \lambda)$ are time-invariant, we estimate the model pooling the panel observations, stacking observations

Table 1

Summary Statistics: 1977-99

Variable[a]	Mean	Mininum	Maximum	Standard deviation
Average capital tax rate (%)	25.8	16.4	42.2	4.4
Population	5,008,300.2	411,530.0	33,499,204.0	5,342,718.2
Population density	366.0	4.2	10,294.9	1,411.7
Percent of urban population	67.7	32.2	100.0	15.1
Percent unemployment	6.2	2.2	18.0	1.6
Per capita real personal income[b]	20,669.7	12,309.0	36,870.3	3,103.5
Percent of working-age population	52.2	36.5	72.6	7.6
Average monthly temperature	60.5	50.7	68.6	2.1

NOTE: [a]Cells are averages over time of statistics computed over cross-sectional units.
[b]In 1996 dollars.

so that the cross-section index runs faster than the time index. Stacking observations, we estimate the model:

(5) $$\boldsymbol{t} = \rho \boldsymbol{Wt} + \boldsymbol{X}\beta + \boldsymbol{u}$$

with

(6) $$\boldsymbol{u} = \lambda \boldsymbol{Wu} + \varepsilon,$$

where $\boldsymbol{t} = (t_1',\ldots,t_T')'$ and $\varepsilon = (\varepsilon_1',\ldots,\varepsilon_T')'$ are $(NT \times 1)$ vectors, with T equal to the total number of periods. $\boldsymbol{X}$ is the $(NT \times K)$ matrix of stacked exogenous variables. $\boldsymbol{W}$ is an $(NT \times NT)$ block-diagonal matrix of spatial weights, with T copies of W along the diagonal, in the case of time-invariant weights, and with matrices $(W_1,\ldots,W_T)$ in the case of time-variant weights. Finally, we assume that the covariance matrix of ε is given by $\sigma_\varepsilon^2 I_{NT}$.

When the spatial weights matrix is the same for both the spatial lag and the error autocorrelation models, as we assume, an application of Kelejian and Robinson's (1993) method of moments (GMM) estimator is outlined as follows. The covariance matrix of the error term is given by

(7) $$E(\boldsymbol{uu}') = \Sigma_{\boldsymbol{u}} = \sigma_\varepsilon^2 (I_{NT} - \lambda \boldsymbol{W})^{-1} (I_{NT} - \lambda \boldsymbol{W}')^{-1}.$$

The moments condition is the orthogonality between the set of instruments $\boldsymbol{H}$ and the error term $\boldsymbol{u}$:

(8) $$E(\boldsymbol{H}'\boldsymbol{u}) = \boldsymbol{0}.$$

Following Kelejian and Prucha's (1999) suggestion, we use as instruments the linearly independent columns of $(\boldsymbol{X}, \boldsymbol{WX})$. The GMM estimator is given by

(9) $$\hat{\gamma}_{GMM} = (\boldsymbol{Z}'\hat{\boldsymbol{D}}\boldsymbol{Z})^{-1} \boldsymbol{Z}'\hat{\boldsymbol{D}}\boldsymbol{t},$$

where $\boldsymbol{Z} = (\boldsymbol{Wt}, \boldsymbol{X})$, $\hat{\boldsymbol{D}} = \boldsymbol{H}(\boldsymbol{H}'\hat{\Sigma}_{\boldsymbol{u}}\boldsymbol{H})^{-1}\boldsymbol{H}'$, and $\hat{\Sigma}_{\boldsymbol{u}}$ is a consistent estimate of $\Sigma_{\boldsymbol{u}}$.[4] Kelejian and Robinson (1993) show that this estimator has an asymptotic normal distribution. The asymptotic variance (used to test for the significance of the coefficients) is given by

(10) $$\text{AsyVar}(\hat{\gamma}_{GMM}) = (\boldsymbol{Z}'\hat{\boldsymbol{D}}\boldsymbol{Z})^{-1}.$$

ANALYSIS

Data

We use yearly panel data on a set of variables reflecting local geographic, demographic, and economic conditions for the contiguous 48 United States and the District of Columbia over the period 1977-99. We construct an average tax rate on capital income using the methodology of Jones (2002) and Mendoza, Razin, and Tesar (1994).[5]

In the usual tax competition model, the tax base in each jurisdiction is determined endogenously by the tax rates set by the competing jurisdictions. In addition, the tax base is determined by local economic and demographic conditions, which reflect the size of the tax base and the nature of the

[4] To obtain a consistent estimate of $\Sigma_{\boldsymbol{u}}$, we use Kelejian and Prucha's (1999) nonlinear least-squares estimators of the error variance and autocorrelation parameters $(\sigma_\varepsilon^2, \lambda)$.

[5] This approach has been used for other purposes by Burnside, Eichenbaum, and Fisher (2000) and Cavallo (2002), among others.

public expenses the local jurisdiction has decided to finance with tax revenues.

Table 1 presents the summary statistics of the variables used in the estimation. Among the explanatory variables, we include real income per capita and the unemployment rate to account for local economic conditions. We include population, the fraction of urban population, and the fraction of working-age population (18-64 years) to account for the characteristics of the local tax base. In addition, we also include states' average monthly temperature to account for differences in the natural attractiveness of locations.

It is important to account for the size of states, as measured by population or personal income, to identify the effect of differentials in tax base size found in the theoretical literature; as we discussed in the introduction, the size of the tax base is commonly seen as indicative of the relative taxation power of local jurisdictions.

For periods between national censuses, we used population projections (including those by age group) provided by the Bureau of the Census. Population density was computed using these estimates and the square mileage of states computed by the Bureau of the Census in 1970, 1980, and 1990, assuming the same values throughout the decade following the census year.[6] We followed the same strategy with the measures of urban population. The data on real per capita personal income are from the Bureau of Economic Analysis (BEA), and the unemployment rate is from the Bureau of Labor Statistics. The data on average monthly temperature is from the National Climate Data Center.

We tried several alternatives for spatial weights. First we identified the set of rival states with a binary scheme for contiguous geographic boundaries (including states that share a common vertex). Second, as in Fredriksson and Millimet (2002), we considered neighboring states as defined by Crone (1998/1999), who updated the BEA's classification of economic regions according to common movements in indices of economic activity. The definition of relevant competitors in this case captures similarities in the composition of the states' industrial base and not merely their physical proximity. In addition to these two binary measures, we weight neighbors by three measures of economic distance: population, distance decay between the states' population centroids, and the distance decay between a set of local economic indicators. In the first case, weights are given by $\omega_{ijt} = P_{jt}$, where P_{jt} is neighbor j's population in year t and the set of neighbors of state i is defined by one of the binary schemes described above. In the second case, the weights are given by $\omega_{ijt} = 1/d_{ijt}^2$, where d_{ijt} is the distance in miles between the population centroids of states i and j.[7] Finally, in the third case, we also compute weights with the Mahalanobis (1930) distance between states, using population density, average temperature, and personal income as indicators.[8] In all cases, the weights matrix was row-normalized so that $\Sigma_{j \neq i} \omega_{ijt} = 1$. This normalization facilitates the interpretation and makes the parameter estimates of alternative models comparable (see Anselin, 2002); it has been used, even when the weights do not follow a binary scheme, by Case, Rosen, and Hines (1993) and Brueckner and Saavedra (2001).

Results

We present the results for the average capital tax rate in Tables 2 and 3. In Table 2 we use border contiguity to indicate neighboring states. In Table 3 we use Crone's (1998/1999) definition of economic regions to define neighboring states.

All variables are measured in logs, so the coefficients are interpreted as elasticities. The first column presents the OLS estimation of the restricted model, where we set $\rho = 0$ and $\lambda = 0$. This model is used to compute the robust LM statistics of Anselin et al. (1996) to evaluate the presence and source of spatial autocorrelation, that is, in the dependent variable or in the error term. $LM_{\rho\lambda}$ tests the null hypothesis of no spatial dependence ($\rho = \lambda = 0$). When this statistic is significant, we also use the tests for spatial lag in the dependent variable ($\rho = 0$), LM_ρ^*, and the test for spatial autocorrelation in the error term ($\lambda = 0$), LM_λ^*. Both of theses tests are robust to misspecification of the spatial dependence

6 The square mileage of states differs only slightly among these census years. The main source of variability in density is, of course, the variability in the population.

7 The distance was computed using the Haversine formula (see Sinnott, 1984) from the geographic coordinates for the population centroids computed by the Bureau of the Census for the years 1960, 1970, 1980, and 1990. We use the same coordinates throughout the decade following the census year.

8 The square *distance* between indicators $z_{it} = [PopDen_{it}, Temp_{it}, Inc_{it}]$ and $z_{jt} = [PopDen_{jt}, Temp_{jt}, Inc_{jt}]$ is computed as $d_{ijt}^2 = (z_{it} - z_{jt})\Sigma_t^{-1}(z_{it} - z_{jt})'$, where Σ_t is the covariance matrix computed using all the cross-section units in period t.

Table 2

Regression Results: Average Capital Tax Rate, Border Contiguity

Dependent variable: average capital tax rate

Variables[a]	OLS	W_{10}	W_{11}	W_{12}	W_{13}
Constant	1.36642	0.60228	–0.04111	1.31492	1.01958
	***(2.88)	(1.11)	(0.08)	**(2.33)	*(1.95)
Population	0.06453	0.06824	0.06343	0.07440	0.06655
	***(13.40)	***(14.51)	***(14.18)	***(15.02)	***(13.95)
Percent of urban population	0.02217	0.04490	0.02131	–0.01191	0.00238
	(1.00)	*(1.91)	(0.94)	(0.49)	(0.10)
Unemployment rate	–0.03044	–0.04167	–0.03163	–0.05814	–0.02532
	**(2.05)	***(2.64)	**(2.03)	***(3.42)	(1.63)
Per capita real personal income	0.12815	0.05523	0.08720	0.08263	0.04523
	***(3.79)	(1.31)	**(2.15)	*(1.87)	(1.16)
Average monthly temperature	–0.68053	–0.37025	–0.40014	–0.41866	–0.32413
	***(21.37)	***(8.52)	***(10.01)	***(8.72)	***(7.10)
Percent of working-age population	0.55510	0.10629	0.31355	0.08769	0.00012
	***(4.15)	(0.77)	**(2.30)	(0.59)	(0.00)
ρ		0.61029	0.52769	0.44188	0.64329
		***(9.12)	***(9.69)	***(5.87)	***(10.51)
λ		0.40673	0.32533	0.35571	0.23710
σ_ε^2	0.01809	0.01386	0.01419	0.01532	0.01554
$LM_{\rho\lambda}$		***386.11	***287.54	***280.68	***257.84
$LM_\rho{}^*$		***106.33	***97.29	***62.65	***120.98
$LM_\lambda{}^*$		*3.13	2.41	1.86	***34.17
$LM_\rho{}^{GMM}$		***83.26	***93.84	***34.47	***110.38
R^2	0.40224	0.55288	0.54206	0.47451	0.59520
No. of observations	1127	1127	1127	1127	1127

NOTE: [a]Dependent variable and regressors in logs. Weights matrices: W_{10}=border contiguity, W_{11}=border contiguity and population weights, W_{12}=border contiguity and geographic distance weights, and W_{13}=border contiguity and Mahalanobis distance weights. Absolute value of z statistics are in parentheses. */**/*** indicate significance at 10/5/1 percent levels, respectively, using the standard normal distribution for the coefficients and the χ^2 distribution for the LM statistics.

in the error term.[9] These tests have several advantages over tests derived from maximum likelihood estimation. The obvious one is that they require estimation of only the restricted model.[10]

In the second and following columns we present the estimation results for the full model using the instrumental variables method of Kelejian and Robinson (1993) and Kelejian and Prucha (1999). With this model we compute Saavedra's (2003) test for the presence of a spatial lag in the dependent variable, $LM_\rho{}^{GMM}$. This test also has good properties and does not require distributional assumptions on the error term.

The results for both definitions of proximity indicate that there is strategic interaction among neighboring states in the determination of capital tax rates. Furthermore, this interaction suggests a positively sloped reaction function in tax policies, as expected. In terms of elasticities, a unitless measure,

[9] In our case, we have assumed an autoregressive (AR) model for the error term; an alternative is the moving average (MA) specification, $u_t = \lambda W\varepsilon_t + \varepsilon_t$.

[10] In addition, these tests have been shown to behave well even for small samples using Monte Carlo simulations.

Table 3

Regression Results: Average Capital Tax Rate, Crone's Regions
Dependent variable: average capital tax rate

Variables[a]	OLS	W_{20}	W_{21}	W_{22}	W_{23}
Constant	1.36642	–1.85057	–2.55462	–0.73050	–1.50925
	***(2.88)	***(3.75)	***(5.12)	(1.35)	***(2.74)
Population	0.06453	0.04480	0.04981	0.04014	0.04529
	***(13.40)	***(9.79)	***(10.52)	***(8.55)	***(9.54)
Percent of urban population	0.02217	–0.00154	–0.03127	0.00227	–0.01588
	(1.00)	(0.07)	(1.37)	(0.10)	(0.68)
Unemployment rate	–0.03044	0.01587	0.01896	0.02212	–0.00061
	**(2.05)	(0.99)	(1.18)	(1.32)	(0.04)
Per capita real personal income	0.12815	0.33132	0.32649	0.27656	0.26924
	***(3.79)	***(7.78)	***(7.81)	***(6.49)	***(6.39)
Average monthly temperature	–0.68053	–0.54038	–0.48311	–0.52082	–0.57077
	***(21.37)	***(16.62)	***(14.80)	***(14.59)	***(15.10)
Percent of working-age population	0.55510	0.74568	0.88353	0.59960	0.86781
	***(4.15)	***(5.85)	***(6.70)	***(4.39)	***(6.46)
ρ		0.05686	0.05727	0.05272	0.04919
		***(5.84)	***(5.97)	***(5.42)	***(4.94)
λ		0.51804	0.43683	0.43159	0.41357
σ_ε^2	0.01809	0.01369	0.01388	0.01410	0.01462
$LM_{\rho\lambda}$		***400.87	***256.67	***269.27	***236.12
$LM_\rho{}^*$		**4.69	**6.59	***9.52	**4.47
$LM_\lambda{}^*$		***366.64	***220.77	***218.77	***200.41
$LM_\rho{}^{GMM}$		***34.1	***35.59	***29.35	***24.44
R^2	0.40224	0.46687	0.44396	0.38293	0.44700
No. of observations	1127	1127	1127	1127	1127

NOTE: [a]Dependent variable and regressors in logs. Weights matrices: W_{20}=Crone's regions, W_{21}=Crone's regions and population weights, W_{22}=Crone's regions and geographic distance weights, and W_{23}=Crone's regions and Mahalanobis distance weights. Absolute value of z statistics are in parentheses. */**/*** indicate significance at 10/5/1 percent levels, respectively, using the standard normal distribution for the coefficients and the χ^2 distribution for the LM statistics.

the response to an increase of 10 percent in the average of its rivals' tax rates (in this case, the geometric average, since we are using logs) results in an increase of about 4.4 to 6.4 percent in a state's initial tax rate when we use border contiguity.[11] The increase in a state's own tax rate in response to a 10 percent increase in the average tax rate of its neighbors is smaller, about 0.5 to 0.6 percent, if we consider economic regions. The spatial lag coefficients are significant according to standard z-statistics, but also according to the LM statistics.[12] The robust LM statistics computed with the OLS residuals suggest that the spatial dependent variable lag is the relevant source of spatial dependence when we use border contiguity in the spatial weights; but the tests identify spatial effects also in the error term when we use economic regions to assess proximity.

Turning to the explanatory variables, we find a positive effect of local population, indicating that

[11] That is, if the state's initial tax rate was 25 percent, according to its reaction function this state will increase its tax rate to 26.1 percent, a percentage change of 4.4 percentage points.

[12] Since the estimators are asymptotically normally distributed, we use the standard normal probabilities to evaluate significance of the coefficients.

more-populated states set higher tax rates. The elasticity is significant and is about 0.06 in the case of border contiguity and about 0.04 in the case of economic regions. Real per capita income also has a positive effect, but it is significant only when economic regions are considered; the elasticity is about seven times as large as the elasticity of population in this case. The signs of these effects are consistent with the view that larger (or richer) jurisdictions have higher market power at taxing local businesses, as is found in the theoretical models of Bucovetsky (1991), Kanbur and Keen (1993), and more recently Peralta and van Ypersele (2002).

The rate of unemployment has a statistically significant effect only in the case of border contiguity, and the effect is negative. The fraction of working-age population has a positive effect on local tax rates. These results are consistent with the positive effect of real per capita income and the size-effect interpretation.[13]

The degree of urbanization was not found to be significant in any of the specifications. We also included monthly average temperature as an explanatory variable to account for differences in the natural attractiveness of locations. The coefficient was found to be negatively significant; this may indicate differential effects of natural local advantages (as in north versus south) in the determination of tax policies.

In assessing the economic significance of these parameter estimates, it is important to notice that, because we have identified that there is strategic interaction, the effects of the exogenous variables will have an impact on the entire configuration of equilibrium tax rates; this is true even if they take place in only one of the competitor states, through their repercussions on the local tax rates and, subsequently, through the responses in the tax-setting behavior of the other states.

CONCLUSION

The identification of strategic interaction in models of tax competition has important implications for the analysis of the comparative statics of the equilibrium configuration of tax rates. In this paper we found evidence of strategic interaction among states in the determination of tax policies on capital income. When we examined competition among states that share common movements in economic activity, the elasticity of the average tax rate of neighboring states was found to be as large as the elasticity of local economic variables such as real per capita income. The elasticity of local population was much larger.

In our investigation we did not model explicitly the source of strategic interaction between state governments. This is a limitation of the analysis in the sense that the source of interaction is not identified. We used as motivation for the analysis the traditional model of capital tax competition, but the empirical results may be consistent with other types of competition, such as yardstick competition among politicians in models of elections.

Our results are encouraging, however, because once the presence of strategic interaction is identified, the natural extension is to try to account for specific models of tax competition to analyze the normative implications of government behavior we discussed in the introduction. One possibility is to test for specific forms of the objective function of competing governments. For example, given existing arguments in the popular press and legislative efforts to curtail horizontal competition among governments in the United States, one could assess the welfare costs of competition by testing the implications of assuming that states behave as welfare maximizers or as revenue maximizers.

Finally, by making the computer code available to the reader, we expect this paper to serve also as an illustration of the method of instrumental variables. This technique is computationally simpler than maximum likelihood and allows for faster estimation of the full model even when spatial effects in the error term have the same spatial weights matrix as the spatial lag in the dependent variable.

[13] Jones (2002) observes a positive contemporaneous correlation between output and average capital rates in the data at the national level.

REFERENCES

Anselin, Luc. *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer Academic Publishers, 1988.

___________. "Under the Hood: Issues in the Specification and Interpretation of Spatial Regression Models." *Agricultural Economics*, November 2002, *27*(3), pp. 247-67.

___________; Bera, A.K.; Florax, R. and Yoon, M.J. "Simple Diagnostic Tests for Spatial Dependence." *Regional Science and Urban Economics*, February 1996, *26*(1), pp. 77-104.

Besley, T. and Case, A. "Incumbent Behavior: Vote-Seeking, Tax Setting, and Yardstick Competition." *American Economic Review*, March 1995, *85*(1), pp. 25-45.

Brueckner, Jan K. "Testing for Strategic Interactions Among Local Governments: The Case of Growth Controls." *Journal of Urban Economics*, November 1998, *44*(3), pp. 438-67.

___________. "A Tiebout/Tax-Competition Model." *Journal of Public Economics*, August 2000, *77*(2), pp. 285-306.

___________. "Strategic Interaction Among Governments: An Overview of Empirical Studies." *International Regional Science Review*, 2003 (forthcoming).

___________ and Saavedra, Luz A. "Do Local Governments Engage in Strategic Property-Tax Competition?" *National Tax Journal*, June 2001, *54*(2), pp. 203-29.

Bucovetsky, S. "Asymmetric Tax Competition." *Journal of Urban Economics*, September 1991, *30*(2), pp. 167-81.

Buettner, Thiess. "Local Business Taxation and Competition for Capital: The Choice of the Tax Rate." *Regional Science and Urban Economics*, April 2001, *31*(2-3), pp. 215-45.

Burnside, Craig; Eichenbaum, Martin and Fisher, Jonas D.M. "Assessing the Effects of Fiscal Shocks." NBER Working Paper No. 7459, National Bureau of Economic Research, January 2000.

Case, Anne C; Rosen, Harvey S. and Hines, James R. Jr. "Budget Spillovers and Fiscal Policy Interdependence: Evidence from the States." *Journal of Public Economics*, October 1993, *52*(3), pp. 285-307.

Cavallo, Michele. "Government Employment and the Dynamic Effects of Fiscal Policy Shocks." Unpublished manuscript, New York University, November 2002.

Crone, Theodore M. "Using State Indexes to Define Economic Regions in the U.S." *Journal of Economic and Social Measurement*, 1998/1999, *25*(3-4), pp. 259-75.

Fredriksson, Per G.; List, John A. and Millimet, Daniel L. "Chasing the Smokestack: Strategic Policy-Making with Multiple Instruments." Unpublished manuscript, Southern Methodist University, 2002.

___________ and Millimet, Daniel L. "Strategic Interaction and the Determination of Environmental Policy Across U.S. States." *Journal of Urban Economics*, January 2002, *51*(1), pp. 101-22.

Garrett, Thomas A. and Marsh, Thomas L. "The Revenue Impacts of Cross-Border Lottery Shopping in the Presence of Spatial Autocorrelation." *Regional Science and Urban Economics*, July 2002, *32*(4), pp. 501-19.

Hoyt, William H. "Market Power of Large Cities and Policy Differences in Metropolitan Areas." *Regional Science and Urban Economics*, November 1992, *22*(4), pp. 539-58.

Jones, John B. "Has Fiscal Policy Helped Stabilize the Postwar U.S. Economy?" *Journal of Monetary Economics*, May 2002, *49*(4), pp. 709-46.

Kanbur, Ravi and Keen, Michael. "Jeux Sans Frontieres: Tax Competition and Tax Coordination When Countries Differ in Size." *American Economic Review*, September 1993, *83*(4), pp. 877-92.

Kelejian, Harry H. and Prucha, Ingmar R. "A Generalized Spatial Two-Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances." *Journal of Real Estate Finance and Economics*, July 1998, *17*(1), pp. 99-121.

___________ and ___________. "A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model." *International Economic Review*, May 1999, *40*(2), pp. 509-33.

___________ and Robinson, D.P. "A Suggested Method of Estimation for Spatial Interdependent Models with Autocorrelated Errors, and an Application to a County Expenditure Model." *Papers in Regional Science*, July 1993, *72*(3), pp. 297-312.

Laussel, Didier and LeBreton, Michel. "Existence of Nash Equilibria in Fiscal Competition Models." *Regional Science and Urban Economics*, May 1998, *28*(3), pp. 283-96.

Levinson, Arik. "NIMBY Taxes Matter: The Case of State Hazardous Waste Disposal Taxes." *Journal of Public Economics*, October 1999a, *74*(1), pp. 31-51.

___________. "State Taxes and Interstate Hazardous Waste Shipments." *American Economic Review*, June 1999b, *89*(3), pp. 666-77.

Mahalanobis, P.C. "On Tests and Measures of Group Divergence." *Journal of the Asiatic Society of Bengal*, 1930, *26*, pp. 541-88.

Mendoza, Enrique G.; Razin, Assaf and Tesar, Linda L. "Effective Tax Rates in Macroeconomics: Cross-Country Estimates of Tax Rates on Factor Incomes and

Consumption." *Journal of Monetary Economics*, December 1994, *34*(3), pp. 297-323.

Mintz, Jack and Tulkens, Henry. "Commodity Tax Competition Between Member States of a Federation: Equilibrium and Efficiency." *Journal of Public Economics*, March 1986, *29*(2), pp. 133-72.

Peralta, Susana and van Ypersele, Tanguy. "Capital Tax Competition Among an Arbitrary Number of Asymmetric Countries." Discussion Paper 2002/31, CORE-UCL, 2002.

Saavedra, Luz A. "A Model of Welfare Competition with Evidence from AFDC." *Journal of Urban Economics*, March 2000, *47*(2), pp. 248-79.

___________. "Tests for Spatial Lag Dependence Based on Method of Moments Estimation." *Regional Science and Urban Economics,* January 2003, *33*(1), pp. 27-58.

Sinnott, R.W. "Virtues of the Haversine." *Sky and Telescope*, 1984, *68*(2), pp. 159.

Wildasin, David E. "Nash Equilibria in Models of Fiscal Competition." *Journal of Public Economics*, March 1988, *35*(2), pp. 229-40.

Wilson, John D. "Tax Competition with Interregional Differences in Factor Endowments." *Regional Science and Urban Economics*, November 1991, *21*(3), pp. 423-51.

___________. "Theories of Tax Competition." *National Tax Journal*, June 1999, *52*(2), pp. 269-304.

Appendix

CALCULATION OF AVERAGE CAPITAL TAX RATES

Appendix B in Jones (2002) describes his approach to calculating national average tax rates on capital income using the national income and product accounts (NIPA). We follow the same methodology and calculate average tax rates on capital income at the state level using the state analog of the variables he describes.

The average tax rate on capital income is defined, using Jones's (2002) notation, as

$$\tau_K = \frac{\tau_P CI + CT + PT}{CI + PT},$$

where *CI* is capital income, *PT* is (corporate) property state and local tax payments, and *CT* is state corporate (income) tax payments. Capital income is defined as

$$CI = PRI/2 + RI + CP + NI,$$

where *PRI* is proprietor's income, *RI* is rents income, *CP* is dividends, and *NI* is net interest income. Finally, the personal income tax rate, τ_P, is defined by

$$\tau_P = \frac{FIT + SIT}{W + PRI/2 + CI},$$

where *FTI* is (personal) federal income tax payments, *SIT* is (personal) state income tax payments, and *W* is wages and salary income.

In our calculations, the corporate income and property tax payments were obtained from state and local government finances data from the Bureau of the Census. The variables used to compute capital income and the average personal income tax rate were obtained from the state personal income calculations of the Bureau of Economic Analysis.