

A Perspective on U.S. International Capital Flows

William Poole

I am very pleased to be here today to visit with the Tucson Chapter of the Association for Investment Management Research. I say “visit with” because I do hope that when I finish speaking we can engage in some questions and answers and comments about my chosen topic. International economic issues—especially trade issues—are hot topics these days. Through my concentration on capital markets issues, my intention is to emphasize just how important international capital flows are to the United States. In the process, I hope to shed some light, and not just add to the heat, on trade issues by exploring the intimate connections between international trade and international capital flows.

Recent economic indicators have suggested that the long-awaited acceleration of the recovery from the 2001 recession is under way. According to the advance estimate from the Department of Commerce, real GDP growth—the broadest measure of the strength of the economy—increased at a 7.2 percent annual rate in the third quarter, and the latest employment data show that the accelerated growth is fueling job creation after many months of stagnation.

Through all the ups and downs of the U.S. economy over the past two decades, a staple of the situation has been a deficit in the U.S. international trade accounts and a corresponding surplus in the international capital accounts. Many observers are troubled by this persistent state of affairs and are concerned that the trade deficit might derail the economic recovery. It is common to refer to the situation as an “imbalance,” which naturally implies that something is wrong. The word “deficit” in “trade deficit” has the same connotation. I intend to use the words “surplus” and “deficit” as simple descriptive words and hope that in listening to me you can consciously ignore the baggage that the words com-

monly carry. My purpose is to analyze the external imbalance to see why we might, or might not, be concerned about it.

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 appreciate comments provided by my colleagues at the Federal Reserve Bank of St. Louis. Michael R. Pakko, senior economist in the Research Division, provided special assistance. I take full responsibility for errors.

To emphasize the importance of thinking through the analysis and not letting the word “deficit” decide the issue, consider the situation faced by many healthy corporations. It is common for a thriving company to spend more than its revenues, making up the difference by borrowing. When a company borrows to finance spending on capital, the company may be said to have a deficit on current account—its total spending on goods and services, including new capital, exceeds its revenues. The company simultaneously has a surplus on capital account—more funds are flowing into the company to buy the company’s shares and bonds than the company is investing in similar securities issued by others. Arithmetically, the company has a current account deficit and a capital account surplus, and thus has an “imbalance.” Whether the company is suffering from an economic imbalance depends on the productivity of its capital investments. Sometimes companies do invest in capital and businesses that turn out not to yield returns sufficient to service the debt financing the investments. Such a situation, when repeated over the years, is not sustainable. For a company, and as I will argue for a country, whether continuing infusions of financial capital are sustainable depends on how the financial capital is employed.

William Poole is the president of the Federal Reserve Bank of St. Louis. The author thanks colleagues at the Federal Reserve Bank of St. Louis for comments: Michael R. Pakko, senior economist in the Research Division, provided special assistance. The views expressed are the author’s and do not necessarily reflect official positions of the Federal Reserve System.

Federal Reserve Bank of St. Louis *Review*, January/February 2004, 86(1), pp. 1-7.
© 2004, The Federal Reserve Bank of St. Louis.

CURRENT AND CAPITAL ACCOUNTS IN THE BALANCE OF PAYMENTS

The most widely cited measure of the U.S. external imbalance is the trade deficit—the difference between U.S. exports and imports. More generally, it is useful to concentrate on the broader concept of the current account, which includes current earnings on capital as well as trade in goods and services. Putting aside errors and omissions in the data, the capital account surplus is necessarily equal to the current account deficit. By the same token, a country with a current account surplus simultaneously has a capital account deficit—that is, it is importing more capital claims than it is exporting. In the official statistics reported by the Bureau of Economic Analysis, this side of the ledger is called the “Capital and Financial Account.”

A country's trade balance and its capital account are clearly very closely related. From an economist's perspective, the flows of goods and services that comprise the trade balance tell only part of the story of a country's international economic relations. I'm going to concentrate on the capital account because that part of the international economic story is commonly neglected.

A common mistake is to treat international capital flows as though they are passively responding to what is happening in the current account. The trade deficit, it is said, is financed by U.S. borrowing abroad. In fact, investors abroad buy U.S. assets not for the purpose of financing the U.S. trade deficit but because they believe these are sound investments promising a good combination of safety and return. Many of these investments have nothing whatsoever to do with borrowing in the conventional meaning of the word, but instead involve purchases of land, businesses, and common stock in the United States. Foreign auto companies, for example, have purchased land and built manufacturing plants in the United States. These simple examples should make clear that a careful analysis of the nature of international capital flows is necessary before offering judgments about the U.S. external imbalance.

RECENT TRENDS IN THE U.S. INTERNATIONAL FINANCIAL POSITION

Examining recent trends in the U.S. international financial position will help to uncover key facts and issues. There is a huge amount of detailed data in official U.S. statistics. I'll draw on some of that information.

The capital account measures the change in the net foreign asset position of a country for a given period, such as a year. For the United States, the capital account includes the accumulation of foreign assets by U.S. residents as well as the accumulation of U.S. assets by foreigners. In the U.S. balance of payments accounts, each of these gross asset flows is broken down into “official” flows—representing asset purchases by governments and central banks—and “private” flows—representing the purchases of individuals and corporate entities. These totals are further broken down by type of asset—government securities, corporate bonds, private equity—in tables reporting the international investment position of the United States.

The sheer volume of international financial flows is truly phenomenal. According to the Bank for International Settlements, in 2001 trade in foreign currencies averaged \$1.2 trillion per *day*, and trading in derivatives averaged \$1.4 trillion per day. Much of this daily activity nets out when measuring quarterly and annual flows, but even the quarterly and annual magnitudes have been quite large. Moreover, they have been rising significantly over the past few years. For example, foreign-owned U.S. assets increased by an average of \$155 billion per year during the 1980s. Since 2000, foreign ownership of U.S. assets increased at an average rate of \$833 billion per year—more than a fivefold increase. In 2000, over \$1 trillion of assets were sold to foreign entities.

Growth of U.S. ownership of foreign assets has shown similar, if not quite so remarkable, growth. Averaging \$95 billion during the 1980s, the U.S. entities have accumulated foreign assets at a rate of \$366 billion per year over the past three years. Over the entire span of this comparison, the volume of U.S. assets owned abroad has outpaced our accumulation of foreign assets—a capital account surplus that has moved our country from a positive to a negative net asset position.

It is sometimes said that the United States has become a net debtor. The word “debtor” is extremely misleading in this context, for the U.S. assets owned by foreigners include equities and physical capital located in the United States, as well as bonds issued by U.S. entities. Moreover, the part of the U.S. international financial position that is debt, by which I mean bonds and other fixed claims such as bank loans, is predominantly denominated in dollars. A country with most of its debt denominated in its own currency is in a very different situation from one whose debt is denominated in other currencies.

The familiar crises experienced by several Asian countries in 1997-98, by Mexico on several occasions, by Argentina, and by numerous other countries have all involved situations in which the impacted countries have had large external debts denominated in foreign currencies.

The balance-of-payments accounts provide estimates of annual international investment flows. These accumulate over time to change the stocks of assets. Data on the stocks are available and are referred to as measures of the U.S. international investment position.

As recently as the early 1980s, the U.S. had a positive net investment position. As a consequence of large capital inflows in the 1980s and late 1990s, the United States today has the world's largest negative net international investment position. By the end of 2002, foreigners owned more than \$9 trillion of U.S. assets, based on market values, while U.S.-owned assets abroad reached a level of not quite \$6.5 trillion. Hence, at the end of last year, the U.S. net international investment position represented a negative net position of \$2.6 trillion, about 25 percent of U.S. GDP.

This new role for the United States, with its negative net international investment position, has been a source of consternation among those who see the globalization of financial markets as a worrisome phenomenon. I am much more sanguine about the U.S. international asset position. To explain why I view the rapid growth of cross-border financial market activity in a positive light, I'll discuss some basic economic principles that underlie changes in the U.S. net international position. It would be a mistake, though, to think that the United States is in uncharted waters; other prosperous countries have had large negative international investment positions without getting into trouble, and the United States itself was in this position for decades prior to World War I.

TRADE AND CAPITAL FLOWS

In today's world, with electronic funds transfers, financial derivatives, and largely unrestricted capital flows, investors have a global marketplace in which to seek profitable returns and diversify risk. In such an environment, we should consider the possibility that aggregate patterns of international trade flows may simply be the by-product of a process through which financial resources are seeking their most efficient allocations in a worldwide capital market. That is, instead of thinking that capital flows are

financing the current account deficit, it may well be that the trade deficit is, so to speak, financing capital flows driven by investors seeking the best combination of risk and return in the international capital market.

While such a conclusion is surely an overstatement, I believe that it does contain an important element of truth. Capital flows are a highly dynamic feature of the international economy; changes in investor attitudes and expectations can alter capital flows quickly and force changes in the trade account. To paint a more complete picture of the broad nexus of forces driving trade and investment patterns around the world, I will describe three complementary views of how cross-border goods and asset flows are jointly determined.¹

Perhaps the most basic model for explaining a country's international position could be called "the trade view," which focuses explicitly on the factors determining the import and export of goods and services. Under this perspective, the emphasis is on the economic conditions that determine whether a country runs a deficit in trade. The capital account simply measures the offsetting financial transactions that take place; investors are treated as passive players who finance what is happening in the dynamic trade sector. This view lends itself naturally to the application of basic principles of demand theory. The quantity of goods and services that a country imports depends on income and the relative price of imports, which is determined importantly by the exchange rate. Exports depend on the responses of a country's trading partners to changes in their income and exchange rate movements.

Economists who have taken an empirical approach to estimating these demand relationships have found that the trade view can explain much about the fluctuations in trade and capital flows that we observe across countries. But their estimates have also presented a puzzle: U.S. import demand responds more strongly to changes in income growth than corresponding income responses in other countries. This finding means that, in the long run, with exchange rates settling at their equilibrium values and U.S. and foreign growth rates equal, the U.S. is predicted to run a *persistently* widening current account deficit. Alternatively, a widening deficit could be halted by a persistent depreciation

¹ In describing these three views and highlighting the importance of international capital flows, I draw on the work of Catherine L. Mann, a former economist at the Fed who is now a Senior Fellow at the Institute for International Economics in Washington, D.C. (Mann, 2002).

of the dollar, or by suffering a persistently slower growth rate than the rest of the world.

The conclusion is that either the United States is destined to face some combination of these undesirable outcomes—a continuously depreciating currency and/or lower GDP growth than the rest of the world—or the demand equations of the trade view are missing something. What might be missing is some important factor outside the trade view that can explain the recent historical trend of a widening U.S. current account deficit in an environment in which U.S. GDP growth is on average higher than growth in much of the rest of the world and in which the dollar, despite short-run fluctuations, is on average relatively strong and not persistently depreciating.

A second perspective of current account/capital account determination is best explained through accounting identities of the National Income and Product Accounts. The National Accounts are structured such that the total output—the GDP—of the United States is divided into principal components of consumption, investment, spending by government on goods and services, and exports. Total income from production can be either consumed or saved. These relationships imply that a current account deficit must equal the difference between U.S. domestic investment, or capital formation, and total U.S. saving by both the private sector and government.

This view suggests several explanations for U.S. current account deficits. One explanation that gained popularity in the 1980s was that large, persistent government budget deficits reduced national saving and thereby induced an inflow of financing from abroad. This “twin-deficit” argument has some appeal, particularly in that it suggests a key role for capital account flows. The argument is that claims on U.S. assets are exported to help finance government budget deficits. Indeed, the growth of the U.S. capital account surplus has included a vast accumulation of U.S. Treasury debt by foreigners. It is estimated that over \$1.4 trillion of U.S. Treasury debt is currently held by foreigners, representing about 37 percent of the total outstanding. It is important to recognize, however, that foreign purchases of any U.S. assets, and not just Treasury bonds, serve to help finance the government budget deficit.

The twin-deficits explanation, however, is clearly inadequate. While this explanation appeared to fit the facts of U.S. experience in the 1980s, the relationship breaks down when examining other episodes.

One recent example is the United States during the late 1990s, when the current account deficit persistently widened while the government budget moved from deficit to surplus. In other countries that have experienced large swings in government deficits and current account deficits, the twin-deficits theory doesn’t seem to hold up in terms of timing or magnitude either.

Another explanation suggested by the savings/investment view is that periods of high investment demand—like the late 1990s in the United States—fully absorb domestic savings and require additional external financing. This explanation puts a completely different spin on current account deficits. If we are exporting claims on U.S. assets—financing abroad by selling bonds, equities, and claims on productive facilities—to fund productive investment opportunities, the payoff from those investments will finance the obligations incurred. This is a classic example of how financial markets can be used to exploit productive opportunities that might otherwise be unavailable.

From this perspective, the profitability of U.S. investment opportunities makes United States something of an “oasis of prosperity,” attracting savings from around the world from those who wish to share in the returns and safety of investing in U.S. markets. On this view, trade and current account deficits are induced by the dynamic role of the United States in world capital markets.

And yet this savings/investment view also appears incomplete and not in accord with recent facts. The U.S. external imbalance has continued to widen in recent years despite the fact that growth in the investment component of GDP dropped precipitously during the recent recession and has only recently shown signs of picking up. Moreover, returns in the U.S. equity market were pretty miserable from early 2000 until quite recently. Again, we seem to be left with only part of the story.

A third view of the U.S. external imbalance can, I believe, help complete the story. Just as the savings/investment view exploits the accounting identities of the National Accounts, an “international capital markets view” can be derived from the identity that one country’s deficit is balanced by another country’s surplus. From this perspective, capital account adjustment can play an important independent role that is determined by the motivations of both foreign and domestic investors. In particular, we can think of capital market flows as the equilibrium outcome of investors worldwide seeking to acquire

portfolios that balance risk and return through diversification.

THE U.S. ROLE IN INTERNATIONAL CAPITAL MARKETS

Current commentary on international economic issues pays far too little attention to the role of the United States in international capital markets. The globalization of financial markets—spurred by technological advances and liberalization of capital flow restrictions worldwide—has created entirely new investment opportunities for investors in both the United States and abroad. These new opportunities have undoubtedly given rise to a re-balancing of portfolios, and there are reasons to believe that this process might be associated with a net export of claims on U.S. assets—a capital account surplus.

U.S. financial markets are among the most highly developed in the world, offering efficiency, transparency, and liquidity. Moreover, the U.S. dollar serves as both a medium of exchange and a unit of account in many international transactions. These factors make dollar-denominated claims attractive assets in any international portfolio. No capital market in the world has a combination of strengths superior to that of the United States. Our advantages include the promise of a good return, safety, secure political institutions, liquidity, and an enormous depth of financial expertise. The United States has worked hard for generations to create outstanding capital markets; our latest efforts to improve corporate governance should be viewed as simply another chapter in an ongoing story.

For some purposes, it is useful to think of U.S. financial markets as serving as a world financial intermediary. Just as a bank, or a mutual fund, channels the savings of many individuals toward productive investments, the U.S. financial markets play a similar role for many investors from around the world. In the process, individuals, companies, and governments around the world accumulate dollar-denominated assets to serve as a vehicle for facilitating transactions and storing liquid wealth safely.

A bank earns its return on capital by paying a lower interest rate to depositors than it earns on its assets. Similarly, the United States earns a higher return on its investments abroad than foreigners do on their investments in the United States. Despite the fact that the U.S. international investment position at the end of 2002 was $-\$2.6$ trillion, U.S. net

income in 2002 on its investments abroad slightly exceeded income payments on foreign-owned assets in the United States.

How is the United States able to earn a significantly higher return on its assets abroad than foreigners earn on their assets in the United States? A very simple example is currency, which pays a zero return. At the end of 2002, U.S. currency held abroad was estimated to be about $\$300$ billion, whereas only a trivial amount of foreign currency is held in the United States.

More generally, many private and governmental investors abroad rely on the U.S. capital market as the best place to invest in extremely safe and highly liquid securities. Along a spectrum of safety and liquidity, these assets include currency, U.S. government obligations, agency debt, and corporate bonds. U.S. equity markets are also highly liquid. The United States as a whole earns a return from providing these safe and liquid investments to the world. Indeed, the desire of foreigners to hold U.S. Treasury securities is a testament to the confidence that the world has in the safety and soundness of our financial system.

Recent data show just how impressive is the world's appetite for safe U.S. assets. Over the six quarters ending with the second quarter of this year, total outstanding U.S. government debt rose by about $\$345$ billion, while foreign holdings of such debt have risen by about $\$304$ billion.

Another force at work may be a gradual breakdown in the home bias to investment. For some years, international economists have noted that investors tend to hold portfolios that are weighted more toward domestic assets than would appear optimal by the principles of diversification—there is home-bias to investor behavior. Business cycles and investment risks are not perfectly synchronized across countries; a balanced international portfolio can help to diversify risk. The opening of global capital markets has allowed investors to exploit these opportunities, particularly foreign investors who are able to participate in the relative openness of U.S. capital markets and the multinational diversification of U.S. corporations.

Another aspect of the situation may be a consequence of demographics. Europe and Japan, especially, have populations that are aging more rapidly than does the United States. Just as a retired household typically consumes more than its income, drawing down retirement savings in the process, so also may an entire country draw down international investments to finance the consumption of its retired

population. Japan especially has a high saving rate relative to its domestic investment rate; it is accumulating assets abroad that may be run down in future years to support its elderly population. This process is one that will work out over many years. What may appear to be an “imbalance” this year may make perfect sense over a long-term horizon.

While the international capital markets view provides a perspective on some unique influences on U.S. current account/capital account imbalances, it is entirely consistent with the alternative perspectives.

As foreigners decide to accumulate dollar-denominated assets, U.S. interest rates are kept lower than they otherwise would be, which tends to increase investment demand in the United States. This investment demand, incidentally, includes both corporate demand for capital formation and household demand for new housing. The total demand for funds also includes the U.S. government’s demand, which may be temporarily high as a consequence of the war on terrorism, Iraq, and the 2001-03 period of recession and slow recovery. These factors are consistent with the savings/investment perspective that helps to understand why the United States has a capital inflow and the associated current account deficit.

IS THE U.S. EXTERNAL IMBALANCE SUSTAINABLE?

When considering widening external imbalances like those that the United States has experienced in recent years, a natural question is whether or not current trends are sustainable. Indeed, with a current account deficit equal to 5 percent of GDP and a negative international investment position that amounts to 25 percent of GDP, some have drawn comparisons with countries such as Argentina, Brazil, and Mexico at times of severe balance of payments crises.

I consider it highly unlikely that such a crisis will befall the United States. As a stable, diversified industrial economy, the United States is not likely to suffer from a sudden lack of confidence by investors. In fact, other industrialized economies have incurred much larger external obligations without precipitating crises. For example, Australia’s negative net investment position reached 60 percent of GDP in the mid-1990s, Ireland’s exceeded 70 percent in the 1980s, and New Zealand accumulated a position amounting to nearly 90 percent of GDP in the late 1990s. Notably, these economies have recently

been among the most successful—in terms of economic growth—in the industrialized world. Indeed, the combination of rising external obligations and prospects for robust growth is entirely consistent with the view of the capital account I have discussed today.

Moreover, the international capital markets view suggests that the United States is not only more like those countries that have experienced high levels of debt without obvious ill effects—but that the U.S. case is, in some sense, unique. The central role of U.S. financial markets—and of the dollar—in the world economy suggests that capital account surpluses are being driven by foreign demand for U.S. assets, rather than by any structural imbalance in the U.S. economy itself.

To be sure, no country can permanently incur rising levels of net external obligations relative to GDP. If sustained indefinitely, service payments on ever-increasing obligations would ultimately exceed national income. Long before that situation of literal insolvency occurred, however, market forces would drive changes in exchange rates, interest rate differentials, and relative growth rates in such a way to move the economy toward a sustainable path. Nevertheless, such adjustments need not be sudden, large, or disruptive as they have sometimes been for countries with severe balance-of-payments crises.

Not only are there market forces that will restore equilibrium, should the current situation not correct itself, but more importantly the international capital markets may well be looking ahead to changing circumstances that will reduce the capital flows to the United States in coming years. I’ve already mentioned the demographic forces at work. Another possibility is that economic growth will rise elsewhere in the world, raising demands both for U.S. exports and for international capital to finance higher growth. Given the strength of U.S. multinational corporations, U.S. firms will share in the profits from higher growth abroad, and some of those earnings will be repatriated to the United States.

CONCLUDING COMMENTS

The international financial markets view of U.S. international capital account determination that I have described today highlights the dynamic role of international capital adjustments as investors exploit the opportunities of globalized financial markets. Because the technological progress and capital-market liberalizations that have driven this process have evolved over time, the process has

been protracted. Ultimately, however, when portfolio adjustments have optimally exploited new diversification opportunities, and as growth abroad rises, the net international investment position of the United States will stabilize.

If this view is correct, the forces driving the U.S. capital account represent a persistent, but ultimately temporary, process that might result in a higher level of net indebtedness without necessarily posing any threat to the sustainability of the U.S. international investment position. Nor will the transition to a sustainable long-run path necessarily require wrenching adjustments in domestic or international markets or in exchange rates.

In the meanwhile, we can all benefit from our good fortune to have access to increasingly safe, liquid, and transparent financial markets. The United States has created for itself a comparative advantage in capital markets, and we should not be surprised that investors all over the world come to buy the product.

Casino Gaming and Local Employment Trends

Thomas A. Garrett

THE U.S. CASINO INDUSTRY

Casino gaming has become a major industry in the United States over the past two decades. Prior to the late 1980s, casino gaming was legal only in Nevada and Atlantic City, New Jersey. Today, casino gaming is available in 29 states. As a consequence, annual gaming revenue has grown from \$9 billion in 1991 to over \$40 billion in 2001.¹ Americans spend more money in casinos than individually on golf, on-screen movies, CDs and sound equipment, and cable TV.²

The casino industry consists of two major parties—Indian tribes and publicly traded private corporations such as Harrah's Entertainment and Trump Hotels and Casino Resorts. The Indian Gaming Regulatory Act (Public Law 100-497, passed in 1988) allows Indian tribes to own and operate casinos on their reservations. Tribal gaming is available in 25 states and generates nearly \$13 billion in annual revenue. Corporate casino gaming is available in 10 states and generated over \$27 billion in revenue in 2001.³ Table 1 provides a listing of these states.⁴

¹ Casino revenues in this report represent revenue to the casinos after subtracting player winnings. In comparison, 38 state lotteries generated nearly \$38 billion in 2001; pari-mutuel horse and greyhound racing generated over \$3.25 billion (legal in 43 states) and \$550 million (legal in 15 states), respectively.

² From the American Gaming Association (www.americangaming.org).

³ West Virginia, Delaware, Rhode Island, Louisiana, Iowa, and New Mexico offer video lottery terminals (VLTs) and slot machines as part of their state lottery. These outlets (often called racinos) are usually located at pari-mutuel racetracks and are not considered corporate casinos because they are run by the state. The revenue from these outlets (roughly \$500 million annually) is considered a portion of total lottery revenue.

⁴ Most data presented in this report are from the following sources: *Bear Stearns 2002-2003 North American Gaming Almanac* (Ader, 2003); The National Gambling Impact Study Commission, *Final Report* (1999), www.casino-gambling-reports.com/GamblingStudy/; the American Gaming Association, www.americangaming.org; and each state's gaming commission.

While tribal gaming is available in more states, corporate casino gaming has traditionally been perceived as a more appropriate tool for fostering general economic development through increased employment and tax revenue.⁵ The primary reason for this perception is that states have no power to tax Indian casino revenue because Indian reservations are sovereign entities, distinct from the state.⁶ While states and Indian tribes do cooperate on regulation and security issues (dictated by state-tribal gaming compacts), the relationship between a tribe and a state is very similar to the relationship between two states—one state generally cannot legally dictate what another state can do. Corporate casinos, however, are private industries that are taxed and regulated by the state. These casinos generate much more revenue and hire more labor from the general labor market than Indian casinos.

The impact of corporate casino gaming on local employment is a major issue in the debate over legalized casino gaming. As a result, the issue has received careful study.⁷ This paper explores how corporate casinos affect employment in six Midwestern counties using various employment data and forecasting models. Changes in both household and payroll employment are examined to separate the effects on the residents and businesses in counties with casinos. Payroll employment changes may allude to possible interindustry substitution resulting from casino gaming. Also, both urban and rural "casino counties" are used in the analysis to

⁵ Indian tribes use their gaming revenue to promote economic development on their reservations. Economic development from corporate casino gaming, however, has the potential to affect a much greater population.

⁶ States have negotiated payments from tribes in return for certain services such as security and improving and maintaining highway access to their casinos.

⁷ Gazel, Thompson, and Rickman (1995), KPMG Management Consulting (1995), and Blois, Cunningham, and Lott (1995).

Thomas A. Garrett is a senior economist at the Federal Reserve Bank of St. Louis. Molly D. Castelazo provided research assistance.

Federal Reserve Bank of St. Louis *Review*, January/February 2004, 86(1), pp. 9-22.
© 2004, The Federal Reserve Bank of St. Louis.

Table 1**State Gaming Summary 2001**

State	Land-based casinos	Riverboat casinos	Native American casinos
Arizona			X
California			X
Colorado			X
Connecticut			X
Florida			X
Idaho			X
Illinois		X	
Indiana		X	
Iowa		X	X
Kansas			X
Louisiana	X	X	X
Michigan	X		X
Minnesota			X
Mississippi		X	X
Missouri		X	
Montana			X
Nebraska			X
Nevada	X		X
New Jersey	X		
New Mexico			X
New York			X
North Carolina			X
North Dakota			X
Oklahoma			X
Oregon			X
South Dakota			X
Texas			X
Washington			X
Wisconsin			X
Total	4	6	25

SOURCE: *Bear Stearns 2002-2003 North American Gaming Almanac* (Ader, 2003, p. 16). The above list does not include those states with casinos operating as part of a state lottery. States not listed have no corporate or tribal casinos.

Table 2**Gaming Revenue for Selected States**

State	2001 revenue (\$ millions)	2000 revenue (\$ millions)	Percent change
Colorado	\$675.3	\$631.7	6.9%
Connecticut	1,401.6	1,308.7	7.1
Illinois	1,783.8	1,657.8	7.6
Indiana	1,841.8	1,689.7	9.0
Iowa	922.9	892.6	3.4
Louisiana	1,883.2	1,708.9	10.2
Michigan	1,007.4	742.9	35.6
Mississippi	2,700.8	2,650.4	1.9
Missouri	1,137.1	996.6	14.1
Nevada	9,466.9	9,599.4	-1.4
New Jersey	4,303.9	4,299.6	0.1
Total	27,124.7	26,178.4	3.6

NOTE: Tribal and corporate casino revenues are considered in the above figures and represent revenues to the casinos after subtracting player winnings.

SOURCE: *Bear Stearns 2002-2003 North American Gaming Almanac* (Ader, 2003, p. 6).

distinguish any potential differences in employment effects and to address the typical perception that casino gaming is more often seen as a savior for rural communities rather than urban communities.

The Spread of Casino Gaming

Nevada was the first state to legalize casino gaming in 1931 and has the largest gaming market in the country. The 210 casinos in Nevada generated over \$9.5 billion in revenue during 2001. The largest concentration of casinos is in Las Vegas, with 14 casinos downtown and 47 on the “Strip” amassing nearly \$5.3 billion in revenue and attracting 35 million visitors annually to fill over 100,000 hotel rooms. Hotels downtown and on the Strip have 75,000 electronic gaming devices (EGDs—which are slot machines, video poker games, and any other electronic game used for wagering) and 3,300 table games that take up 3.3 million square feet of casino floor space. Other major markets in Nevada include Reno (\$1 billion in revenue), Laughlin (\$500 million in revenue), and Lake Tahoe (\$330 million in revenue).

In 1976, New Jersey became the second state to legalize casino gaming, but restricted the activity to Atlantic City. Today there are 13 casinos in Atlantic City generating nearly \$4.3 billion in annual revenue

and attracting 32 million visitors, making Atlantic City the second largest casino gaming market in the United States. The Atlantic City market is generally characterized as a “day-trip” destination, whereas Las Vegas is typically considered a vacation destination. Atlantic City casinos have 12,000 hotel rooms and offer 37,000 EGDs, over 1,200 table games, and nearly 1.3 million square feet of casino floor space.

The early 1990s saw a marked increase in the number of states that legalized casino gaming. Riverboat casino gaming first began in Iowa and Illinois in 1991 and quickly spread throughout the Midwest.⁸ Riverboat gaming now also exists in Indiana, Mississippi, and Missouri. Louisiana and Michigan legalized land-based casino gaming within the last decade. States cite the gaming industry’s potential to create economic growth in the state as the primary reason for their approval of corporate casino gaming. The greatest perceived benefits are employment growth, greater tax revenue to state and local governments, and growth in local retail sales. In addition, the increasing fiscal pressures on state budgets during the 1990-91 recession, the fear

⁸ In many cases, the riverboat casinos are “fixed in dock,” meaning that they cannot move freely along a river.

of lost revenue to neighboring states' casinos, and a more favorable public attitude regarding casino gaming have all increased the appeal and acceptance of casinos.

Saddled with their current state budget crises, state legislators have taken up the issue of casino gaming with a renewed interest. The National Conference of State Legislatures reported in April 2003 that collective state budget gaps will approach \$70 billion in fiscal year 2004.⁹ As seen in Table 2, casino revenues are sizeable, making them an attractive revenue source for state and local governments. Many states are considering the expansion of casino gaming, while others such as Pennsylvania are debating whether to introduce slot machines at pari-mutuel racetracks. Several states with casino gaming have increased casino gaming tax rates within the past year or two. Also, states with Indian gaming are considering measures to extract revenue from traditionally tax-exempt Indian casinos. However, the direct taxation of tribal gaming revenue is likely to be met with serious legal challenges involving the sovereignty of Indian tribes.

CASINOS AND EMPLOYMENT

Local officials and casino proponents often claim that casinos increase local employment simply because they create additional jobs within the local area. However, several factors should be considered when evaluating the employment effects of casino gaming. These factors are applicable to any business or industry, not just casino gaming.

The relationship between casinos and employment involves the location of the casino and the required skill level of its work force. The general premise is that casinos increase employment because a casino's operation requires labor and this labor will come from the local area, thus reducing local unemployment. The question to ask is not only whether casinos decrease unemployment, but also for *whom* they decrease unemployment. Most casino jobs require some skill, be it accounting, dealing cards, security, or other expertise. If a casino is planning to move to a rural area that has a relatively less-skilled work force, the casino probably will draw skilled labor from outside of the area. If this labor remains outside of the local area and workers commute to the casinos, then unemployment in the local area will remain unchanged. If some of this skilled

labor decides to move near the casino, then the unemployment rate in the local area will fall because the labor force has increased.¹⁰ However, unemployment for the original population has remained essentially unchanged—only the new arrivals have found employment with the casino. It is the employment of these new arrivals that has decreased the unemployment rate. Thus, the promise of increased employment for the original population, which is used as an argument for the construction of casinos, may not be realized. In a relatively urban area, there is probably enough variety in the work force to ensure that skilled labor will be provided locally. In rural areas, however, most labor may be from outside of the area, thus leaving unemployment for the original population unchanged.

While casino employment is used as an indicator of economic development, true economic development occurs only when there is increased value to society.¹¹ The introduction of casino gaming may cause local businesses to close, which will result in layoffs.¹² The net increase in employment in the local area may thus be less than the number of new casino jobs. However, casino gaming may increase total employment when casinos indirectly generate non-casino jobs in the local area as a result of increased demand for non-casino goods and services. Casino employees who were previously unemployed or who recently moved into the area now generate income, and this income will be spent on goods and services such as housing and entertainment. An increase in demand for these services will increase firms' demand for labor, thereby increasing employment.

These employment "spillovers" essentially result in a positive or negative multiplier effect. The degree of this multiplier effect has been disputed in the literature. Research by Gazel, Thompson, and Rickman (1995), KPMG Management Consulting (1995), and Blois, Cunningham, and Lott (1995) suggests that a positive multiplier effect exists. However, studies by Anders, Siegel, and Yacoub (1998), Przybylski et al. (1998), and Siegel and Anders (1999) provide evidence that there is, at least to some degree,

⁹ Of this total, California accounts for \$26 billion. See the National Conference of State Legislatures *State Budget Update* (April 2003).

¹⁰ Recall that the unemployment rate is the number of unemployed divided by the labor force (employed + unemployed).

¹¹ An evaluation of the social welfare effects of casino gaming should consider the benefits of casino gaming beyond employment, as well as possible costs such as addiction and crime. See Grinols (forthcoming).

¹² The influence of casino gaming on businesses and employment certainly reaches beyond county borders. Thus, economic development in one county could come at the expense of a reduction in economic activity in a neighboring county.

Table 3**Statistics for Counties with Casinos**

	Warren County, MS	Tunica County, MS	Massac County, IL	St. Clair County, IL	Lee County, IA	St. Louis County, MO
Casino employment	2,443	12,689	883	1,184	367	2,050
Number of casinos in county	4	9	1	1	1	1
County employment	25,030	5,636	7,665	108,270	16,708	540,981
County population	49,343	9,635	15,081	256,599	33,313	1,015,417
Percent (number) of casino employees from home county	75 (1,832)	30 (3,807)	44 (389)	80 (947)	45 (165)	32 (656)
County unemployment rate, pre-casino (%)	8.8	10.7	9.5	9.4	3.7	3.1
County unemployment rate, post-casino (%)	4.2	7.1	4.4	5.9	3.3	3.9
State unemployment rate, pre-casino (%)	8.2	8.2	7.5	7.5	3.7	4.2
State unemployment rate, post-casino (%)	5.5	5.5	5.4	5.4	3.3	4.7
Employment/population ratio, pre-casino	0.430	0.302	0.381	0.400	0.477	0.534
Employment/population ratio, post-casino	0.507	0.591	0.509	0.418	0.449	0.531
First casino opened	2/93	8/92	2/93	7/93	11/94	3/97

NOTE: All employment data are from December 2001, and population data are from 1999. Home-county casino employment was obtained from contacting casinos. For multiple casinos in a county, the above figure represents the county average.

interindustry substitution between casino gaming and certain local businesses.

Local Employment—Empirical Analysis

Several academic studies have explored the impact of casinos on local employment. While their conclusions are somewhat mixed, the studies generally suggest that casinos do increase employment in the local area (or at least they do not lead to an employment decrease). Grinols (1994) studied Illinois casinos and found that, of the eight casinos in the state, six have had no significant impact on total employment since their introduction. Using a different methodology, Hewings, Schindler, and Nafziger (1996) found that Illinois casinos generated over 17,000 new jobs. In a study of Colorado casinos, the Center for Business and Economic Forecasting (1995)

found that Native American gaming led to 6,100 new jobs. Leven and Phares (1997) found that nearly 12,200 new jobs were created as a result of Missouri casinos.

This section presents two analyses of county employment changes after the introduction of casinos. The first analysis uses monthly household employment data to explore the effect of casinos on resident employment in each county; the second analysis uses annual payroll employment data to detect employment changes in specific industries. It is important to consider that household and payroll employment data measure employment in different ways, so the figures for each will be neither equal nor directly comparable. Household employment is derived from a survey of households and is the number of people in the county who are employed, regardless of where they are employed; county pay-

roll employment is derived from a survey of firms and is the number of jobs in the county.

The counties used in the analyses are Warren County (Vicksburg casino market) and Tunica County (Tunica casino market) in Mississippi; Massac County (one casino) and St. Clair County (one casino) in Illinois; Lee County (one casino) in Iowa; and St. Louis County (one casino) in Missouri. Of these six counties, St. Louis and St. Clair counties are classified as urban counties and the rest as rural. Detailed employment statistics for each of the six counties are shown in Table 3.

For the first analysis, total household employment is compared before and after casino introduction. For each county, an empirical model is developed to capture employment changes several years before casino introduction. These changes are then used to forecast employment changes from the date of casino introduction through December 2001 (the end of the sample period). These forecasts represent the level of employment that would have existed if the casinos had not been opened. The difference between the actual and forecasted employment is the estimated effect of the casinos.

The second analysis uses payroll employment data to compare county employment in construction, manufacturing, retail trade, services, and finance before and after casino introduction.¹³ If casinos cause an influx of new businesses and/or residents to the county, employment in these sectors may have increased since the introduction of casino gaming. Based on previous studies, this may be especially true for services and retail trade employment. Conversely, if casinos cannibalize existing retail and services sector businesses, then employment in these sectors may have decreased since casino gaming was introduced.

Data and Methodology for Household Employment Forecasts. Seasonally adjusted monthly household employment was obtained from the U.S. Bureau of Labor Statistics for each of the six counties over the period 1986:01-2001:12. Since the six counties introduced casino gaming in the early to mid-1990s, the length of the data series was chosen to ensure an adequate sample of observations pre- and post-casino adoption.

The behavior of the employment series for each county prior to casino adoption (see bottom of

Table 3 for casino opening dates) is captured using an ARIMA(p,d,q) model, which is defined as

$$x(t) = \gamma + \alpha_1 x(t-1) + \dots + \alpha_p x(t-p) + e(t) + \beta_1 e(t-1) + \dots + \beta_q e(t-q),$$

where x is county household employment, γ is a constant term, $e(t)$ is the error term, p is the number of autoregressive lags, and q is the number of moving-average lags. Augmented Dickey-Fuller (ADF) tests for stationarity were conducted on the employment series (pre-casino adoption) for each of the six counties to determine the order of integration (d).¹⁴ The ADF tests reveal that employment for St. Clair County and St. Louis County is stationary in levels, employment for Lee County is trend stationary (a linear time trend is included in the above equation), and employment for Tunica, Massac, and Warren counties is first-difference stationary (where $x(t)$ becomes $\Delta x(t) = x(t) - x(t-1)$).

The Akaike information criterion was used to determine the model order for each county's employment series. The appropriate ARIMA models are as follows: Tunica County ARIMA(2,1,1), Massac County ARIMA(1,1,0), Warren County ARIMA(2,1,0), St. Clair County ARIMA(1,0,4), Lee County ARIMA(1,0,0), and St. Louis County ARIMA(1,0,2).¹⁵

Visual inspection of the St. Louis County and St. Clair County employment series reveals marked business cycle effects. No discernable effects are present for the four rural counties. To capture these effects in the empirical models, the coincident index for both Missouri and Illinois is included as a variable in their respective ARIMA model. Based on model order tests, contemporaneous values of the Illinois coincident index are included in the St. Clair County model and contemporaneous and one-lag values of the Missouri coincident index are included in the St. Louis County model.¹⁶

¹⁴ The ADF test results are available from the author.

¹⁵ Following Perron (1989), the ARIMA models also account for, if necessary, structural changes in the employment series prior to casino introduction by including binary dummy variables. Visual inspection of the data reveals structural breaks in the Massac County and St. Louis County data. Thus, the equation for Massac County includes a binary dummy variable that has a value of 1 for 1990:01 and the St. Louis County equation includes a binary dummy variable that has a value of 1 over the period 1990:01-1993:01.

¹⁶ The coincident index for each state weights changes in payroll employment, average hours worked in manufacturing, the unemployment rate, and real wages paid. Monthly state-level coincident indices are available from the Federal Reserve Bank of Philadelphia at www.phil.frb.org/econ/stateindexes/index.html. See Crone (2002) for a discussion of regional coincident indices.

¹³ The analysis does not consider employment in agriculture, government, or transportation because employment in these sectors is unlikely to be significantly affected by casino gaming. A description of the businesses included in each of the five sectors can be found at www.census.gov/epcd/cbp/map/97data/29/189.txt.

Figure 1

Warren County, MS—Household Employment

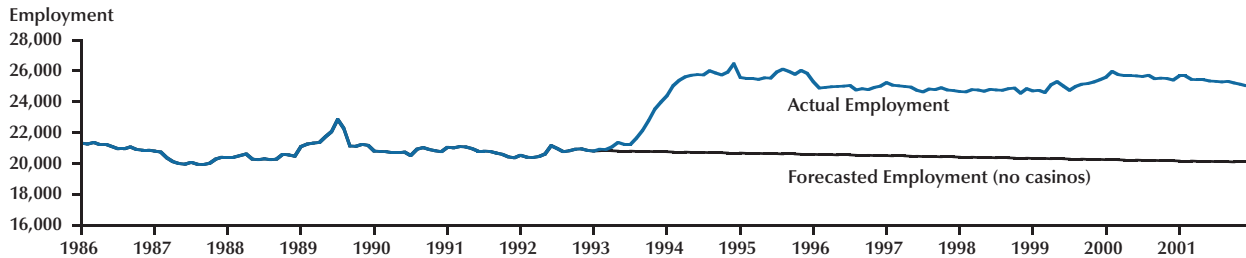
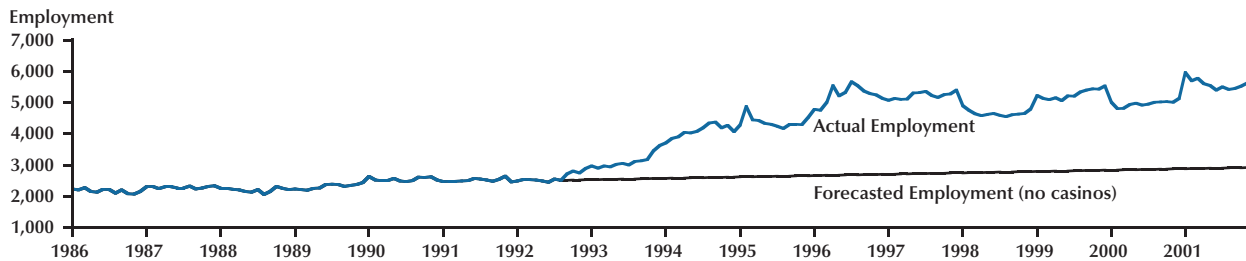


Figure 2

Tunica County, MS—Household Employment



Using the coefficient estimates from the ARIMA models, employment is forecasted dynamically from the month of casino adoption through 2001:12 and is compared with actual employment since the beginning of casino operation.¹⁷

Results for Household Employment Forecasts

Warren and Tunica Counties, Mississippi. Actual and forecasted household employment for Warren and Tunica counties in Mississippi are shown in Figures 1 and 2. Both are rural counties, and casino gaming constitutes a major industry in each county. The figures reveal that employment in both counties significantly increased since the adoption of casino gaming. There was a dramatic jump in employment in Warren County in 1993 and 1994, the two years in which casinos began operations. Since that time, employment growth has been relatively flat. Employment in Tunica County has grown steadily since the first casino was introduced in late 1992, reflecting the steady increase in the number of casinos in Tunica County throughout the middle and late 1990s. Forecasted employment for the 1990s reveals that,

without casino gaming, employment would have decreased slightly in Warren County (about 7 jobs per month), but would have increased slightly in Tunica County (about 3 jobs per month).

As of December 2001, Tunica County household employment increased by 3,144 since the introduction of the first casino, while the population increased by 1,172. Warren County employment increased by 4,225 since the introduction of the first casino, while its population increased by 910. Therefore, much of the increase in household employment occurred for pre-casino residents rather than new residents. The employment-to-population ratios for both counties have also increased since the introduction of casino gaming (see Table 3). The employment-to-population ratio increased by nearly 27 percentage points in Tunica County and by over 7 percentage points in Warren County. Casino employment in Tunica County is greater than the population of the county, so the bulk of employees who work in Tunica casinos live outside of the county. In Warren County, total casino employment for residents is about 1,800, but the increase in employment since casino introduction was nearly 5,000 with little change (910) in population. This suggests that over the

¹⁷ Estimates from the ARIMA models are available from the author.

Figure 3

St. Clair County, IL—Household Employment

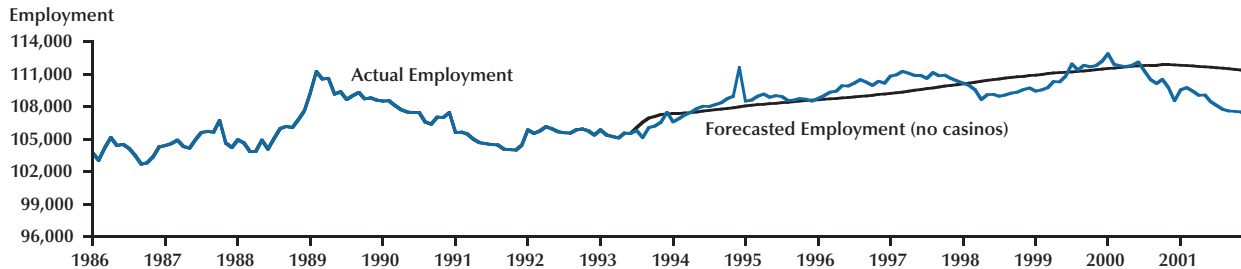
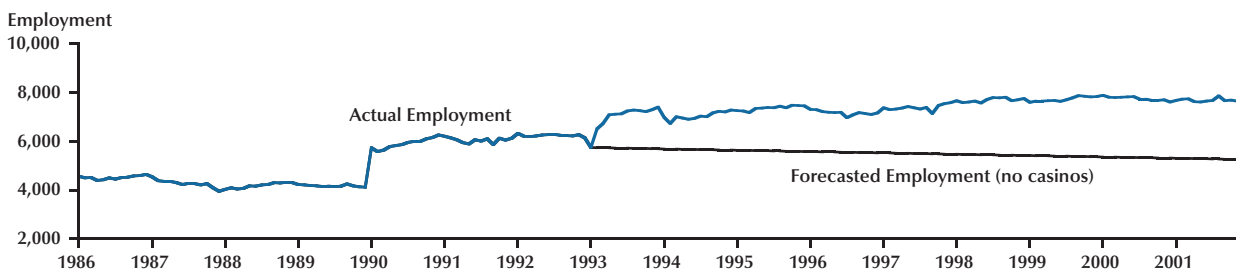


Figure 4

Massac County, IL—Household Employment



sample period there was employment growth in Warren County outside of the casino industry.

St. Clair and Massac Counties, Illinois. Figures 3 and 4 show actual and forecasted employment for urban St. Clair County and rural Massac County in Illinois. St. Clair County employment is relatively more volatile. In a county with over 250,000 people and household employment near 108,000, the exact impact of one casino on local employment is hard to determine given the relative smallness of casino employment to total employment and the volatility of overall county employment. Prior to 2000, actual and forecasted employment trended upward. Beyond this point, however, actual employment fell below the forecasted decrease in employment. Total employment has risen 1,601 since the introduction of the casino (the casino employs 1,184 people, 947 of whom are from St. Clair County), but the population of St. Clair County has decreased nearly 7,500. As a result, the employment-to-population ratio has increased slightly from the pre-casino period (1.7 percentage points). It thus appears that casino gaming has not hurt St. Clair County employment, but the volatility of total employment and the loss in

population leads one to question the overall ability of one casino to maintain or foster employment growth in an urban area.

Employment in rural Massac County markedly increased when the casino began operations and has increased steadily since then. Without the introduction of casino gaming, employment forecasts show a decrease at a rate of about 5 jobs per month. By the end of 2001, actual employment was higher than forecasted employment, but the growth in actual employment has been relatively slow since the introduction of casino gaming. Employment increased by 1,927 since the introduction of casino gaming, which employs 389 persons from Massac County, and the population of Massac County increased by 18. As in Warren and Tunica counties, the vast majority of employment growth in Massac County occurred for pre-casino residents rather than new arrivals. In addition, the bulk of employment growth occurred outside of the casino industry. The employment-to-population ratio for Massac County increased nearly 13 percentage points since the introduction of casino gaming.

Lee County, Iowa. Forecasted and actual employ-

Figure 5

Lee County, IA—Household Employment

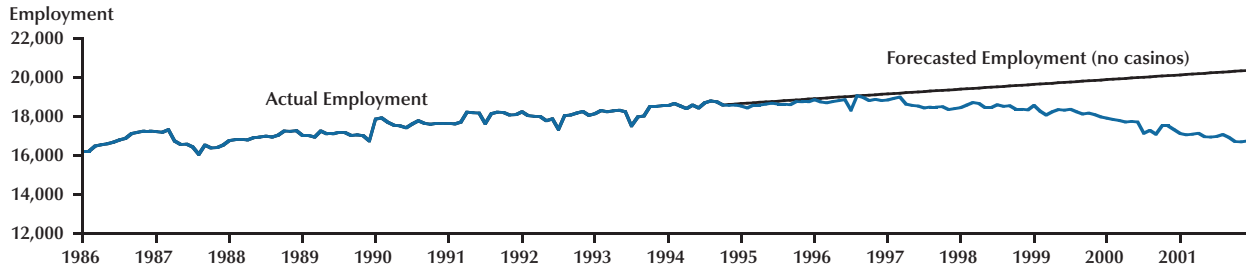


Figure 6

St. Louis County, MO—Household Employment



ment for rural Lee County, Iowa, are shown in Figure 5. Actual employment remained relatively constant around the time the casino began operations but has steadily decreased since then. Forecasted employment continued a gradual increase (about 20 jobs per month) since the date of casino introduction. At the end of 2001, Lee County had lost 1,846 jobs since the casino began operations and experienced a population decrease of 1,652. As a result, the employment-to-population ratio decreased by nearly 3 percentage points since the casino was introduced. Unlike rural counties such as Massac County in Illinois and Tunica and Warren counties in Mississippi, the introduction of casino gaming in Lee County has not corresponded to an increase in employment. It is possible, however, that the introduction of casino gaming has slowed the decrease in employment and population in Lee County.

St. Louis County, Missouri. St. Louis County’s total household employment is nearly 550,000. As in urban St. Clair County, household employment in St. Louis County is quite variable over the sample period (Figure 6). The impact of one casino employ-

ing 2,050 people, only 32 percent of whom are from St. Louis County, cannot be accurately inferred from the data. Employment continued to fall after the casino was introduced but then slightly increased above forecasted levels in 2000. It is possible that the casino has created some jobs, but the direct impact of the casino on total employment is masked by highly variable total employment and the relatively small employment contribution made by a single casino.

Payroll Employment Changes by Sector.

Because the payroll employment data are listed on an annual basis, this study’s small sample size is not adequate for running forecast models. Thus, the analysis involves comparing employment levels in each sector before and after the introduction of casinos. Sector employment changes in these two periods for each of the six counties are shown in Tables 4 through 9. For each county, services sector employment excludes casino employment, which is listed as a separate sector. Recall that changes in sector employment cannot be directly compared with household employment changes in the previous section because the two employment measures are different.

Table 4**Warren County, MS, Payroll Employment Changes by Sector**

Sector	Employment		
	1993	2001	Gain or loss (% change)
Manufacturing	3,408	5,090	1,682 (49.4)
Retail trade	3,926	2,932	-994 (-25.3)
Services	4,114	7,674	3,560 (86.5)
Financial	597	487	-110 (-18.4)
Construction	363	590	227 (62.5)
Casinos	—	2,443	2,443
Total gain or loss			6,808 (54.9)

NOTE: The first casino opened in February 1993. Services sector employment excludes casino employment.

SOURCE: Data are from the U.S. Census Bureau. 1993: www.census.gov/epcd/cbp/map/93data/28/149.txt.

2001: www.census.gov/epcd/cbp/map/01data/28/149.txt.

Table 5**Tunica County, MS, Payroll Employment Changes by Sector**

Sector	Employment		
	1992	2001	Gain or loss (% change)
Manufacturing	375	614	239 (63.7)
Retail trade	204	374	170 (83.3)
Services	123	2,441	2,318 (1884.5)
Financial	77	99	22 (28.6)
Construction	15	60	45 (300.0)
Casinos	—	12,689	12,689
Total gain or loss			15,483 (1950.0)

NOTE: The first casino opened in August 1992. Services sector employment excludes casino employment.

SOURCE: 1992 data are from GovStats: <http://govinfo.kerr.orst.edu/php/commerce/state/show.php>. 2001 data are from the U.S. Census Bureau: www.census.gov/epcd/cbp/map/01data/28/143.txt.

Warren County, Mississippi (Table 4), experienced a large increase in manufacturing, services, and construction employment since the introduction of casino gaming, which constitutes 36 percent of the total increase in payroll employment. The increase in manufacturing employment is quite large, given the national decrease in manufacturing employment during the 1990s. Moderate decreases in retail trade and financial employment occurred within the county over the same time period.

Because casino gaming is a relatively large industry in Warren County, the findings suggest that the increase in services sector employment and decrease in retail trade employment may be attributed to casino gaming.

Tunica County, Mississippi (Table 5), had employment increases in all five sectors, with the largest increase in the services sector. Given that casino gaming is the predominant industry in Tunica County, the data in Table 5 suggest that employment

Table 6**St. Clair County, IL, Payroll Employment Changes by Sector**

Sector	Employment		
	1993	2001	Gain or loss (% change)
Manufacturing	7,318	6,724	-594 (-8.1)
Retail trade	18,154	12,872	-5,282 (-29.1)
Services	25,922	39,102	13,180 (50.8)
Financial	3,158	4,274	1,116 (35.3)
Construction	2,626	3,927	1,301 (49.5)
Casino	—	1,184	1,184
Total gain or loss			10,905 (19.1)

NOTE: The casino opened in July 1993. Services sector employment excludes casino employment.

SOURCE: Data are from the U.S. Census Bureau. 1993: www.census.gov/epcd/cbp/map/93data/17/163.txt.

2001: www.census.gov/epcd/cbp/map/01data/17/163.txt.

Table 7**Massac County, IL, Payroll Employment Changes by Sector**

Sector	Employment		
	1993	2001	Gain or loss (% change)
Manufacturing	736	638	-98 (-13.3)
Retail trade	695	353	-342 (-49.2)
Services	874	1,379	505 (57.8)
Financial	108	198	90 (83.3)
Construction	126	100	-26 (-20.6)
Casino	—	883	883
Total gain or loss			1,012 (39.9)

NOTE: The casino opened in February 1993. Services sector employment excludes casino employment.

SOURCE: Data are from the U.S. Census Bureau. 1993: www.census.gov/epcd/cbp/map/93data/17/127.txt.

2001: www.census.gov/epcd/cbp/map/01data/17/127.txt.

increases in the various sectors can be attributed, in large part, to the introduction of casino gaming. Overall payroll employment increased by over 1,900 percent since casino gaming was introduced in 1992; a large portion of this increase (82 percent) is attributed to casino employment.

Casino gaming in St. Clair County, Illinois (Table 6), contributed to roughly 11 percent of the gain in payroll employment since casino gaming

was introduced. Services, finance, and construction employment all increased by an average of 43 percent, but manufacturing and retail trade decreased by 8 percent and 29 percent, respectively. The employment impact of casino gaming has been much smaller than changes in other sectors, but it still has contributed moderately to net changes in total payroll employment.

Massac County, Illinois (Table 7), experienced

Table 8**Lee County, IA, Payroll Employment Changes by Sector**

Sector	Employment		
	1994	2001	Gain or loss (% change)
Manufacturing	5,991	5,130	-861 (-14.4)
Retail trade	3,144	2,057	-1,087 (-34.6)
Services	3,573	5,229	1,656 (46.3)
Financial	482	558	76 (15.8)
Construction	796	669	-127 (-16.0)
Casino	—	367	367
Total gain or loss			24 (0.17)

NOTE: The casino opened in November 1994. Services sector employment excludes casino employment.

SOURCE: 1994 data are from GovStats: <http://govinfo.kerr.orst.edu/php/commerce/state/show.php>. 2001 data are from the U.S. Census Bureau: www.census.gov/epcd/cbp/map/01data/19/111.txt.

Table 9**St. Louis County, MO, Payroll Employment Changes by Sector**

Sector	Employment		
	1997	2001	Gain or loss (% change)
Manufacturing	97,608	59,048	-38,560 (-39.5)
Retail trade	113,407	74,170	-39,237 (-34.6)
Services	207,947	287,982	80,035 (38.5)
Financial	45,162	49,869	4,707 (10.4)
Construction	32,087	39,876	7,789 (24.3)
Casino	—	2,050	2,050
Total gain or loss			16,784 (3.38)

NOTE: The casino opened in March 1997. Services sector employment excludes casino employment.

SOURCE: Data are from the U.S. Census Bureau. 1997: www.census.gov/epcd/cbp/map/97data/29/189.txt. 2001: www.census.gov/epcd/cbp/map/01data/29/189.txt.

an increase in services and financial employment but a decrease in other sectors. With the introduction of casino gaming, payroll employment increased nearly 40 percent. Casino gaming has provided the largest contribution to the increase in total payroll employment in Massac County (87 percent). Without it, the gain in total payroll employment in Massac County would have been roughly 130 persons.

Without casino gaming, Lee County, Iowa

(Table 8), would have experienced an overall decrease in payroll employment since casino introduction. While services and financial employment increased over the sample period, these increases were met by larger decreases in manufacturing, retail trade, and construction employment, resulting in the loss of 343 jobs. Casino employment of 367 persons provided a net gain of 24 jobs in Lee County. Household employment and population fell for Lee County,

but payroll employment remained relatively constant since casino introduction. This suggests that either some of the original population in Lee County moved outside of the county and continues to work in Lee County or more residents of neighboring counties now work in Lee County.

Like St. Clair County in Illinois, the casino industry is a relatively minor employer in St. Louis County, Missouri (Table 9). However, the 2,050 casino jobs contributed to roughly 12 percent of the increase in total payroll employment in St. Louis County (similar to the 11 percent in urban St. Clair County). Large decreases in manufacturing and retail trade occurred, but these decreases were met with slightly larger increases in services, financial, and construction employment. Thus, even though casino gaming may be a minor industry in urban areas, casino gaming can make up a moderate portion of net payroll employment gains or losses.

SUMMARY AND CONCLUSIONS

The employment effects of casino gaming are difficult to quantify. A casino may draw labor from outside of the local area, thus leaving local employment conditions unchanged if that labor does not relocate to the local area. Casinos are only synonymous with economic development if they create a greater value to society. It is possible that casino gaming may reduce employment in other local industries if consumers substitute casino gaming for other consumption. The net effect of gaming could be positive or negative depending upon the degree to which casino gaming substitutes for or complements consumption at other local businesses.

Determining the possible impact of casino gaming on local employment involves an examination of employment changes in the local area before and after the introduction of casino gaming. The empirical analysis presented here reveals that, in three of four cases, rural counties that adopted casino gaming experienced increases in household and payroll employment. This seems to hold even though casino employment is dispersed over several counties rather than just the home county. Also, employment gains are much greater in rural counties that have adopted casino gaming as a major or predominant industry. It is harder to detect the impact of casino gaming in more-metropolitan counties, where employment is highly variable and casino gaming constitutes a small portion of total employment. However, casino gaming in urban areas can still constitute a moderate portion of net payroll employ-

ment gains or losses even though casino gaming is a minor industry; still, the impact is much greater in rural counties with casino gaming.

One question that remains is, How much will the gaming industry grow in the future? The current budget crises facing state and local governments may generate further expansion of casino gaming across the country. There is little evidence that the industry has reached the saturation point—a “build it and they will come” attitude pervades the industry at the current time. While the evidence here suggests that rural counties that adopt casino gaming as a major industry are likely to see large employment gains, this does not suggest that every county can become like Tunica, Mississippi. Attitudes regarding the spread of casino gaming in a given local area, costly industry regulation, and increasing casino competition may hinder the growth of gaming in rural areas.

The degree to which state and local governments currently rely on casino revenue raises the question of whether or not the casino industry is recession-proof. One may expect that the growth of the casino industry is contingent upon economic conditions; if the industry is highly procyclical, then casino revenues may do little to lessen the budgetary impacts of an economic slowdown. This may be true: In fact, many states with casinos are facing budget crises similar to those of states without casinos. However, little research has been done on this issue. Regardless of what the future holds, there is little doubt that casinos are here to stay and that more communities will be faced with the question of whether to adopt casino gaming.

REFERENCES

- Ader, Jason N. *Bear Stearns 2002-2003 North American Gaming Almanac*. Las Vegas: Huntington Press, 2003.
- Anders, Gary; Siegel, Donald and Yacoub, Munther. “Does Indian Casino Gambling Reduce State Revenues? Evidence from Arizona.” *Contemporary Economic Policy*, July 1998, 16(3), pp. 347-55.
- Blois, T.; Cunningham, S. and Lott, W. “The Bridgeport Casino Proposals: An Economic Evaluation.” Prepared for the Connecticut Department of Economic and Community Development and the Connecticut Department of Special Revenue. Storrs, CT: Connecticut Center for Economic Analysis, October 1995.

- Center for Business and Economic Forecasting, Inc. *Economic Impact of Limited Gaming in Colorado*. Report prepared for the Colorado Casino Owners Association, Regis University, January 1995.
- Crone, Theodore. "Consistent Economic Indexes for the 50 States." Working Paper No. 02-7/R, Federal Reserve Bank of Philadelphia, May 2002.
- Gazel, R.; Thompson, W. and Rickman, D. "The Economic Impact of Native American Gaming in Wisconsin." *Gaming Research and Review Journal*, 1995, 2(2), pp. 43-62.
- GovStats. <http://govinfo.kerr.orst.edu/php/commerce/state/show.php>.
- Grinols, Earl. "Bluff or Winning Hand? Riverboat Gambling and Regional Employment and Unemployment." *Illinois Business Review*, Spring 1994, 51(1), pp. 8-11.
- Grinols, Earl. "Cutting the Cards and Craps: Right Thinking About Gambling Economics" in G. Reith, ed., *Gambling: Debating the Issues*, Prometheus Contemporary Issues Series. New York: Prometheus (forthcoming).
- Hewings, G.; Schindler, G. and Nafziger, B. "The Impact of Riverboat Casino Gambling on the Illinois Economy 1991-1995." Report to the Illinois Gaming Board. Chicago: Regional Economics Applications Laboratory, 1996.
- KPMG Management Consulting. *One Year Review of Casino Windsor*. Report prepared for Ontario Casino Corporation, Toronto, 1995.
- Leven, Charles and Phares, Donald. "Casino Gaming in Missouri: The Spending Displacement Effect and Net Economic Impact." Proceedings of the 90th Annual Conference on Taxation, National Tax Association, Chicago, November 1997, pp. 431-38.
- National Conference of State Legislatures. *State Budget Update*. Washington, DC: National Conference of State Legislatures, February 2003.
- National Gambling Impact Study Commission. *Final Report*. Washington, DC: National Gambling Impact Study Commission, 1999.
- Perron, Pierre. "The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis." *Econometrica*, November 1989, 57(6), pp. 1361-401.
- Przybylski, Michael; Felsenstein, Daniel; Freeman, Daniel and Littlepage, Laura. "Does Gambling Complement the Tourist Industry? Some Empirical Evidence of Import Substitution and Demand Displacement." *Tourism Economics*, September 1998, 4(3), pp. 213-31.
- Siegel, Donald and Anders, Gary. "Public Policy and the Displacement Effects of Casinos: A Case Study of Riverboat Gambling in Missouri." *Journal of Gambling Studies*, Summer 1999, 15(2), pp. 105-21.
- U.S. Census Bureau. www.census.gov.

U.S. Regional Business Cycles and the Natural Rate of Unemployment

Howard J. Wall and Gylfi Zoëga

In the late 1960s, Milton Friedman and Edmund Phelps convinced the economics profession that there was no long-run trade-off between inflation and unemployment. A policy that tries to maintain the unemployment rate below a certain threshold (dubbed the natural rate of unemployment by Friedman) would lead to rising inflation, while trying to maintain it above the threshold would lead to ever-declining rates of inflation. The proposition of long-run neutrality of inflation and money growth soon gained wide acceptance, and work in this area has focused on making the natural rate of unemployment fully endogenous in general-equilibrium models (Pissarides, 2000; Layard, Nickell, and Jackman, 1991; and Phelps, 1994). This theory can be used to show how a variety of macroeconomic shocks—such as the rate of technical progress, real interest rates, and oil prices—affect the natural rate and social welfare.

Inflation-targeting central banks often monitor employment and wage changes in the hope of preventing wage inflation in the labor market from generating general price inflation.¹ The use of the notion of an equilibrium level of unemployment that is independent of current and past monetary variables has made the estimation of the natural rate important. This practice relies on representative-agent type models—the ones used to provide microeconomic foundations for the inflation/unemployment relationship—to assess the state of the economy using aggregate data. A central banker may then use data on aggregate employment, unemployment, and average wage inflation across all sectors of the economy to assess the position of the economy in relation to an estimate of the natural

rate of unemployment. Most often, the estimate is the implied natural rate in an econometric model of the aggregate Phillips curve.

The objective of this article is to show that the sole reliance on aggregate data may lead to incorrect inferences about the natural rate of unemployment. We show how regional business cycles might affect aggregate wage inflation and how attention to regional labor market trends can be useful for understanding the aggregate labor market. Moreover, we show how the natural rate of unemployment may depend directly on the dispersion of economic activity across regions.

Our regional approach is in some ways parallel to the sectoral approach of Lilien (1982), Abraham and Katz (1986), and Brainard and Cutler (1993). Lilien (1982) found that a measure of sector-specific disturbances accounted for a significant portion of the variation in aggregate employment: When an industry sheds redundant labor in less time than it takes for the affected workers to find employment elsewhere, unemployment rises when the pace of sectoral reallocation of labor (and capital) increases. Abraham and Katz (1986) pointed out that Lilien's estimates might exaggerate the role of sectoral disturbances by failing to take into account differences in the sensitivity of different industries to macroeconomic shocks. Brainard and Cutler (1993) developed a data series to measure the intensity of reallocation shocks. They constructed a time series of the variance of sectoral stock market excess returns and found that they had a modest—though statistically significant—role in explaining aggregate employment fluctuations.

Our paper follows recent work illustrating the significant regional differences in economic conditions, business cycle dynamics, and reactions to monetary policy. Overman and Puga (2002) demonstrate the increased polarization of unemployment

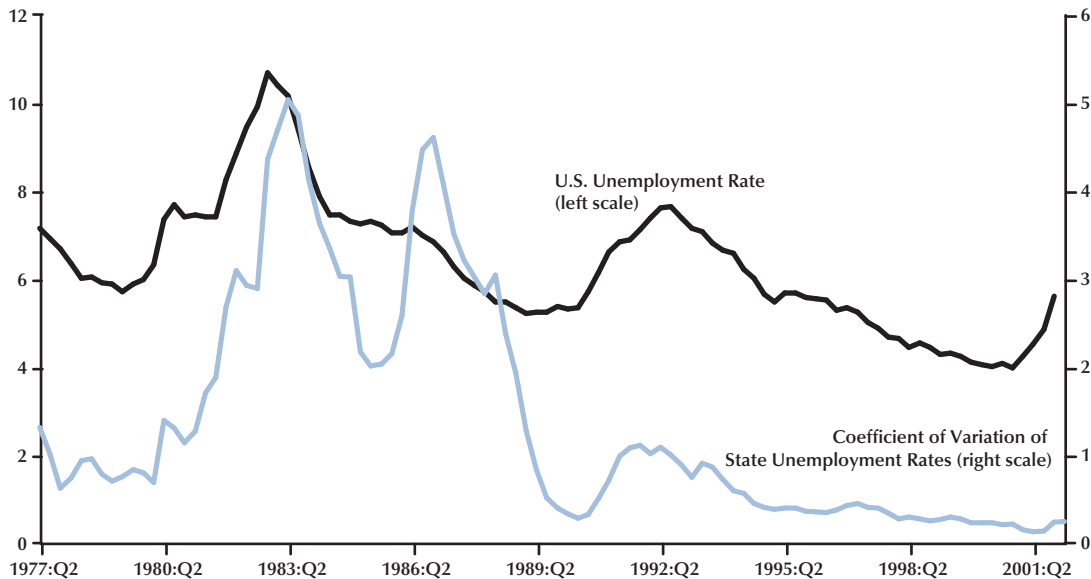
¹ Such considerations have led to the appointment of a labor economist, Steve Nickell, to Britain's Monetary Policy Committee.

Howard J. Wall is a research officer at the Federal Reserve Bank of St. Louis. Gylfi Zoëga is a professor of economics at the University of Iceland; a reader of economics at Birkbeck College, University of London; and an academic adviser to the Central Bank of Iceland. The authors thank seminar participants at the Central Bank of Iceland, the Federal Reserve Bank of St. Louis, and the Institute for Monetary and Economic Studies of the Bank of Japan for their comments and suggestions. The views expressed are those of the authors and do not necessarily represent official positions of the Central Bank of Iceland, the Federal Reserve Bank of St. Louis, or the Federal Reserve System.

Federal Reserve Bank of St. Louis *Review*, January/February 2004, 86(1), pp. 23-31.
© 2004, The Federal Reserve Bank of St. Louis.

Figure 1

U.S. Unemployment and Its Coefficient of Variation Across States



within Europe, where unemployment increasingly appears in regional clusters that cross national borders. Crone (1998/1999) groups the U.S. states into regions based on common cyclical behavior, while Carlino and Sill (2001) find considerable state differences in the volatility of regional cycles. Owyang, Piger, and Wall (2003) identify distinct state-level recession/expansion phases, finding a great deal of business cycle discord among the states and between individual states and the country as a whole. They also find significant cross-state differences in the depths of recessions and the speed of expansions. Recent research has also found that states and regions respond differently to monetary policy (Carlino and DeFina, 1998; Fratantoni and Schuh, 2003; and Owyang and Wall, 2003).

It follows from these studies that the national economy of the United States is a composite of significantly diverse but interrelated regional economies. In this paper, we show how the diversity in regional labor market conditions can be used to enrich policymakers' understanding of the aggregate economy. In the next two sections, we briefly lay out a state-level view of recent U.S. labor market trends and then describe how differences in regional business cycles can lead to changes in aggregate wage inflation. In the final section, we test for the underlying conditions for this to occur and demonstrate how region-level data can be used to estimate

the aggregate natural rate of unemployment in the United States.

A STATE-LEVEL VIEW OF U.S. UNEMPLOYMENT

This article relies on two suppositions about the dispersion of regional labor market conditions: (i) that the dispersion is related to aggregate labor market conditions and (ii) that the dispersion changes over time. Both suppositions are supported by Figure 1, which illustrates that the movements in the aggregate unemployment rate over the past 25 years have largely been in synch with changes in the dispersion of state unemployment rates (as measured by the cross-state coefficient of variation). Correspondingly, the 1990s saw steadily declining aggregate unemployment alongside a convergence of state unemployment rates. The only period during which aggregate unemployment was out of synch with the coefficient of variation was in 1986-87, when a handful of states had sudden increases in unemployment following the crash of energy prices in 1986.² Along with the country as a whole, all other states saw falling unemployment during this period.

² These states were Alaska, Alabama, Colorado, Louisiana, Mississippi, Texas, and Wyoming.

Figure 2

Changes in State Unemployment Rates Around Recessions

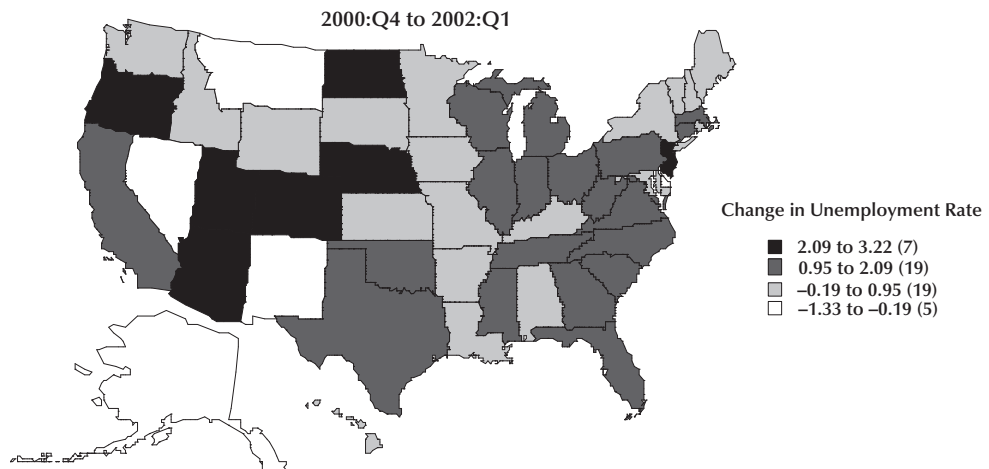
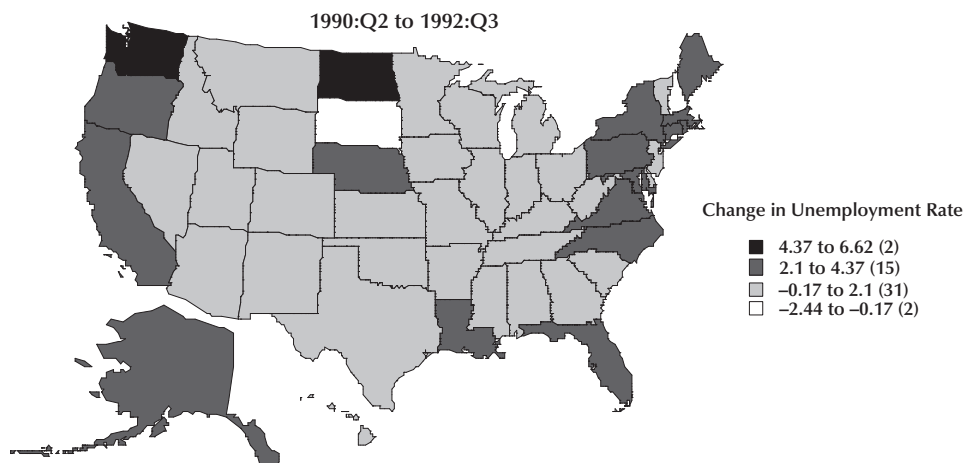
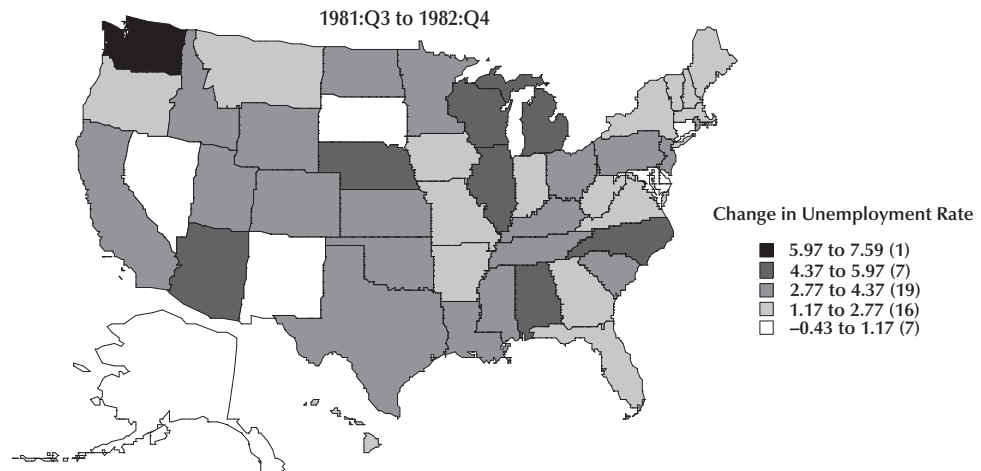


Figure 2 provides more evidence of the potential importance of regional labor market variation by showing the distribution of changes in state unemployment surrounding the three most recent recession episodes.³ Associated with the 1981-82 recession, the U.S. unemployment rate rose by about 3.3 percentage points from the third quarter of 1981 to the fourth quarter of 1982. Over the same period, 29 states saw smaller increases in their unemployment rates, 14 of which saw increases that were less than half as large (Nevada actually saw a small decrease). On the other hand, of the 21 states whose unemployment rates rose relatively more than the national average, 6 states saw a rise of at least 4.8 percentage points.

The period surrounding the 1990-91 recession is perhaps the most regionally distinct of the three most recent recessions. The aggregate unemployment rate rose by 2.3 percentage points from the second quarter of 1990 to the third quarter of 1992. The brunt of the increase was felt on the coasts, where most states saw increases in their unemployment rates that were much larger than average—particularly the large states of California, New York, North Carolina, and Washington. In contrast, a significant majority of states (36), mostly located in the vast middle of the country, saw a milder than average increase in unemployment. In fact, four states actually saw their unemployment rates fall during the period.

Associated with the 2001 recession was a run-up in unemployment that began in the fourth quarter of 2000 and continued well after the end of the recession. By the first quarter of 2002, the fact of a regionally diverse unemployment experience, and an increasing coefficient of variation, had become clear. By that time, the aggregate unemployment rate had risen by 1.6 percentage points, although 35 states saw smaller increases than this, and 6 saw declines. The states hit most severely were scattered across the country, with pockets in the Great Lakes region, the Atlantic seaboard, the western Plains, and the Southwest.

HOW REGIONAL BUSINESS CYCLES MIGHT MATTER

Here, we describe how a nonlinear relationship between inflation and measures of labor market

pressures—such as vacancies, unemployment, and employment growth—would mean that differences in regional business cycles are able to affect measures of aggregate conditions. Such nonlinearities are standard in the theory of unemployment, and it is not difficult to find empirical evidence backing them up.

Numerous statistical studies of the distribution of wage changes point to a potential role for asymmetric wage adjustments and heterogeneity (see, for example, McLaughlin, 1999; and Card and Hyslop, 1997). These studies show that the distribution of wage changes is skewed away from small increases and absolute cuts and toward large increases. There is a thinning of the left-hand tail to the left of the zero-inflation point, thereby indicating nominal wage rigidity. As McLaughlin (1999) documents, the skewness of the distribution exists even in the absence of any nominal wage rigidity: Even if the distribution is truncated at zero wage increases, the distribution is still skewed. According to survey results from Truman Bewley (1999), managers are hesitant to cut wages because of considerations about worker morale. Wage cuts are likely to introduce personnel and incentive problems beyond the intended effect on turnover. It follows that in an economy where some sectors and/or regions are declining and others are expanding, the relative wage cuts occurring in the former are smaller than the wage increases offered in the latter.

This microeconomic evidence suggests that regional labor market disaggregation may have a role in illuminating aggregate labor market outcomes. This can be illustrated most simply with the textbook version of the Phillips curve that traces its origins to Phelps (1968). In this model, wage inflation persists because firms cannot adjust instantaneously to changes in vacancies. This might be due to the costs of setting wages or because wage setting is staggered over time. Thus, for a given unemployment rate, the rate of wage inflation is an increasing function of the number of vacancies that firms would like to fill and of inflation expectations. There is a critical vacancy rate, \bar{v} , at which actual wage inflation equals expected wage inflation. When the vacancy rate is above \bar{v} , there is unexpected wage inflation. Conversely, when the vacancy rate is below \bar{v} , there is unexpected wage deflation.

The microeconomic evidence we cite above suggests that the slope of the relationship between wage inflation and the vacancy rate differs above and below \bar{v} . This is because firms are more reluctant

³ As determined by the NBER, the dates for these recessions are July 1981 through November 1982, July 1990 through March 1991, and March 2001 through November 2001.

to cut expected wages than to raise them. So, starting from \bar{v} , a decrease in the vacancy rate will lead to wage deflation that is smaller in absolute terms than the wage inflation that would follow an equivalent increase in the vacancy rate. In other words, the relationship between wage inflation and the vacancy rate is convex because it is flatter for vacancy rates below \bar{v} .

To see how this convexity matters, consider an economy with two equal-sized regions, both with vacancy rates of \bar{v} . Now consider equal but opposite-signed changes in the regions' vacancy rates (i.e., the changes are mean-preserving). One region experiences unexpected wage inflation that is greater in absolute terms than the unexpected wage deflation in the other. Thus, a mean-preserving increase in the dispersion of regional vacancy rates is associated with higher average wage inflation. More generally, with a strictly convex relationship between wage inflation and the vacancy rate, the larger is the dispersion of regional vacancy rates, the higher is the aggregate wage inflation for any given aggregate vacancy rate.

CONVEXITY AND THE NATURAL RATE IN THE UNITED STATES

The discussion above describes how aggregate wage inflation can be affected by the dispersion of regional labor market conditions when the region-level relationship between wage inflation and labor market conditions is convex. To test for this convexity, we use state unemployment rates and rates of growth of employment as our measures of state labor market conditions. Unfortunately, there are no state-level data for vacancies. This gives rise to the following equation, which we estimate with state-level panel data:

$$(1) \quad \frac{\dot{w}_{it}}{w_{it}} = \alpha_0^i + \alpha_1 \frac{\dot{N}_{it}}{N_{it}} + \alpha_1 \left(\frac{\dot{N}_{it}}{N_{it}} \right)^2 + \alpha_2 u_{it} + \alpha_2 u_{it}^2 + \alpha_3 \frac{\dot{w}_t^e}{w_t} + \varepsilon_{it}.$$

In (1), i refers to the state, t refers to the time period, \dot{w}_{it}/w_{it} is wage inflation, α_0^i is a state fixed effect, \dot{N}_{it}/N_{it} is employment growth, u_{it} is the unemployment rate, and \dot{w}_t^e/w_t is expected aggregate wage inflation. We use quarterly data from 1977:Q3 to 2002:Q1. Our wage measure is hourly earnings in manufacturing, employment data are from the establishment survey, and the unemployment rate

is from the household survey. Expected wage inflation at the aggregate level is measured by actual consumer price inflation (CPI) lagged one quarter. We estimate (1) with feasible generalized least squares (FGLS) to correct for state-specific autocorrelation and heteroskedasticity that is correlated across states.⁴

As reported in Table 1 and illustrated by Figures 3 and 4, the coefficients for employment growth and the unemployment rate (in levels and squared) imply a convex relationship between wage inflation and regional labor market conditions. However, the coefficient on the squared employment term is not statistically significant at traditional levels, so the relationship is not statistically different from linearity. On the other hand, the convexity of the relationship between wage inflation and the unemployment rate is statistically significant.

The weight of this empirical evidence indicates that the relationship between labor market conditions and wage inflation is convex, meaning that changes in the dispersion of conditions across states will have repercussions at the aggregate level. In particular, divergent regional business cycles cause measured wage inflation to rise for a given aggregate unemployment rate. In other words, the aggregate unemployment rate at which wage inflation is unchanged will be higher. These results suggest one possible reason for the non-inflationary boom that took place in the United States in the 1990s. Recall Figure 1, which shows that the coefficient of variation of state unemployment rates fell throughout the period, indicating a convergence of economic activity. Consistent with our discussion, this decreased dispersion was accompanied by a falling aggregate unemployment rate but no increase in wage inflation.

To explore this possibility further, we estimate a relatively simple Phillips curve for the United States, including features common to Phillips curve models⁵:

⁴ We are able to correct for this most-general form of heteroskedasticity because our time series is relatively long for a cross-state panel. A useful rule of thumb is that this is possible if there are twice as many time periods as cross-sectional units (Beck and Katz, 1995), which our panel just satisfies.

⁵ The variety of Phillips curve specifications is vast; Staiger, Stock, and Watson (2001) alone includes dozens of different Phillips curve specifications and estimates. As Phelps (1968) noted 35 years ago, and which is no less true today, "The numerous Phillips-curve studies of the past ten years have...[offered] countless independent variables in numerous combinations to explain wage movements. But it is difficult to choose among these econometric models, and rarely is there a clear rationale for the model used" (p. 678).

Figure 3

Wage Inflation and Employment Growth



Figure 4

Wage Inflation and the Unemployment Rate

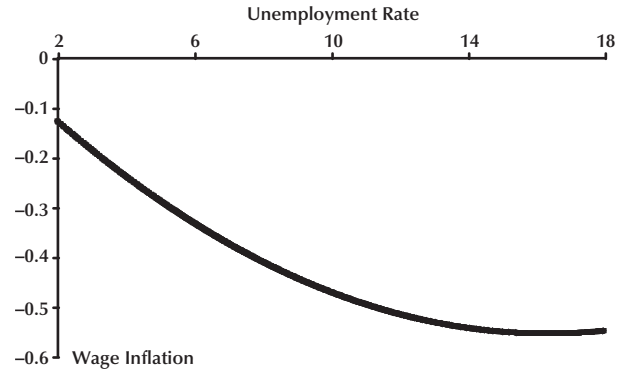


Table 1

Wage Inflation and Vacancies in a State Panel

	Coefficient	Standard error	t
Employment growth	0.0365*	0.0146	2.50
Employment growth squared	0.0047	0.0036	1.31
Unemployment rate	-0.0679*	0.0207	3.28
Unemployment rate squared	0.0021*	0.0012	1.66
Expected wage inflation	0.5907*	0.0293	20.14
State fixed effects (48)		Yes	
Observations		4,752	
Estimated covariances		1,176	
Estimated autocorrelations		48	
Log-likelihood		-4,587.42	

NOTE: * Indicates statistical significance at the 10 percent level. The estimator is FGLS and corrects for state-specific autocorrelation and heteroskedasticity with cross-state correlations. Quarterly state-level data, 1977:Q3–2002:Q1. Indiana and Kansas are excluded because of missing earnings data in early years of the sample. For space considerations, we do not report the estimates of the state fixed effects.

$$(2) \quad \frac{\dot{w}_t}{w_t} - \theta_t^e = \alpha_0 - \alpha_1 \ln u_t + \lambda X_t + \mathbf{B}\Phi + \pi_t^e + \varepsilon_t.$$

In (2), the dependent variable is nominal hourly wage growth averaged over years t and $t + 1$ net of expected productivity growth, θ_t^e , measured by the trend growth of output per worker in the nonfarm business sector. We also include a vector of demographic variables, Φ , to control for changes in the composition of the labor force (Phelps and Zoëga, 1997; Shimer, 1999; Francesconi et al., 2000; and

Staiger, Stock, and Watson, 2001). Following Staiger, Stock, and Watson (2001), these variables are the percentages of the adult population that are high school dropouts, college graduates, white, female, and aged 25–54. Expected wage inflation, π_t^e , is measured by average CPI inflation for years $t - 1$ and $t - 2$.

Our innovation is to include X_t , the coefficient of variation of state unemployment rates, which we expect to be positively related to wage inflation: Even if the aggregate unemployment rate is

Table 2

U.S. Phillips Curve Estimation

	Coefficient of variation and demographics	Demographics only
Constant	-112.326* (57.196)	-51.097 (68.513)
Log unemployment rate	-3.300* (0.741)	-3.036* (0.958)
Coefficient of variation of state unemployment rates	0.263* (0.080)	—
High school dropouts	0.957* (0.542)	1.098 (0.817)
College graduates	0.542* (0.218)	0.486* (0.242)
White	-0.044 (1.057)	-0.082 (1.389)
Female	1.645 (1.751)	0.798 (2.561)
Aged 25-54	0.240* (0.116)	-0.005 (0.119)
Expected wage inflation	0.452* (0.059)	0.326* (0.090)
Observations	19	19
Log-likelihood	7.781	1.566
R ²	0.944	0.8927

NOTE: White-corrected standard errors are in parentheses. * Indicates significance at the 10 percent level. Yearly aggregate data, 1982-2000.

unchanged, an increase in the dispersion of labor market conditions will raise the aggregate rate of wage inflation.

In choosing the time frame for estimating (2), we are hampered by the lack of state-level data before 1977 and demographic variables after 2000. In addition, to eliminate the estimation problems associated with the so-called monetarist experiment period, we include only 1982 and later. Despite these restrictions, we obtain the fairly reasonable results reported by Table 2.

Results for our more general specification—which includes demographic variables and the coefficient of variation of state unemployment rates—indicate that the education and age variables have all been important in determining the rate of wage inflation. More importantly for our present purposes, the results are consistent with our hypothesis that the regional dispersion of economic activity can affect aggregate wage inflation: The coefficient on the coefficient of variation of state unemployment is positive and statistically significant.

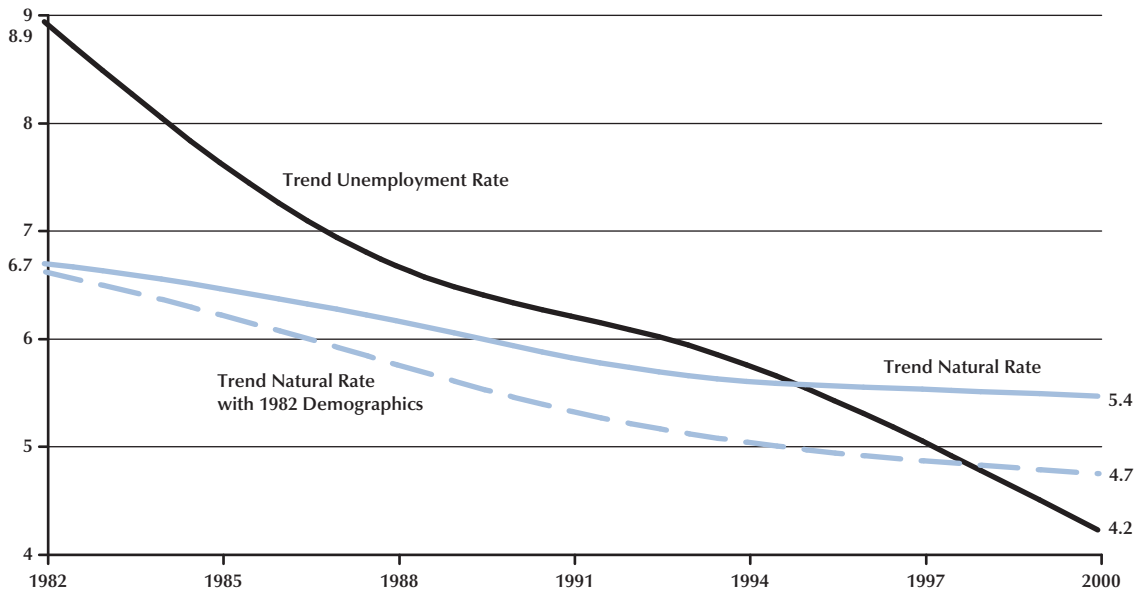
Table 2 also reports the results when the aggregate Phillips curve is estimated under the restriction that the coefficient of variation of state unemployment does not matter statistically. From these results it is clear that this restriction is not supported. When

the coefficient of variation is excluded, the coefficient on only one of the demographic variables—the share of college graduates—is anywhere close to being statistically significant. In addition, the constant term becomes smaller and statistically insignificant, making it very difficult to use the results to calculate a natural rate of unemployment. In sum, as supported by a likelihood-ratio test rejecting the null hypothesis that the restriction does not have a statistically significant effect, the estimates with the coefficient of variation are preferred.

According to Ball and Mankiw (2002), the primary source of the changes in the natural rate of unemployment in the 1990s was the acceleration of productivity growth (see also Pissarides, 2000; and Hoon and Phelps, 1997). An additional factor was the changing composition of the labor force (Phelps and Zoëga, 1997; Shimer, 1999; and Francesconi et al., 2000). Our Phillips curve estimation indicates that the convergence of state labor market conditions also had a role. The extent of this role can be obtained by examining the natural rates of unemployment implied by our Phillips curve estimation. Specifically, solving equation (2) by assuming that expected wage inflation is equal to last year's wage inflation, it can be rewritten as

Figure 5

The Falling U.S. Natural Rate, 1982-2000



$$(3) \quad \Delta \left(\frac{\dot{w}_t}{w_t} \right) = -\alpha_1 (\log(u_t) - \log(u_t^n)) + \varepsilon_{it},$$

where $u_t^n = \exp[(\alpha_0 + \lambda X + \mathbf{B}\Phi)/\alpha_1]$ is the time-variant natural rate of unemployment.⁶

The trend natural rate from our estimation and the actual trend unemployment rate are illustrated in Figure 5. According to our results, the natural rate fell steadily between 1982 and 2000, from 6.7 percent to 5.4 percent. Although relatively large, because the period's demographic changes worked to increase the natural rate, this 1.3-percentage-point drop understates the importance of changes in the dispersion of state-level unemployment rates. To remove the effect of these changes, the dashed blue line in Figure 5 is what the trend natural rate of unemployment would have been if the demographic variables had remained fixed at their 1982 levels.⁷ As the figure indicates, if all else had remained constant, changes in the dispersion of state unem-

ployment rates would have lowered the trend natural rate to 4.7 percent by 2000.

CONCLUSIONS

Using state-level data, we find that there is a convex relationship between unexpected wage inflation and labor market conditions—as measured by the unemployment rate and employment growth. This convexity suggests that increases in the cross-state dispersion of unemployment rates and employment growth mean a higher level of aggregate wage inflation even if aggregate unemployment and employment growth are unchanged. Finally, we include the coefficient of variation of state unemployment rates in our estimation of an aggregate Phillips curve. From this, we find that the convergence of state labor market performance between 1982 and 2000 was responsible for a 2-percentage-point drop in the natural rate of aggregate unemployment.

REFERENCES

- Abraham, Katharine G. and Katz, Lawrence F. "Cyclical Unemployment: Sectoral Shifts or Aggregate Disturbances?" *Journal of Political Economy*, June 1986, 94(3, Part 1), pp. 507-22.

⁶ See Staiger, Stock, and Watson (1997) for an analysis of the precision of natural rate estimates.

⁷ Note that because our dependent variable in (2) is wage inflation net of productivity growth, the trend natural rates shown in Figure 5 are also net of the effect of productivity changes.

- Ball, Laurence and Mankiw, N. Gregory. "The NAIRU in Theory and Practice." *Journal of Economic Perspectives*, Fall 2002, 16(4), pp. 115-36.
- Beck, Nathaniel and Katz, Jonathan N. "What To Do (and Not To Do) with Time-Series Cross-Section Data." *American Political Science Review*, September 1995, 89(3), pp. 634-47.
- Bewley, Truman F. *Why Wages Don't Fall During a Recession*. Cambridge, MA: Harvard University Press, 1999.
- Brainard, S. Lael and Cutler, David M. "Sectoral Shifts and Cyclical Unemployment Reconsidered." *Quarterly Journal of Economics*, February 1993, 108(1), pp. 219-43.
- Card, David and Hyslop, Dean. "Does Inflation 'Grease the Wheels of the Labor Market?'" in Christina D. Romer and David H. Romer, eds., *Reducing Inflation: Motivation and Strategy*. Chicago: University of Chicago Press, 1997.
- Carlino, Gerald and DeFina, Robert. "The Differential Regional Effects of Monetary Policy." *Review of Economics and Statistics*, November 1998, 80(4), pp. 572-87.
- Carlino, Gerald and Sill, Keith. "Regional Income Fluctuations: Common Trends and Common Cycles." *Review of Economics and Statistics*, August 2001, 83(3), pp. 446-56.
- Crone, Theodore M. "Using State Indexes to Define Economic Regions in the US." *Journal of Economic and Social Measurement*, 1998/1999, 25(3/4), pp. 259-75.
- Francesconi, Marco; Orszag, J. Michael; Phelps, Edmund S. and Zoëga, Gylfi. "Education and the Natural Rate of Unemployment." *Oxford Economic Papers*, January 2000, 52(1), pp. 204-23.
- Fratantoni, Michael and Schuh, Scott. "Monetary Policy, Housing, and Heterogeneous Regional Markets." *Journal of Money, Credit, and Banking*, August 2003, 35(4), pp. 557-89.
- Hoon, Hian Teck and Phelps, Edmund S. "Growth, Wealth and the Natural Rate: Is Europe's Jobs Crisis a Growth Crisis?" *European Economic Review*, April 1997, 41(3-5), pp. 549-57.
- Layard, Richard; Nickell, Stephen and Jackman, Richard. *Unemployment*. Oxford: Oxford University Press, 1991.
- Lilien, David M. "Sectoral Shifts and Cyclical Unemployment." *Journal of Political Economy*, August 1982, 90(4), pp. 777-93.
- McLaughlin, Kenneth J. "Are Nominal Wage Changes Skewed Away From Wage Cuts?" *Federal Reserve Bank of St. Louis Review*, May/June 1999, 81(3), pp. 117-32.
- Owyang, Michael T.; Piger, Jeremy and Wall, Howard J. "Business Cycle Phases in U.S. States." Working Paper 2003-011, Federal Reserve Bank of St. Louis, October 2003.
- Owyang, Michael T. and Wall, Howard J. "Regional Disparities in the Transmission of Monetary Policy." Working Paper 2003-008, Federal Reserve Bank of St. Louis, April 2003.
- Overman, Henry G. and Puga, Diego. "Unemployment Clusters Across Europe's Regions and Countries." *Economic Policy*, April 2002, 34, pp. 115-47.
- Phelps, Edmund S. "Money-Wage Dynamics and Labor-Market Equilibrium." *Journal of Political Economy*, July-August 1968, 76(4, Part 2), pp. 678-711.
- Phelps, Edmund S. *Structural Slumps*. Cambridge, MA: Harvard University Press, 1994.
- Phelps, Edmund S. and Zoëga, Gylfi. "The Rise and Downward Trend of the Natural Rate." *American Economic Review*, May 1997, 87(2), pp. 283-89.
- Pissarides, Christopher A. *Equilibrium Unemployment Theory*. Cambridge, MA: MIT Press, 2000.
- Shimer, Robert. "Why Is the U.S. Unemployment Rate So Much Lower?" in Benjamin S. Bernanke and Julio Rotemberg, eds., *NBER Macroeconomics Annual 1998*. Volume 13. Cambridge, MA: MIT Press, 1999, pp. 11-61.
- Staiger, Douglas; Stock, James H. and Watson, Mark W. "How Precise Are Estimates of the Natural Rate of Unemployment?" in Christina D. Romer and David H. Romer, eds., *Reducing Inflation: Motivation and Strategy*. Chicago: University of Chicago Press, 1997.
- Staiger, Douglas; Stock, James H. and Watson, Mark W. "Prices, Wages, and the U.S. NAIRU in the 1990s," in Alan B. Krueger and Robert M. Solow, eds., *The Roaring Nineties: Can Full Employment Be Sustained?* New York: Russell Sage Foundation, 2001, pp. 3-60.

Subjective Probabilities: Psychological Theories and Economic Applications

Abbigail J. Chiodo, Massimo Guidolin, Michael T. Owyang, and Makoto Shimoji

Conventional economic analysis of individual behavior begins with the assumption that consumers maximize expected utility, optimizing their planning for the future. Economists incorporate this assumption in models by endowing consumers in those models with the skills of a good statistician—that is, the ability to make rational (and often complicated) calculations. While not always realistic (perhaps never), this assumption facilitates the use of economic models that may work well in the real world. However, in some cases, these models cannot explain some of the evidence uncovered in psychological experiments. In other words, the traditional statistics-based approach sometimes fails to predict individual behavior and aggregate market outcomes that are consistent with the empirical evidence. For instance, observed stock prices and portfolio choices fail to conform to the implications of well-known frameworks, such as the capital asset pricing model (CAPM). Such cases have encouraged a branch of economics that borrows ideas from psychology to explain these discrepancies.¹

In this area of study, researchers replace the assumption that individuals use complicated statistical formulas to maximize expected utility with the likelihood that they use simple rules of thumb instead, rules that have been identified by psychological research. Psychologists have found evidence that individuals estimate the probability of future outcomes in a nonstatistical, or subjective, manner. Kahneman and Tversky (1973) and Kahneman, Slovic, and Tversky (1982), among others, have introduced the idea of subjective probability heuristics—rules that people tend to rely on when assessing the likelihood of alternative events. Psychological

research has shown that the use of these rules can create different outcomes from what statisticians (and economists) might expect, both in the estimated probabilities and in observed behavioral patterns.

Behavioral theories of decisionmaking therefore ask whether economic phenomena may be explained by models in which

- *Some*, but not necessarily all, agents either fail to update their probabilistic beliefs by applying the appropriate statistical rules or subsequently fail to maximize a standard expected utility objective.
- The remaining fully rational agents, then, cannot completely exploit and eliminate the biases caused by the actions of agents who are not perfectly rational.

While these heuristics are drawn from psychological studies, they may be supported by economic models with *boundedly rational* agents (Simon, 1955). In other words, agents do not always have the time or the cognitive ability to process all of the data provided by the economic environment with the necessary accuracy. Instead, people might employ these heuristics to arrive at analyses that are less costly to calculate than optimal decisions (Evans and Ramey, 1992); and, often, the optimal decisions themselves are impossible to calculate for difficult problems. Thus, boundedly rational agents do not maximize expected utility as an economist would generally assume. Instead, they maximize *perceived* expected utility, a quantity based not on actual probabilities but on their *beliefs* about those probabilities (Rabin, 1998, 2002).

In this article, we focus on the nature and application of psychological rules for probability formation and the biases from anticipated economic

¹ This vein of research is, in some part, attributed to the cross-disciplinary work of Amos Tversky and 2002 Nobel laureate Daniel Kahneman.

Abbigail J. Chiodo was a research associate at the Federal Reserve Bank of St. Louis while this article was written. Massimo Guidolin is an assistant professor of economics at the University of Virginia. Michael T. Owyang is an economist at the Federal Reserve Bank of St. Louis. Makoto Shimoji is an associate professor, International Graduate School of Social Sciences, Yokohama National University. Kristie M. Engemann and Mark L. Opitz provided research assistance.

outcomes that can result from their use.² We examine three heuristics that have been identified by psychologists: the representativeness heuristic (RH), the availability heuristic (AH), and anchoring and adjustment (AA). We review the psychological evidence supporting the common use of these heuristics in estimating subjective probabilities. Finally, we consider a financial application that uses heuristics to estimate probabilities with potentially important economic implications. We then show the effect of these heuristics on people's probability judgments.

PSYCHOLOGICAL EVIDENCE

Economics has a long history of exploring human behavior in decisionmaking. Economic models often require agents to form expectations under uncertainty, e.g., expected inflation in macroeconomic models, expected returns in financial models, or expected utility in decision/choice models. However, when faced with calculating expectations, economists often assume that the probabilities are known or can be inferred (rationally) through learning. What is meant by this? An economic agent might maximize his expected utility over n uncertain outcomes, defined as

$$(1) \quad EU = \sum_{i=1}^n p_i U_i,$$

where p_i is the probability of outcome i and U_i is the utility from outcome i .

Psychologists, however, have found that people neglect some available information in their decision-making process—that is, they do not update probabilities as new information arrives, as an agent adhering to rational expectations would. Consistent with Rabin's idea of perceived expected utility, agents might maximize

$$(2) \quad EU^p = \sum_{i=1}^n \hat{p}_i U_i,$$

where \hat{p}_i is the subjective probability of outcome i . The difference between equations (1) and (2) is solely in the agent's assessment of the likelihood that i will be realized. In this section, we explore how economists and psychologists view p_i and \hat{p}_i differently.

² Sherman and Carty (1984) and Camerer (1995) provide surveys of the psychological evidence on the heuristics discussed here. Another strand of the recent behavioral literature focuses on the effects of nonexpected utility preferences for optimal decisions. We do not discuss these contributions and concentrate instead on the process of belief formation.

Representativeness

Tversky and Kahneman (1974) suggest that people typically rely on the representativeness heuristic (RH) when answering "probabilistic questions" such as "What is the probability that event A originates from process B?" That is, the RH is used when a person must determine such probabilities based on the degree to which A resembles B.

RH is used when an agent must update a subjective probability with new information. Economists sometimes assume that agents employ Bayes's law when updating probabilities. Bayes's law defines the probability of an event X , conditional on observing A , as

$$(3) \quad p(X|A) = \frac{p(A|X)p(X)}{p(A)},$$

where $p(A|X)$ is the conditional probability of A given X and $p(X)$ and $p(A)$ are population parameters typically referred to as base rates.

While Bayes's law is a useful statistical rule, psychologists have found that people tend to act in a decidedly non-Bayesian fashion and have identified a number of subjective probability biases grouped under the umbrella of RH.

Tversky and Kahneman (1974) and Kahneman, Slovic, and Tversky (1982) note that using RH when determining probability can lead to insensitivity to prior probability, or base-rate frequency, of the outcomes.³ In one example, subjects were asked to identify a described individual as either a lawyer or an engineer. The subjects were given descriptions that included phrases such as "he wears glasses" or "he wears a pocket protector." Subjects were first told that the individual in question was drawn from a random sample composed of 100 people, 70 of which were engineers and 30 of which were lawyers. Then, the base rates were reversed. The subjects were told that, of the 100 people in the sample, 70 were lawyers and 30 were engineers. Kahneman, Slovic, and Tversky found that the subjects' probability judgments did not differ when the base rate was changed, even though Bayes's law indicates that the conditional probabilities cannot be equal if the base rates change.

Grether (1980, 1992) and El-Gamal and Grether (1995) designed experiments that determine that RH "is a good descriptive model of behavior under

³ Sherman and Carty (1984) provide a comprehensive review of the biases that are attributed to RH.

uncertainty for untutored and unmotivated (or at least not financially motivated) individuals” (Grether, 1980, p. 538). Specifically, they show that subjects under-use base rate information when making subjective probability judgments for events that have little or no consequence or cost. Borgida and Brekke (1981) have also shown that, while most people do not neglect base rates entirely, they are typically under-used.

Availability

The availability heuristic (AH) describes a method by which a person determines the likelihood of an event according to the ease with which he or she can recall instances that match the event. That is, one’s experiences and conditioning affect how a person determines the likelihood that an event will occur. For example, one might estimate the risk of a burglary in a certain neighborhood by the number of burglaries one can recall (including any personal experience).

A similar method is the simulation heuristic, by which people will determine the likelihood of events based on the ease with which they can simulate (or imagine) the outcome in their minds. An example of this is a person who determines the probability that the value of a certain stock will decline based on the number of different scenarios he or she can easily imagine that would cause such an occurrence.

While AH can often be helpful in making decisions and estimates, Tversky and Kahneman (1974) list several biases that can result from AH. These include biases (i) due to the retrievability of instances (examples easily brought to mind are often judged to be more likely than they actually are), (ii) due to imaginability (easily imagined outcomes can give the illusion that they are more common), and (iii) due to illusory correlation (one event more strongly implies another if the two events frequently occur simultaneously).

Tversky and Kahneman (1973) outline several studies used to demonstrate the AH and its subsequent biases. For example, subjects were read lists of names of both sexes, some of which were names of very famous people (Richard Nixon and Elizabeth Taylor, for example). Afterward, some were asked to estimate if there were more males or females on the list. People tended to estimate, sometimes incorrectly, that there were more of whichever sex had more famous people in the list. The famous names were easier to remember and therefore more prominent in the minds of the subjects. Tversky

and Kahneman conclude, then, that people make estimates based on AH, which (in this case) led to the retrievability bias.

AH has been applied to marketing and advertising to investigate the effect of retrieval on the subjective assessment of product failure. Folkes (1988) presents four studies in which subjects were asked to predict how likely various products were to fail. Different scenarios and distinctive brand names were used to make some products or instances more memorable. Folkes found that these judgments were biased in ways described by the AH—that more memorable products (memorable for various reasons) influenced the subjects’ decisions. Rabin (1998) points out that people often give too much weight to memorable evidence, even when better sources of information are available. He notes that one may allow a dramatic personal story from a friend regarding an instance of product failure to be more influential than consumer reports with general statistics regarding that product.

Recently, Mullainathan (2002) developed a model of memory limitations based on two psychological concepts that can have properties similar to AH. The first concept, rehearsal, is the assumption that remembering an event, story, or some form of information one time makes it easier to remember again. Mullainathan points out that rehearsal is used by students who study for a test by reading over the material and then quizzing themselves to help them remember it. The second concept, associativeness, is the process by which current events can trigger memories of past events that have similar aspects. Thus, even uninformative information—information that does not change the likelihood of an event—can influence beliefs by changing perceptions of the past. Mullainathan suggests that people respond “too much” because news resurrects memories that reinforce beliefs.⁴

Anchoring and Adjustment

The final heuristic we address is anchoring and adjustment (AA), which Tversky and Kahneman (1974) define. According to this heuristic, individuals make estimates based on a starting point (the anchor) and update (adjust) their subjective probability based on new information. While this does not seem to differ from RH or even from Bayesian updating,

⁴ Mullainathan considers an application of this model to individuals’ consumption decisions. He suggests that individuals react more to their private information than to aggregate information because aggregate information is forgotten.

psychologists have shown that individuals have a propensity to bias their estimated probabilities toward the anchor. That is, individuals do not adjust *enough* to new information, making the value of the anchor more critical.

An individual's initial guess (the anchor) can be subjective (interpreted) or objective (e.g., taken from base rates). Often, the anchor depends on the manner in which the question is asked or how the information is given. For example, Tversky and Kahneman (1974) asked subjects to estimate the percentage of African countries in the United Nations by first giving them a number (determined randomly by spinning a wheel) and then asking the subjects whether that number was higher or lower than the percentage of African countries. Different initial values led to strikingly different estimates. While the median estimate was 25 percent for groups that received 10 as the starting value, the median estimate was 45 percent for those given a starting value of 65—illustrating the bias toward the anchor.

A starting value can also be the result of a subject's (usually incomplete) computation. Tversky and Kahneman (1974) give the example of two groups being asked to estimate 8! in a limited amount of time. While one group was given $1 * 2 * 3 * 4 * 5 * 6 * 7 * 8$ as the problem, the other group was given $8 * 7 * 6 * 5 * 4 * 3 * 2 * 1$. Note that the product of the first few steps of multiplication (performed left to right) of the descending sequence is much higher than that of the ascending sequence. As predicted, the median estimate of the group shown the descending sequence was much higher than the median estimate of the group shown the ascending sequence.

This example illustrates how subjects tend to focus on only part of a problem. For instance, the probability of success at any one stage of an event is often used as a starting point (an anchor) to determine the probability of overall success. However, their assessment of the probability for an event with multiple stages is often skewed because they do not deviate enough from that anchor. Tversky and Kahneman (1974) refer to research that shows how anchoring biases the estimation of probability for different types of events—specifically, that subjects overestimate *conjunctive* events and underestimate *disjunctive* events. For example, suppose there is a bag of marbles, half of which are red and half of which are black: People will overestimate the probability of, for instance, drawing a red marble from the bag seven times in succession after replacing the drawn marble (a conjunctive event); they will

underestimate the probability of drawing a red marble from the bag at least once in seven successive tries after replacing the drawn marble (a disjunctive event). The anchor for both events is the probability of drawing a red marble on any try. Success in a conjunctive event may be likely for only *one* of several required outcomes, yet subjects stick close to their anchor and thus overestimate the probability of overall success. Conversely, subjects tend to underestimate the likelihood of success beyond the anchor when multiple attempts are allowed to achieve merely one successful outcome.

ILLUSTRATIVE APPLICATIONS

To demonstrate the effect of heuristic biases on probability judgments, we offer the following illustrations.

Disaster Insurance

The biases resulting from these heuristics have some implications in the earthquake insurance market. A large earthquake in one area certainly qualifies as the kind of salient event mentioned in the discussion of the AH. After all, graphic pictures and information from the media or personal stories from friends affected by the earthquake are likely to be easily remembered when estimating one's own need for earthquake insurance. Psychology theory implies, then, that a large earthquake should cause people to overestimate the probability that they will need earthquake insurance, which could explain the "gains by losses" phenomenon: In the event of an earthquake, an insurance company must pay out on claims, incurring a loss; if an earthquake causes an increase in demand for insurance, however, insurance companies can benefit, overall, by experiencing significant gains during the period after the earthquake.

Consider an actuarially fair earthquake insurance policy with premium π and payout Y . By definition, the actuarially fair premium must be a function of the payout and the risk of the event being insured against. In this case, the premium should exactly offset the expected payouts.

Suppose now that, given the premium, a person must decide whether to purchase insurance based on the perceived likelihood of a loss. Irrelevant information, such as an earthquake in another part of the country, does not affect the probability of a loss. However, a person employing heuristics might assume a greater likelihood of a local earthquake—

making the insurance contract more attractive to her. Thus, if persons employed the heuristics, we would expect demand for insurance to rise after the occurrence of a similar event.

In fact, Kunreuther et al. (1978) finds that people tend to discount the likelihood of a disaster (e.g., a flood or an earthquake) until the event occurs. After people “update” their assessment, purchases of insurance contracts rise. Moreover, Shelor, Anderson, and Cross (1992) and Aiuppa, Carney, and Krueger (1993) found that insurers’ stock prices increased after the 1989 earthquake in San Francisco, due to an increased demand for coverage. However, Yamori and Kobayashi (2002) find no such benefit to insurance companies in Japan after the 1995 Hanshin-Awaji earthquake. Yamori and Kobayashi note that unique attributes of Japanese earthquake insurance may be the reason for the difference between the United States and Japan in stock market reactions to large earthquakes. Namely, the Japanese government sets the insurance industry premium levels at “no loss and no profit.” Interestingly, while studies have shown the positive link between earthquakes and insurance stock prices in the United States, other studies indicate no such relationship for hurricanes.⁵

Product Liability

A second application for the heuristics involves market attitude with regard to product reputation, specifically, shocks to reputation. We can model market behavior after a product failure as a temporary shift in demand that results in lower sales and falling retail and stock prices.

Airplane Crashes. News agencies report airplane crashes in detail, often exhaustively and over an extended period of time. These reports provide vivid images to the public.^{6,7} As a consequence, people may avoid air travel, at least for some amount of time. Without an AH, such tragedies would have little effect on people’s belief regarding air travel safety because these events are rather uncommon and it is widely known that air transportation has been much safer than any type of

ground transportation in the United States. Because the change in people’s beliefs regarding the overall safety of air travel would be minimal, this type of tragedy would be interpreted as idiosyncratic to a particular airline. In this case, it is therefore possible that other airlines (the rivals of the airline that experienced a crash) would benefit from such an event. If an AH does exist among the potential customers, however, the demand for air travel as a whole would decline. This externality would harm the market as a whole, and, as a consequence, other carriers would lose profits as well.

Borenstein and Zimmerman (1988) found that “an airline’s shareholders suffer a statistically significant wealth loss when the airline experiences a serious accident...[although] the average loss in equity value is much smaller than the total social costs of an accident” (p. 913). In addition, they found that (i) such accidents have little or no effect on demand and (ii) that there is little evidence of a externality effect (positive or negative) caused by such accidents on the demand for other airlines. This study suggests that the market barely reacts to these events.

A subsequent study by Mitchell and Maloney (1989) partitioned the sample into “at-fault” crashes (those caused by pilot error) and all other crashes and tested whether these two distinct groups receive different reactions from the market. Contrary to the study by Borenstein and Zimmerman (1988), they found a statistically significant *negative* reaction in the former group.⁸ However, these studies do not offer an insight regarding the effect of an AH.

Nethercutt and Pruitt (1997) reported a finding similar to that of Mitchell and Maloney (1989) by examining the ValuJet accident in 1996. In their study, they found two things: (i) not only the shareholders of ValuJet but also those of other “low cost” carriers suffered losses due to this accident and (ii) the shareholders of the major airlines indeed received statistically significant gains after this event. At first sight, this result seems to suggest the nonexistence of an AH. However, their study does not distinguish the switching effect from the spillover effect; hence, it does confirm that the former dominates the latter, but has nothing to say about the effect of an AH.

The study by Bosch, Eckard, and Singal (1998) partially answers the question raised above. The authors consider the market overlaps of airlines in

⁵ See Lamb (1995) and Cagle (1996) for details.

⁶ By “airplane crashes,” we do not mean the consequence of terrorism or hijacking, which are due to external forces.

⁷ In the previous case, we argued that some of the events are consistent with the presence of the AH. In this section, we also survey several examples of product detection, e.g., product recalls. Although most of the failures are also life-threatening, like the examples above, there is no evidence of spillover effects in the industry as a whole, unlike the examples above.

⁸ See also Broder (1990).

the context of a recent airplane crash and examine whether customers respond to a commercial airline crash by switching to rival airlines and/or flying less. They find that passengers who do choose to fly will travel on the airlines that have not had a recent crash, but even these airlines suffer a negative spillover—namely, fewer passengers overall. With market overlap, the coexistence of the switching effect and the spillover effect offset each other and we can observe only the net effect of these two together. However, with little overlap, the switching effect is limited and hence we can test if the spillover effect exists. Indeed, Bosch, Eckard, and Singal found negative spillover effects after airplane crashes, consistent with the existence of the AH.

Firm Bankruptcy. Lang and Stulz (1992) studied the effect of one firm's bankruptcy announcement on the other firms in the same industry. They listed two effects of such an announcement⁹:

Contagion Effect. A change in the value of competitors that cannot be attributed to wealth redistribution from the bankrupt firm. This may happen because investors think that firms with characteristics similar to those of the bankrupt firm are less profitable than expected.

Competitive Effect. A change in the value of competitors that can be attributed to wealth redistribution from the bankrupt firm. This may happen, for example, if investors think that the bankrupt firm is doing poorly because other firms are doing well.

As for the first effect, they found that “on average, the market value of a value-weighted portfolio of the common stock of the bankrupt firm's competitors decreases by 1% at the time of the bankruptcy announcement and the decline is statistically significant” (p. 46). They also reported that “the effect appears to be greater for highly leveraged industries” (p. 46). For the second effect, they found that “the value of competitors' equity actually increases by 2.2% in more concentrated industries with low leverage” (p. 47).

These types of effects may be due to other announcements or events such as defective products and recalls.¹⁰ In addition, even though the same types of effects are observed, they may arise from other sources. In the following, we discuss such

⁹ Note that contagion affects all the relevant firms in the same negative direction, whereas the competitive effect does not.

¹⁰ For example, Ford Motor Company experienced a decline in sales after the Firestone tires used on Ford products were declared faulty by the media.

possibilities, as well as the possibility that some of the events may be attributed to the existence of heuristics we study in this paper.

A FINANCIAL APPLICATION

We now consider an application of heuristic probability judgments in an asset pricing model.¹¹ A formal description can be found in the appendix. Recently, Barberis and Thaler (2002) have stressed how behavioral approaches that focus on the mechanism of expectation formation cannot be applied to explanations of well-known aggregate puzzles in finance, such as the equity premium, excess volatility, and predictability issues. Although it is acknowledged that many models developed to investigate the cross-section of asset returns may often be used to also explain aggregate puzzles, much remains to be achieved by this strand of literature.¹² In this section we discuss a number of asset pricing puzzles that can be explained by subjective probability biases.¹³

How Do Subjective Probabilities Affect Asset Prices?

Assume there are two assets: a single-period, risk-free discount bond in zero net supply and a publicly traded stock (or stock index) in exogenous, unit supply. The stock pays out an infinite stream of perishable, real dividends, the growth rates of which randomly switch between two values: d_h in the good state (an expansion) and d_l in the bad state (a recession).

Individuals use both informative and uninformative variables to determine probability estimates. For simplicity, assume that the only informative variable is dividends. Dividends are informative because they directly relate to the payouts produced by the stock. Therefore, the information set is com-

¹¹ Barberis and Thaler (2002) and Hirshleifer (2001) are recent survey papers on the field of behavioral finance. See Mullainathan (1998) for an application of the AH to finance.

¹² Several papers have used nonexpected utility preferences consistent with the psychological and experimental evidence to approach the same phenomena (for instance, Barberis, Huang, and Santos, 2001, and Benartzi and Thaler, 1995). These papers are often considered to belong to the behavioral camp.

¹³ It goes without saying that a vast literature has developed over the past two decades that approaches the same puzzles we discuss; these have used types of frictions (transaction costs, information asymmetries and incomplete information, nonstandard preferences, etc.) that do not involve either the process of expectation formation or the ability of investors to rationally use the available information. The surveys in Campbell (2000) and Cochrane (2001) offer highly readable accounts and references.

posed of the sequence of realized high and low dividend growth rates plus a set that collects all the relevant realizations of the uninformative variables. Examples of uninformative variables are past stock prices because stock prices fail to add any predictive power for future cash flows produced by the stock currently owned. Alternatively, investors might use past levels of the price-dividend ratio to forecast future dividend growth because this ratio has been found to successfully predict stock prices in the empirical literature. In practice, stock market participants will *directly* care about the probability for dividends only. However, depending on the way subjective probabilities are formed, investors might *indirectly* also care about the joint probability distribution of dividends and the uninformative variables, in the sense that they might use uninformative events to predict dividends.

The Heuristics-Based Solution

Suppose that the probability of an increase in dividends is unknown and must be subjectively calculated based on past observations.¹⁴ A representative agent believes that the value of the stock depends not only on past dividend payments but also on the irrelevant information she has in her information set.¹⁵ Since events that are recent are more likely to be remembered, the further back in time an observation on the dividend growth rate is, the more unlikely it is that it will belong to the recalled information set. However, agents recall events that bear a high resemblance to current events, even when the similarity is defined not only in terms of dividends, but also in terms of other, irrelevant variables (e.g., past asset prices). For instance, an investor is more likely to recall a big drop in a company's dividend when it is associated

with a deep international crisis, even though the political variables need not carry information useful in predicting future economic conditions or the profitability of the company.

The appendix shows that, under these assumptions, the heuristic-based equilibrium stock price differs from the full-information equilibrium price because it stops being a fixed multiple of dividends; on the contrary, the heuristic-based equilibrium price-dividend ratio now contains a time-varying component, fitting the empirical finding that price-dividend ratios are subject to long swings. The variation in the price-dividend ratio derives from changes in the investors' memory-based (or subjective) expectation of dividends. For instance, particularly bad dividends may depress stock prices to the point that investors start recollecting previous bad times when the observed mean dividend growth was low. This happens because, when the AH is present, low growth tends to make other episodes of low growth salient and to bring up memories of other recessions; this happens also because poor dividend growth depresses stock prices and makes other recessionary episodes more memorable. These biases depress the subjective dividend expectations and cause deeper persistent declines in stock prices.

Excess Volatility of Stock Prices

We now investigate a few qualitative implications for stock prices. Since Shiller (1981) and LeRoy and Porter (1981), researchers have observed that stock prices tend to be much more volatile than the underlying economic fundamentals (dividends or aggregate consumption) would dictate. Recent research has examined this issue with mixed success (see Brennan and Xia, 2001, Bullard and Duffy, 2001, and Timmermann, 2001). Under full-information rational expectations, this finding represents a puzzle.¹⁶ The heuristics-based approach illustrates how the excess stock volatility puzzle can be easily resolved when the price-dividend ratio is time-varying as a result of limited memory and of avail-

¹⁴ Barsky and DeLong (1993) present a discounted model in which investors form extrapolative expectations and generate excess volatility of stock prices. However, their paper does not impose much structure on belief formation and fails to link the extrapolation process to the experimental psychology literature.

¹⁵ For a heuristic rule to have an effect on equilibrium outcomes, irrational traders need not be completely weeded out of the economic system (through bankruptcy or reduction to a marginal role in determining equilibrium results). Although the debate in the economics literature is not settled yet, important papers in finance (DeLong et al., 1990a,b, 1991) shed light on the subject: namely, that the price biases created by simple heuristics (such as the case with random trading of securities) create situations in which the exploitation of less-rational investors is risky and therefore fails to be implemented on the scale necessary to completely annihilate the effects of the biases (Shleifer and Vishny, 1997). Therefore, heuristics might appear in the aggregate and a representative agent is a useful shortcut to model such a situation.

¹⁶ Barberis and Thaler (2002) informally discuss a psychological model that could explain the excess volatility puzzle: Investors perceive a disproportionate volatility of the dividend growth rate when they are exuberant, i.e., when they observe dividend increases that convince them, too quickly, that mean dividend growth has increased. Although Barberis and Thaler notice that a similar story may be derived as an application of the RH, they do not present a formal model that maps heuristics into beliefs. Shiller (2003) has recently used the excess volatility puzzle as a workhorse to introduce behavioral finance research to overcome the traditional efficient market hypothesis.

ability, representativeness, and anchoring biases. The appendix provides a formal treatment.

Since high dividends are generally accompanied by high stock prices and low dividends by low stock prices, a high growth realization will make past high dividend growth rates more memorable because of the AH. This is because, if a high dividend growth rate causes an increase in stock prices, other episodes of bull markets and good fundamentals will be recalled. Such an event is likely to increase the subjective expectation of dividends and the price-dividend ratio. A similar reasoning applies to situations of low fundamentals and stock prices, i.e., they will generally make “bad times” more memorable and depress the expectation of future dividends. Therefore we expect positive covariation between dividend growth and the price-dividend ratio, which makes stock prices much more volatile than what is implied by full-information rational expectations. In this sense, heuristics-based asset pricing makes the solution of the volatility puzzle not only possible, but likely.

Bubbles and Crashes

A related topic is the tendency of stock markets to experience long periods of sustained (but hardly rational) increases in prices, followed by quick outbursts that often lead to sudden crashes. With reference to these phenomena, economists have developed both a literature on the theoretical conditions under which price bubbles may form and thrive (see Tirole, 1985) and a more recent empirical literature that describes markets as going through a sequence of “bulls” and “bears” (see Perez-Quiros and Timmermann, 2000). Unfortunately, the former mostly stresses the delicacy of bubbles, while the latter falls short of providing answers to our questions because it focuses on the microfoundations of bulls and bears. When investors use heuristics, bubbles and crashes occur in equilibrium more frequently.

Suppose the current period is characterized by good economic fundamentals and hence positive stock returns. In particular, some degree of exogenous optimism may easily project good dividend growth in high stock returns. At this point the following mechanism is triggered: A high current stock price elicits memories of previous periods of fast economic growth and “good” fundamentals. When past stock prices are also used to calculate expectations of future dividend growth, high current prices will also make past bull market periods more mem-

orable. Hence, past high-dividend periods will be assigned an abnormally high probability and will end up being over-represented in the recalled information set. As a result, expected dividends will be irrationally high. Unless the next-period dividend is particularly unfavorable, this sustains high demand for equities and stock prices: This is the beginning of the bubble. In such an environment it would be possible for stock prices to increase at such a pace that (given agents’ imperfect memory), in practice, only very recent bull periods would be recalled and used in forming expectations. Here, it is as if the market enters an entirely different world: Optimism dominates to the point where price increases are a foregone conclusion (c.f., the “New Economy”).¹⁷

The effect is further enhanced when anchoring is strong: If the run of price increases is sufficiently protracted, agents’ subjective perception of the probability of good economic fundamentals will become increasingly difficult to move.¹⁸ What ends a bubble? Potentially, a sufficiently negative realization of fundamentals growth. Such an epiphany could suddenly make investors recall past cases in which bull markets turned into bear markets. In other cases, it is sufficient that some variables, although irrelevant for pricing (political variables, for instance), may suddenly make investors aware that bad outcomes are possible. When this happens, the bubble bursts, often plunging into a catastrophic crash.

One phenomenon that has not been well explained by the theoretical literature on bubbles is the possibility of protracted periods of depressed stock prices, far below their most moderate rational levels—a sort of negative bubble (Weil, 1990). An advantage of heuristic-based asset pricing is the ability to generate episodes of irrationally low stock prices. Starting with poor underlying growth and some pessimism, markets may quickly plunge into spells in which investors focus only on past negative news and periods and, hence, systematically underestimate the mean dividend growth rate so that stock prices are too low given the quality of the underlying fundamentals. Strong anchoring may complete the picture, thus damping investors’ expectations for growth prospects.

¹⁷ See, for instance, *The Economist*, “Beyond the Business Cycle?” October 23, 1999, and the “new era” theories discussed in Shiller (2000).

¹⁸ Intuitively, anchoring makes bubbles harder to ignite but also harder to burst. Given the available empirical evidence, the behavior of financial markets is highly consistent with strong anchoring.

The Inflation–Stock Returns Puzzle

It is conventional wisdom to prefer nominal stock returns and inflation to be positively and highly correlated; rational markets, then, should price equities based on their discounted, expected nominal cash flow payments. Therefore, ruling out deeper macroeconomic effects (e.g., sectoral shifts and other distortions), an increase in current and expected inflation ought to increase expected nominal dividend payments and cause upward adjustment of observed stock prices. Empirical research in the past 20 years has found very limited support for the hypothesis that stock returns can protect shareholders from inflation. Normally, positive but moderate correlations have been found. In other words, the Fisher equation systematically fails for nominal stock returns.¹⁹ Heuristic-based asset pricing offers an easy way to rationalize such a phenomenon.

Suppose that inflation not only influences nominal dividend levels, but also acts as a variable in the set of uninformative information. In particular, assume that investors have convinced themselves that high inflation is always accompanied by subsequent increases in the level of real interest rates that depress economic growth. Interestingly, this conjecture does not need to be supported by the data, or it might be supported only by old data. Inflation is just an additional variable that becomes informative of future economic growth only because investors *think* it is. In this case, a high current inflation rate is essentially bad news: It makes past periods of poor growth and recession more memorable (via availability) and accelerates inflation (via representativeness). In practice, two effects take place at once: On one hand, inflation raises expected nominal dividends; on the other hand, inflation induces a pessimistic change in the agent’s recalled information set, lowering expected real dividends. The net effect is unclear but is consistent with the fact that nominal stock returns do not seem to react much to inflation news.²⁰

¹⁹ Equivalently, empirical studies have found negative correlation between real stock returns and inflation, both expected and unexpected. See, for instance, Nelson (1976) and Fama and Schwert (1977) for early evidence.

²⁰ Geske and Roll (1983) propose a theoretical model in which the only exogenous shocks are disturbances to the real level of economic activity. Since they assume monetary policy acts counter-cyclically, the negative correlation between real stock returns and inflation derives from the fact that bad news on fundamentals forecasts higher future monetary growth and inflation. However, the recent debate

CONCLUSION

In this article, we surveyed some of the research that has highlighted the crossover between economics and psychology. The assumptions economists have traditionally imposed in their models maintain that individuals are rational (and selfish) and construct their beliefs according to probability theory, following Bayes’s law. For most economic applications, this type of assumption fits well. However, there remain situations in which nonrational or quasirational behavior on the part of the median agent is observed. In these situations (e.g., hazard insurance and asset pricing), assuming that people behave rationally leads to puzzles—such as the inflation–stock returns puzzle, bubbles and crashes, and excess stock price volatility—that are yet unexplained using standard economic theories.

Economists have more recently begun to acknowledge irrationality as a source of interest for these economic applications. Accounting for all idiosyncratic effects is literally impossible and, moreover, undesirable. Economic theory adequately explains many types of behavior, including consumption behavior, for example. However, there remain some *systematic* deviations from rational behavior, which the *standard* models do not fully capture. The heuristics that psychologists suggest are examples of this. Incorporating these types of behavioral rules in our research could not only broaden how we approach and analyze subjects but also may greatly increase the power of our conclusions. We find, for example, that the puzzles in the asset pricing literature (such as those listed above) can be accounted for by adding a heuristic probability rule to the standard asset pricing framework. Thus, while behavior might not be a solution that is broadly cast, we propose that its importance, in some circumstances, may warrant further investigation.

REFERENCES

Aiuppa, Thomas A.; Carney, Robert J. and Krueger, Thomas M. “An Examination of Insurance Stock Prices

over monetary policy reaction functions (see Clarida, Gali, and Gertler, 1999) make a belief-based mechanism more appealing. Marshall (1992) and Bakshi and Chen (1996) offer rational expectations models that generate plausible predictions for the inflation–stock returns relationship. McDevitt (1989) focuses on a nominal tax explanation: In countries with nominal tax components to their tax systems, inflation increases the effective tax rate; an imperfect correlation between stock returns and inflation could then obtain. With the use of an international data set, McDevitt finds little empirical support for the nominal tax approach.

- Following the 1989 Loma Prieta Earthquake." *Journal of Insurance Issues and Practices*, March 1993, 16(1), pp. 1-14.
- Bakshi, Gurdip S. and Chen, Zhiwu. "Inflation, Asset Prices, and the Term Structure of Interest Rates in Monetary Economies." *Review of Financial Studies*, Spring 1996, 9(1), pp. 241-75.
- Barberis, Nicholas; Huang, Ming and Santos, Tano. "Prospect Theory and Asset Prices." *Quarterly Journal of Economics*, February 2001, 116(1), pp. 1-53.
- Barberis, Nicholas, and Thaler, Richard. "A Survey of Behavioral Finance." NBER Working Paper No. w9222, National Bureau of Economic Research, September 2002.
- Barsky, Robert B. and DeLong, Bradford J. "Why Does the Stock Market Fluctuate?" *Quarterly Journal of Economics*, May 1993, 108(2), pp. 291-311.
- Benartzi, Shlomo and Thaler, Richard H. "Myopic Loss Aversion and the Equity Premium Puzzle." *Quarterly Journal of Economics*, February 1995, 110(1), pp. 73-92.
- Borgida, Eugene and Brekke, Nancy. "The Base Rate Fallacy in Attribution and Prediction," in John H. Harvey, W. Ickes, and R. Kidd, eds., *New Directions in Attribution Research*. Volume 3. Hillsdale, NJ: Erlbaum, 1981.
- Borenstein, Severin and Zimmerman, Martin B. "Market Incentives for Safe Commercial Airline Operation." *American Economic Review*, December 1988, 78(5), pp. 913-35.
- Bosch, Jean-Claude E.; Eckard, Woodrow and Singal, Vijay. "The Competitive Impact of Air Crashes: Stock Market Evidence." *Journal of Law and Economics*, October 1998, 41(2, Part 2), pp. 503-19.
- Brennan, Michael J. and Xia, Yihong. "Stock Price Volatility and Equity Premium." *Journal of Monetary Economics*, April 2001, 47(2), pp. 249-83.
- Broder, Ivy E. "The Cost of Accidental Death: A Capital Market Approach." *Journal of Risk and Uncertainty*, March 1990, 3(1), pp. 51-63.
- Bullard, James and Duffy, John. "Learning and Excess Volatility." *Macroeconomic Dynamics*, April 2001, 5(2), pp. 272-302.
- Camerer, Colin. "Individual Decision Making," in John H. Kagel and Alvin E. Roth, eds., *The Handbook of Experimental Economics*. Chap. 8. Princeton: Princeton University Press, 1995, pp. 587-703.
- Campbell, John Y. "Asset Pricing at the Millennium." *Journal of Finance*, August 2000, 55(4), pp. 1515-67.
- Cagle, Julie A.B. "Natural Disasters, Insurer Stock Prices, and Market Discrimination: The Case of Hurricane Hugo." *Journal of Insurance Issues*, 1996, 19, pp. 53-68.
- Clarida, Richard; Gali, Jordi and Gertler, Mark. "The Science of Monetary Policy: A New Keynesian Perspective." *Journal of Economic Literature*, December 1999, 37(4), pp. 1661-707.
- Cochrane, John H. *Asset Pricing*. Princeton, NJ: Princeton University Press, January 2001.
- DeLong, J. Bradford; Shleifer, Andrei; Summers, Lawrence H. and Waldmann, Richard J. "Noise Trader Risk in Financial Markets." *Journal of Political Economy*, August 1990a, 98(4), pp. 703-38.
- DeLong, J. Bradford; Shleifer, Andrei; Summers, Lawrence H. and Waldmann, Richard J. "Positive Feedback Investment Strategies and Destabilizing Rational Speculation." *Journal of Finance*, June 1990b, 45(2), pp. 375-95.
- DeLong, J. Bradford; Shleifer, Andrei; Summers, Lawrence H. and Waldmann, Richard J. "The Survival of Noise Traders in Financial Markets." *Journal of Business*, January 1991, 64(1), pp. 1-19.
- Economist*. "Beyond the Business Cycle?" 23 October 1999.
- El-Gamal, Mahmoud A. and Grether, David M. "Are People Bayesian? Uncovering Behavioral Strategies." *Journal of the American Statistical Association*, December 1995, 90(432), pp. 1137-45.
- Evans, George and Ramey, Garey. "Expectations Calculation and Macroeconomic Dynamics." *American Economic Review*, March 1992, 82(1), pp. 207-24.
- Fama, Eugene F. and Schwert, William G. "Asset Returns and Inflation." *Journal of Financial Economics*, November 1977, 5(2), pp. 115-46.
- Folkes, Valerie S. "The Availability Heuristic and Perceived Risk." *Journal of Consumer Research*, June 1988, 15(1), pp. 13-23.

- Geske, Robert and Roll, Richard. "The Fiscal and Monetary Linkage between Stock Returns and Inflation." *Journal of Finance*, March 1983, 38(1), pp. 1-33.
- Gordon, Myron J. "Dividends, Earnings, and Stock Prices." *Review of Economics and Statistics*, May 1959, 41(2, Part 1), pp. 99-105.
- Grether, David M. "Testing Bayes Rule and the Representativeness Heuristic: Some Experimental Evidence." *Journal of Economic Behavior and Organization*, January 1992, 17(1), pp. 31-57.
- Grether, David M. "Bayes Rule as a Descriptive Model: The Representativeness Heuristic." *Quarterly Journal of Economics*, November 1980, 95(3), pp. 537-57.
- Hirshleifer, David. "Investor Psychology and Asset Pricing." *Journal of Finance*, August 2001, 56(4), pp. 1533-97.
- Kahneman, Daniel and Tversky, Amos. "On the Psychology of Prediction." *Psychological Review*, 1973, 80, pp. 237-51.
- Kahneman, Daniel; Slovic, Paul and Tversky, Amos, eds. *Judgement Under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press, 1982.
- Kunreuther, Howard; Ginsberg, R.; Miller, L.; Sagi, P.; Borhan, B. and Katz, N., eds. *Disaster Insurance Protection: Public Policy Lessons*. New York: John Wiley Interscience, 1978.
- Lamb, Reinhold P. "An Exposure-Based Analysis of Property-Liability Insurer Stock Values Around Hurricane Andrew." *Journal of Risk and Insurance*, March 1995, 62(1), pp. 111-23.
- Lang, Larry H.P. and Stulz, Rene M. "Contagion and Competitive Intra-industry Effects of Bankruptcy Announcements." *Journal of Financial Economics*, August 1992, 32(1) pp. 45-60.
- Lehmann, Bruce N. "Asset Pricing and Intrinsic Values: Review Essay." *Journal of Monetary Economics*, December 1991, 28(3), pp. 485-500.
- LeRoy, Stephen F. and Porter, Richard D. "The Present Value Relation: Tests Based on Implied Variance Bounds." *Econometrica*, May 1981, 49(3), pp. 555-74.
- Lucas, Robert E. Jr. "Asset Prices in an Exchange Economy." *Econometrica*, November 1978, 46(6), pp. 1429-45.
- Marshall, David A. "Inflation and Asset Returns in a Monetary Economy." *Journal of Finance*, September 1992, 47(4), pp. 1315-42.
- McDevitt, Catherine. "The Role of the Nominal Tax System in the Common Stock Returns/Expected Inflation Relationship." *Journal of Monetary Economics*, July 1989, 24(1), pp. 93-107.
- Mitchell, Mark L. and Maloney, Michael T. "Crisis in the Cockpit? The Role of Market Forces in Promoting Air Travel Safety." *Journal of Law and Economics*, October 1989, 32(2, Part 1), pp. 329-55.
- Mullainathan, Sendhil. "A Memory Based Model of Bounded Rationality." Unpublished manuscript, MIT, April 1998.
- Mullainathan, Sendhil. "A Memory-Based Model of Bounded Rationality." *Quarterly Journal of Economics*, August 2002, 117(3), pp. 735-75.
- Nelson, Charles R. "Inflation and Rates of Return on Common Stocks." *Journal of Finance*, May 1976, 31(2), pp. 471-83.
- Nethercutt, Leonard L. and Pruitt, Stephen W. "Touched by Tragedy: Capital Market Lessons from the Crash of Valuejet Flight 592." *Economics Letters*, November 1997, 56(3), pp. 351-58.
- Perez-Quiros, Gabriel and Timmermann, Allan. "Firm Size and Cyclical Variation in Stock Returns." *Journal of Finance*, June 2000, 55(3), pp. 1229-62.
- Rabin, Matthew. "Psychology and Economics." *Journal of Economic Literature*, March 1998, 36(1), pp. 11-46.
- Rabin, Matthew. "A Perspective on Psychology and Economics." *European Economic Review*, April 2002, 46(4-5), pp. 657-85.
- Shelor, Roger M.; Anderson, Dwight C. and Cross, Mark L. "Gaining from Loss: Property-Liability Insurer Stock Values in the Aftermath of the 1989 California Earthquake." *Journal of Risk and Insurance*, September 1992, 59(3), pp. 476-88.
- Sherman, Steven J. and Corty, Eric. "Cognitive Heuristics," in *The Handbook of Social Cognition*. Volume 3, Chap. 6. New York: Lawrence Earlbaum Associates, 1984, pp. 189-286.

- Shiller, Robert J. "Do Stock Prices Move Too Much To Be Justified by Subsequent Changes in Dividends?" *American Economic Review*, June 1981, 71(3), pp. 421-36.
- Shiller, Robert J. *Irrational Exuberance*. Princeton, NJ: Princeton University Press, 2000.
- Shiller, Robert J. "From Efficient Markets Theory to Behavioral Finance." *Journal of Economic Perspectives*, Winter 2003, 17(1), pp. 83-104.
- Shleifer, Andrei and Vishny, Robert. "The Limits to Arbitrage." *Journal of Finance*, March 1997, 52(1), pp. 35-55.
- Simon, Herbert A. "A Behavioral Model of Rational Choice." *Quarterly Journal of Economics*, February 1955, 69(1), pp. 99-118.
- Timmermann, Allan. "Structural Breaks, Incomplete Information, and Stock Prices." *Journal of Business and Economic Statistics*, July 2001, 19(3), pp. 299-314.
- Tirole, Jean. "Asset Bubbles and Overlapping Generations." *Econometrica*, November 1985, 53(6), pp. 1499-528.
- Tversky, Amos and Kahneman, Daniel. "Availability: A Heuristic for Judging Frequency and Probability." *Cognitive Psychology*, 1973, 5, pp. 207-32.
- Tversky, Amos and Kahneman, Daniel. "Judgment under Uncertainty: Heuristics and Biases." *Science*, 1974, 185, pp. 1124-31.
- Weil, Philippe. "On the Possibility of Price Decreasing Bubbles." *Econometrica*, November 1990, 58(6), pp. 1467-74.
- Yamori, Nobuyoshi and Kobayashi, Takeshi. "Do Japanese Insurers Benefit from a Catastrophic Event? Market Reactions to the 1995 Hanshin-Awaji Earthquake." *Journal of the Japanese and International Economies*, March 2002, 16(1), pp. 92-108.

Appendix

A MODEL

Consider an event χ that agents are attempting to forecast with an associated indicator X_T member of $\{0, 1\}$, where $X_T = 1$ if event χ occurs in period T . An agent's subjective probability, her estimate of the probability that $X_T = 1$ conditional on the information set Ω_T , is determined by a function $P: \Omega_T \rightarrow [0, 1]$, which maps the information set into the probability space. The agent's information set consists of past realizations of X_T ,

$$\tilde{X}_{T-1} = \{X_1, X_2, \dots, X_{T-1}\},$$

as well as current informative and uninformative information,

$$\tilde{Y}_T = \{Y_1, Y_2, \dots, Y_T\}, \text{ and}$$

$$\tilde{Z}_T = \{Z_1, Z_2, \dots, Z_T\},$$

respectively.²¹ Suppose further that the event χ is serially uncorrelated and that information useful in forecasting χ in period t is useful only for that period.²² That is, we assume that

²¹ Here, information is uninformative if $\Pr[X_T = 1 | Z_T] = \Pr[X_T = 1]$.

²² These assumptions are employed for simplicity of exposition and are not necessary for the development of the model.

$$\Pr[X_T = 1 | \tilde{X}_{T-1}, \tilde{Y}_T] = \Pr[X_T = 1 | Y_T].$$

The rational expectations solution can then be written

$$P^R: Y_T \rightarrow [0, 1],$$

where the function P^R follows Bayes's law (3). Thus, a rational agent with information Ω_T has subjective probability

$$\Pr[X_T = 1 | Y_T] = \frac{\Pr[Y_T | X_T = 1] \Pr[X_T = 1]}{\Pr[Y_T]}.$$

Now consider a model in which the agent employs the heuristics outlined in the previous sections. We follow Mullainathan (2002) in assuming that the memory processes of agents is incomplete, i.e., that agents *forget* some realizations of X_T . That is, the agent's recalled information set, $\hat{\Omega}_T$, can be written as

$$(4) \quad \hat{\Omega}_T = \{\hat{X}_T, Y_T, Z_T\},$$

where

$$(5) \quad \hat{X}_T = \{X_{1,T}^\dagger, X_{2,T}^\dagger, \dots, X_{T-1,T}^\dagger\}$$

and

$$X_{t,T}^\ddagger = \begin{cases} X_t & \text{with probability } p_{t,T} \\ 0 & \text{with probability } 1 - p_{t,T} \end{cases}$$

Essentially, $X_{t,T}^\ddagger$ is a combination of two indicators: whether the event occurs in time t and whether time t is remembered in time T . The likelihood that an event that occurred in period t is recalled in period T is a function of the time since its last recall, associated events in time t (that is, Y_t and Z_t), and the current environment, Y_T and Z_T . Define $\alpha_{t,T} = \alpha(Y_t, Z_t, Y_T, Z_T)$, where the function $\alpha(\cdot)$ measures the distance between the points (Y_t, Z_t) and (Y_T, Z_T) in Cartesian space. We can then write the likelihood of recalling the event of time t at a later time T as

$$(6) \quad p_{t,T} = F(X_{t,T-1}^\ddagger, \alpha_{t,T}),$$

where the function $F(\cdot)$ has the following properties: $F_1(\cdot) > 0$ and $F_2(\cdot) < 0$. The former indicates that periods recalled in period $T-1$ are more likely to be recalled in T . The latter stipulation indicates that it is more likely that t will be recalled the closer that the current environment is to elements temporally associated with period t .

In this framework, the agent forms an estimate of the likelihood of χ in time T based, in part, on how closely T resembles any time $t < T$ in which χ occurred through the closeness function $\alpha(\cdot)$.

Limited memory makes the probability judgments noisy and biased toward salient events that may or may not be informative. Elements of the agent's information set are a subset of the total information available. Thus, the agent's update has the property that forgetting the occurrence of an event in the past will decrease the subjective probability estimate.

Additionally, salience increases the perceived probability, since salience increases the likelihood of recall. Moreover, since the information set varies over time, the volatility of the estimated probability is greater than the volatility of a learned probability with perfect recall (e.g., OLS learning). In the perfect recall case, information gathered over time reduces the volatility. In the limited-memory case, information that is forgotten biases the subjective probability down, while recalled probability biases it up, each period inducing higher volatility.

Agents make errors in neglecting base rates and consequently bias subjective probabilities upward when they perceive that new information is relevant. It can be shown that, regardless of the direction new information should move the posterior probability,

agents employing an updating function that neglects base rates necessarily overestimate the value of the new information. Agents' subjective probabilities are biased toward their anchor.

A Formal Financial Application

This appendix develops a formal application of the heuristic probability judgments in an asset pricing model. We will initially follow Lucas (1978) to develop a simple general equilibrium framework to study the effects of subjective probabilities.

There are two assets: a risk-free, discount bond with a price B_t and interest rate $r_t^f = \frac{1}{B_t} - 1$; and a stock with price S_t . The stock pays out an infinite stream of *perishable*, real dividends, $\{D_s\}_{s=t}^\infty$, whose growth rates, $d_t \equiv \frac{D_t}{D_{t-1}} - 1$, follow a Bernoulli process:

$$d_t = \begin{cases} d_h & \text{with probability } \pi_D \\ d_l & \text{with probability } 1 - \pi_D \end{cases} \quad 0 \leq \pi_D \leq 1.^{23}$$

Agents' information set consists of a finite sample space, Ω_T , comprising all sequences of the form

$$\omega_T = \{X_{\{d_1=d_h\}}, X_{\{d_2=d_h\}}, \dots, X_{\{d_T=d_h\}}, X_{\{z_1=z_h\}}, X_{\{z_2=z_h\}}, \dots, X_{\{z_T=z_h\}}\},$$

where $X_{\{\cdot\}}$ denotes a standard indicator function. Each ω_T provides a record of possible sequences of dividend growth rates and realizations of an uninformative variable Z_t . In our previous notation, $\Omega_T = \{\bar{X}_T, \bar{Z}_T\}$. For simplicity, the only informative variable is dividends. Also, assume that the uninformative variable Z_t follows another binomial distribution *independent* of dividends, i.e., the rate of change of Z_t , $z_t \equiv \frac{Z_t}{Z_{t-1}} - 1$, follows a Bernoulli process:

$$z_t = \begin{cases} z_h & \text{with probability } \pi_Z \\ z_l & \text{with probability } 1 - \pi_Z \end{cases} \quad 0 \leq \pi_Z \leq 1.$$

Therefore, the joint probability measure of each realization ω_T is given by

$$(9) \quad P(\omega_T) = [\pi_D^j (1 - \pi_D)^{T-j}] \cdot [\pi_Z^i (1 - \pi_Z)^{T-i}] \quad 0 \leq j \leq T \quad 0 \leq i \leq T,$$

where ω_T is any state characterized by both j occur-

²³ Without loss of generality, assume that $d_h > d_l > -1$, so that dividends are nonnegative provided $D_0 > 0$.

rences of high dividend growth and i occurrences of a high rate of growth of Z . While the marginal probability for dividends is

$$P(\{X_{\{d_1=d_h\}}, X_{\{d_2=d_h\}}, \dots, X_{\{d_T=d_h\}}\}) = [\pi_D^j (1 - \pi_D)^{T-j}],$$

since D and Z are independent.

From basic asset pricing principles, in the absence of risk aversion, we can find the price of both assets as the present discounted value of the future stream of cash flows generated by each of them:

$$S_t = E_t[\beta(S_{t+1} + D_{t+1})]$$

$$B_t = E_t[\beta] = \beta,$$

where $\beta = \frac{1}{1 + \rho}$, $\rho > 0$ is the subjective rate of impatience and $E_t[\cdot]$ denotes the conditional expectation operator measurable with respect to available information. Under the assumption of risk neutrality, this simple asset pricing model is a specialization of a classical present-discounted value dividend model (see Lehmann, 1991) to the binomial distribution case.

In the full-information case where the parameters $(\pi_D, d_h, d_l, \pi_Z, z_h, z_l)$ of the joint process for D and Z are known to the agent, a solution for asset prices can be obtained easily using the method of undetermined coefficients. Since the lattices for D and Z are independent, Z does fail to convey any useful information concerning D and an agent will rationally base her portfolio and pricing decisions on the marginal probability measure for D only, a standard (transformation of a) binomial distribution parameterized by $\{\pi_D, d_h, d_l\}$. It is then possible to demonstrate that

$$(10) \quad S_t^{FI} = \lim_{T \rightarrow \infty} E_t \left[\sum_{j=1}^T \left(\beta^j \prod_{i=1}^j D_{t+i} / D_{t+i-1} \right) \right] \cdot D_t.$$

This solution for the equilibrium stock price under full information shows that the stock price is a simple, constant multiple of dividends. Ψ^{FI} denotes the constant pricing kernel or, equivalently, the price-dividend ratio. The explicit solution to (10) can then be derived as

$$(11) \quad S_t^{FI} = \Psi^{FI} D_t = \frac{1 + d_l + \pi(d_h - d_l)}{\rho - d_l - \pi(d_h - d_l)} D_t = \frac{1 + E[d_t]}{\rho - E[d_t]} D_t,$$

while the full-information bond price, B_t^{FI} , is

$$B^{FI} = \frac{1}{1 + \rho} > 0.$$

Equivalently, the time-invariant equilibrium risk-free rate, r^{FI} , is simply $\rho > 0$. Since $r^{FI} = \rho$, it is straightforward to rewrite (11) as

$$S_t^{FI} = \frac{1 + E[d_t]}{r^{FI} - E[d_t]} D_t,$$

which shows the exact equivalence between the solution under full-information rational expectations and the classic Gordon's (1959) formula, popular in applied corporate finance.

The Heuristics-Based Solution

Suppose that π_D is unknown and must be subjectively calculated in a recursive fashion. A subjective assessment of π_D at time t is equivalent to calculating a probability function $\hat{P}^S(\hat{\Omega}_t) \equiv \Pr[X_{t+k} = 1 | \hat{\Omega}_t]$ for all $k \geq 1$, where $\hat{\Omega}_t$ is as defined in (4). The agent believes that the value of the stock depends not only on future dividend payments, i.e.,

$$\hat{P}^S(\hat{\Omega}_t) \equiv \Pr[X_{t+k} = 1 | \hat{X}_t, \hat{Z}_t] \text{ for all } k \geq 1,$$

but also on irrelevant information she uses. Assume that the agent's information set is $\hat{\Omega}_t = \{\hat{X}_t, \hat{Z}_t\}$, where

$$\hat{X}_t \equiv \{X_{j,t}^*\}_{j=1}^t$$

$$(12) \quad X_{j,t}^* = \begin{cases} X_t & \text{with probability } p_{j,t} \\ 0 & \text{with probability } 1 - p_{j,t} \end{cases}$$

and

$$p_{j,t} = \theta^{t-j} \alpha(Z_j, Z_t) \text{ with } \frac{d\alpha(Z_j - Z_t)}{d(Z_j - Z_t)} < 0, \text{ and } \theta < 1.$$

The parameter θ is an additional, subjective discount factor applied to information flows: Since $\theta < 1$, the further back in time an observation on the dividend growth rate is, the more unlikely it is that it will belong to the recalled information set $\hat{\Omega}_t$. $\alpha(Z_j - Z_t)$ might be simply taken to be the inverse of the Euclidean distance between Z_j and Z_t ²⁴:

$$\alpha(Z_j - Z_t) \equiv \frac{1}{1 + \sqrt{(Z_j - Z_t)^2}}.$$

²⁴ In the presence of a drifting process for Z , it would be advisable to de-mean Z_t by subtracting δ_t and de-mean Z_j by subtracting δ_j , where δ is the drift.

The current observation on the dividend growth rate belongs to $\hat{\Omega}_t$ since $\alpha(Z_j - Z_t) = \theta^0 = 1$. Also, the probability that $X_{j,t}$ is a member of $\hat{\Omega}_t$ is a function not only of the growth rate of Z but its level as well.

Under our binomial assumptions, maximum-likelihood delivers the following (recalled) sample proportion estimator:

$$\hat{P}^S(\hat{\Omega}_t) = \hat{\pi}_D^S(\hat{\Omega}_t) = \frac{\sum_{j=1}^t X_{j,t}^*}{\sum_{j=1}^t I_{\{x_{j,t}^* = x_{j,t}\}}},$$

where $I_{\{x_{j,t}^* = x_{j,t}\}}$ is another indicator variable that takes a value of 1 when $X_{j,t}$ is a member of $\hat{\Omega}_t$ (it was recalled) and zero otherwise.

It is now straightforward to calculate the incomplete information, equilibrium asset prices in the presence of heuristic biases. When $\rho > \hat{\pi}_D^S(\hat{\Omega}_t) d_h + (1 - \hat{\pi}_D^S(\hat{\Omega}_t)) d_l$, the heuristic-based stock price, S_t^H , is given by

$$\begin{aligned} S_t^H(\theta, \gamma, \alpha(\cdot)) &= \Psi^H(\hat{\Omega}_t; \theta, \gamma, \alpha(\cdot)) D_t \\ &= \frac{1 + d_l + \hat{\pi}_D^S(\hat{\Omega}_t)(d_h - d_l)}{\rho - d_l - \hat{\pi}_D^S(\hat{\Omega}_t)(d_h - d_l)} D_t \\ (13) \quad &= \frac{1 + \hat{E}_t^S[d_{t+1}]}{r^H - \hat{E}_t^S[d_{t+1}]} D_t, \end{aligned}$$

while the full-information bond price, B_t^H , is $r^H = \rho > 0$.²⁵ Equation (13) differs from the full-information result as the equilibrium stock price stops being a fixed multiple of dividends; on the contrary, Ψ_t^H is time-varying, fitting the empirical observations that price-dividend ratios are subject to long swings. The variation in the price-dividend ratio derives from changes in the memory-based conditional expectation, $\hat{E}_t^S[d_{t+1}]$. More generally, observe that even in the absence of strong changes in dividends, (12) itself implies that Ψ_t^H ought to display considerable variability.

Excess Volatility of Stock Prices

From (10) it follows that²⁶

$$\begin{aligned} 1 + r_t^{FI} &\equiv \frac{S_t^{FI} + D_t}{S_{t-1}^{FI}} = \frac{\Psi^{FI} D_t + D_t}{\Psi^{FI} D_{t-1}} = \\ &\Psi^{FI} d_t + \frac{1}{\Psi^{FI}} d_t = \left(\Psi^{FI} + \frac{1}{\Psi^{FI}} \right) d_t \approx \Psi^{FI} d_t, \end{aligned}$$

so that the volatility of gross stock returns, $1 + r_t^{FI}$, is a constant factor $\sqrt{\Psi^{FI}}$ times the volatility of the rate of growth of fundamentals. Since empirical research has shown stock returns to be over ten times more volatile than fundamentals, this reveals an inconsistency, as $\sqrt{\Psi^{FI}} > 10$ implies $\Psi^{FI} > 100$, too high a price-dividend ratio.

To the contrary, (13) shows that the excess stock volatility puzzle disappears when the price-dividend ratio, Ψ , that maps dividends into equilibrium stock prices is time-varying, as a result of limited recall capabilities. In this case

$$1 + r_t^H = \left(\Psi_t^H + \frac{1}{\Psi_t^H} \right) d_t \approx \Psi_t^H d_t$$

so that

$$Var[1 + r_t^H] = Var[\Psi_t^H] + Var[d_t] + 2Cov[\Psi_t^H, d_t].$$

When $Cov[\Psi_t^H, d_t] > 0$, an increase of the volatility of stock returns (as a result of heuristic biases) will obtain in a full-information framework. Such a case is highly likely under the heuristic rules of our framework.

²⁵ The notation stresses that heuristic-based stock prices do depend on the strength and structure of the assumed biases, as represented by the parameters θ, γ , and the functional form of $\alpha(\cdot)$. $\hat{E}_t^S[\cdot]$ is an expectation taken with respect to the subjective probability assessment of the agent.

²⁶ The approximation is justified by realistic values of the price-dividend ratio in excess of 20 to 30.

The Efficient Market Hypothesis and Identification in Structural VARs

Lucio Sarno and Daniel L. Thornton

For a variety of reasons economists have long been interested in measuring the economy's response to exogenous shocks. The shocks are thought to result, for example, from specific unexpected policy actions, sources that are exogenous to the domestic economy (such as an oil price shock), or sudden changes in technology. The economic structure (or data-generating process) that determines any economic outcome must be inferred from the observed data, and a structural interpretation of the data is obtained from economic theory. However, there are alternative economic theories and, consequently, alternative structural interpretations of the same observations. Hence, economists are faced with the very difficult problem of discriminating among these interpretations and, consequently, identifying the specific source of the shock or the economy's response to it.

Before a structural model can be evaluated, it must be identified. A structural model is identified when one can obtain the structural parameters from the estimates of the reduced-form parameters. A model is "just identified" when there is a one-to-one correspondence between the structural parameters and the reduced-form parameters. On the other hand, a model is over-identified if there is more than one set of structural parameters that is consistent with a given set of reduced-form parameters, whereas it is unidentified when there is no way to obtain the structural parameters from the estimated reduced-form parameters.¹

Generally speaking, there have been two broad approaches to identification, the Cowles Commission (CC) methodology and the so-called structural vector

autoregression (SVAR) methodology.² As a consequence of Sims's (1980) critique of the CC methodology, the SVAR methodology has become arguably the most widely used method of structural analysis. Both methodologies assume that the structural economy can be approximated by a linear, dynamic system of structural equations with an additive stochastic structure. In applications of the CC methodology, identification was typically achieved by placing restrictions (typically homogenous, i.e., zero, restrictions) on some of the coefficients of a dynamic structural model of the economy. While it was well understood that identification could be achieved by placing restrictions on the stochastic structure of the model, this was seldom done in practice.³

In contrast, in the SVAR methodology (which is attributed to Bernanke, 1986; Blanchard and Watson, 1986; and Sims, 1986) identification is achieved by imposing contemporaneous restrictions on both the structure of the economy and the stochastic structure of the model.⁴ Exclusion restrictions on the structural dynamics—which were frequently imposed in applications of the CC methodology—are never imposed.

The restrictions that the SVAR methodology imposes on the structural shocks have often been

¹ When a model is over-identified, there is a set of over-identifying restrictions that can be tested as part of a structural model evaluation.

² The Cowles Commission methodology is attributable to various researchers who were in one way or another connected to the Cowles Commission for Research in Economics. For a summary of this methodology, see Koopmans (1949). For an early application of it, see Klein (1950).

³ See Koopmans (1949) for a discussion of variance-covariance restrictions.

⁴ We note that there are identification schemes that impose no contemporaneous restrictions. This literature includes the work of Blanchard and Quah (1989) and Shapiro and Watson (1988). This methodology is not discussed here. See Keating (1992) for an excellent survey of structural VAR approaches to identification.

Lucio Sarno is a professor of finance, Finance Group, Warwick Business School, University of Warwick, and a research affiliate, Centre for Economic Policy Research (CEPR), London. Daniel L. Thornton is a vice president and economic advisor at the Federal Reserve Bank of St. Louis. The authors thank Martin Sola, Hiroshi Fujiki, Oscar Jorda, and Yi Wen for comments. John Zhu provided research assistance.

Federal Reserve Bank of St. Louis *Review*, January/February 2004, 86(1), pp. 49-60.
© 2004, The Federal Reserve Bank of St. Louis.

criticized (e.g., Bernanke, 1986; Stock and Watson, 2001), and Cooley and LeRoy (1985) have noted that, in the absence of these restrictions, the estimated shocks from the SVAR would be linear combinations of all the structural shocks in the reduced-form VAR. This paper extends and refines Cooley and LeRoy's observation by noting that if the VAR includes one or more efficient market variables (EMVs)—variables that reflect all information relevant for their determination—the covariance restrictions that are typically employed in a SVAR identification are inappropriate and may have to be replaced with alternative restrictions. Our paper is close in spirit to those of Wallis (1980) and Pesaran (1981) in the rational-expectations literature; however, we focus on SVARs rather than on more general structural rational-expectations models.

Strictly speaking, our analysis applies only to VARs that include variables that are efficient in the strong form of the efficient market hypothesis (EMH). We argue, however, that our analysis is likely to have implications for VARs that include variables that meet the less stringent requirements of semi-strong market efficiency. The potential importance of our critique for applied work is illustrated with a SVAR model that is widely used to identify the effects of monetary policy shocks on the economy.

The paper begins with a brief discussion of the CC and SVAR approaches to identification. We then discuss (i) the EMH and the various forms of market efficiency and (ii) the effect of including an EMV in a SVAR model. The implications of our analysis for applied work are illustrated with a widely used SVAR model.

THE CC AND SVAR METHODS OF IDENTIFICATION

Both the CC and SVAR methods of identification assume that the economy can be approximated by a general linear structural model of the economy of the form

$$(1) \quad AY_t = BY_{t-1} + Dv_t,$$

where Y_t is an $N \times 1$ vector of endogenous variables and v_t is a vector of i.i.d. structural shocks, with mean zero and a constant covariance matrix.⁵ Bernanke

(1986, p. 52) notes that these shocks are “primitive” exogenous forces, not directly observed by the econometrician, which “buffet the system and cause oscillations.” He notes that “because these shocks are primitive, i.e., they do not have a common cause, it is natural to treat them as approximately uncorrelated.” Hence, it is reasonable to assume that $Ev_t v_t' = \Lambda$, where Λ is a diagonal matrix. This noncontroversial assumption is common to both the CC and SVAR approaches.

The reduced-form of the structural model (i.e., what economists observe) is given by

$$(2) \quad Y_t = A^{-1}BY_{t-1} + A^{-1}Dv_t$$

or

$$(3) \quad Y_t = \Gamma Y_{t-1} + u_t,$$

where $\Gamma = A^{-1}B$ and $u_t = A^{-1}Dv_t$.

The economic model is (exactly) identified when it is possible to obtain estimates of the structural parameters (i.e., the elements of A , B , D , and Λ from the reduced-form parameters) and vice versa (i.e., when there is a one-to-one correspondence between the structural and reduced-form parameters). Identification is achieved by placing restrictions on A , B , and D .⁶

In the CC methodology, identification was typically achieved by imposing restrictions on A and B . While it was widely understood that identification could be achieved by imposing restrictions on D or Λ , such restrictions were seldom imposed in practice. There are N^2 unique elements of $\Gamma = A^{-1}B$, but N^2 elements in each of A and B . Hence, the necessary (order) condition for (exact) identification using the CC methodology is that there are as many zero elements (in the case of homogenous restrictions) in B as there are non-zero elements in A . That is, there must be a total of N^2 restrictions imposed on A and B —the fewer the restrictions imposed on A , the more restrictions must be imposed on B . If these restrictions are linearly independent (the rank condition for identification), it is possible to go from the reduced-form parameters to the structural parameters and vice versa.

In response to Sims's (1980) claim that the restrictions placed on B were “incredible,” the SVAR literature has taken a different approach to identification. No restrictions are placed on B . Instead, identification is achieved by placing restrictions on

⁵ The first-order autoregressive structure is used because any higher-order autoregressive process can be written as a first-order process. For presentation purposes, however, we will assume that the model is strictly first order.

⁶ Restrictions can also be imposed on Λ . For example, one might assume that the variance of one structural shock is some multiple of another. This possibility is ignored for ease of presentation.

the elements of A and D . To see how the model is identified in the SVAR literature, note that $Eu_t u_t' = A^{-1} D \Lambda D' A^{-1} = \Sigma$, where Σ is a real symmetric matrix of rank N . An estimate of Σ is obtained by estimating the reduced-form model, i.e.,

$$\sum_{t=1}^T E \hat{u}_t \hat{u}_t' = \hat{\Sigma},$$

where \hat{u}_t is the vector of residuals obtained by estimating equation (3). There are at most $N(N + 1)/2$ unique, non-zero elements of $\hat{\Sigma}$. In contrast, there are N^2 parameters in A , N elements in Λ , and N^2 elements in D . Consequently, there are $2N^2 + N$ structural parameters, so that $(3N^2 + N)/2$ restrictions are needed to satisfy the necessary (order) conditions for identification. Hence, identification can be achieved by imposing $(3N^2 + N)/2$ restrictions on the $2N^2$ elements of A and D .

It is frequently assumed in the SVAR literature that $D = I$. With this assumption, there are only $(N^2 + N)/2$ restrictions that need to be imposed on A . N of these restrictions can be obtained by assuming that the diagonal elements of A are equal to unity (these are normalization restrictions), which leaves $N(N-1)/2$ required restrictions.⁷ In the case of recursive structural VARs (RSVARs), these restrictions come from assuming that A is lower triangular.⁸

THE EMH

The assumption that SVAR models impose on A and D to achieve identification may not hold if the VAR includes one or more EMVs. To see why, it is useful to briefly discuss the EMH (Samuelson, 1965; and Campbell, Lo, and MacKinlay, 1997). Malkiel (1992, p. 739) states that a “market is said to be efficient with respect to an information set, ϕ , if security prices would be unaffected by revealing that information to all participants.” The degree of market efficiency is usually categorized by the nature of the information set. Markets are said to be efficient in the weak form if the information set only includes the history of prices or returns. For the semi-strong form of market efficiency, the information set is all

publicly available information. When market prices reflect the information known to any market participant, they are said to be efficient in the strong form.⁹

Market efficiency is also characterized by the speed with which information is reflected in market prices (e.g., Chordia, Roll, and Subrahmanyam, 2002; and Schwert, 2002). A shock that is initially reflected in only one asset price may, over time, be reflected in other asset prices. The faster the information is reflected in other prices, the more efficient the market is said to be. Financial markets are thought to be efficient with respect to publicly announced (or known) information, in that such information is thought to be rapidly, if not immediately, reflected in asset prices (Malkiel, 1992; and Campbell, Lo, and MacKinlay, 1997). It may take longer for information that is not publicly announced to be incorporated in asset prices; however, a shock that initially affects only one asset price may create arbitrage opportunities. As market participants respond to such opportunities, prices of other assets change. Hence, the longer the period of time over which economic data are averaged, the more likely it is that asset prices will reflect information that is and is not publicly announced.

THE EMH AND THE SVAR IDENTIFICATION

By definition, an EMV responds contemporaneously to all shocks that are relevant for its determination. This means that none of the elements of the row of $A^{-1}D$ corresponding to the EMV are zero. It is not important whether the response of the EMV to structural shocks is due to the form of A or D ; nevertheless, if the assumptions made about the form of A are such that the rows of A^{-1} corresponding to the EMVs are zero, the elements of the rows of D corresponding to these variables must be non-zero.

To better understand why this is so, consider a simple three-variable structural model of the economy represented by equation (1). We initially assume that no identifying restrictions are imposed, so that

$$A^{-1} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix},$$

⁷ Even in cases where $D \neq I$, there is a one-to-one correspondence between the structural shocks and the variables in the VAR (e.g., Bernanke, 1986). This is a consequence of the requirement that D must be $N \times N$.

⁸ This is often referred to as a Wold causal chain in honor of Herman Wold, who advocated the theoretical desirability of recursive models in economics, e.g., Wold (1954). In the case of non-recursive structural VARs, the necessary condition for identification is usually achieved by imposing $N(N-1)/2$ homogenous (or in some cases, non-homogenous) restrictions that are rationalized on the basis of economic theory.

⁹ See Malkiel (1992) for details.

$$B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix},$$

and

$$D = \begin{bmatrix} 1 & d_{12} & d_{13} \\ d_{21} & 1 & d_{23} \\ d_{31} & d_{32} & 1 \end{bmatrix}.$$

The diagonal elements of D are normalized to unity under the assumption that the structural shocks are unique. Now assume that the second variable in the VAR, Y_2 , is an EMV.

For the sake of illustration, assume that $D = I$, so that the reduced-form error is given by

$$(4) \quad u_t = \begin{bmatrix} u_{1,t} \\ u_{2,t} \\ u_{3,t} \end{bmatrix} = \begin{bmatrix} a_{11}v_{1,t} + a_{12}v_{2,t} + a_{13}v_{3,t} \\ a_{21}v_{1,t} + a_{22}v_{2,t} + a_{23}v_{3,t} \\ a_{31}v_{1,t} + a_{32}v_{2,t} + a_{33}v_{3,t} \end{bmatrix},$$

where v_1 , v_2 , and v_3 denote the first, second, and third primitive structural shocks, respectively.

Note that the reduced-form shocks are related to the structural shocks solely by the structure of A . While the point made above applies to any SVAR model, for ease of illustration, we assume a RSVAR, i.e., A is assumed to be lower triangular. With this assumption, equation (4) reduces to

$$(5) \quad u_t = \begin{bmatrix} a_{11}v_{1,t} \\ a_{21}v_{1,t} + a_{22}v_{2,t} \\ a_{31}v_{1,t} + a_{32}v_{2,t} + a_{33}v_{3,t} \end{bmatrix}.$$

Under these assumptions, the first shock is reflected only in the first reduced-form residual, the first and second structural shocks are reflected in the second reduced-form residual, and so on and so forth.

Note that equation (5) is incompatible with our assumption that Y_2 is an EMV because, under the assumptions made about A and D , Y_2 responds only to the first and second structural shocks. Hence, given the assumptions made about the structure of A , the EMH requires alternative assumptions be made about the structure of D .

In the case of a RSVAR, one way the model can be made consistent with the EMH is by letting the EMV appear last in the Choleski ordering. The placement of the variables in the ordering in RSVARs is usually based on economic arguments, however. Hence, changing the recursive ordering in a RSVAR

is tantamount to making different assumptions about the structure of the economy. Hence, while placing the EMV last in the recursive ordering overcomes the problem we discuss in this paper, it need not be the "correct" solution.

Alternatively, one could maintain the Choleski ordering and relax the assumptions on D . In this example this can be achieved by assuming that $d_{23} \neq 0$. In this case, the reduced-form errors would now be given by

$$(6) \quad u_t = \begin{bmatrix} v_{1,t} \\ a_{21}v_{1,t} + v_{2,t} + d_{23}v_{3,t} \\ a_{31}d_{21}v_{1,t} + a_{32}v_{2,t} + (a_{32}d_{23} + 1)v_{3,t} \end{bmatrix}.$$

$Y_{2,t}$ responds to the third structural shock due to the assumption that $d_{23} \neq 0$. If the model were exactly identified, however, an additional restriction must be imposed on either A or Λ (e.g., $\sigma_{v_i}^2 = \sigma_{v_j}^2$ for some i and j) to satisfy the necessary conditions for identification.

Note that if there are two or more EMVs in the VAR, placing these variables last in the Choleski ordering will not overcome the problem unless one of the EMVs deviates from the other by an idiosyncratic shock.¹⁰ If the recursive structure of A is maintained, identification will have to be achieved by imposing additional restrictions on either A or Λ .

IMPLICATIONS OF THE EMH FOR APPLIED WORK

How important is our analysis of the EMH for applied work? This is a difficult question to answer for at least two reasons. First, in general, the answer depends on the nature of the variables included in the SVAR and the structural restrictions imposed for identification. Consequently, the importance of including an EMV in the SVAR must be analyzed on a case-by-case basis.

Second, strictly speaking, our analysis holds only if the SVAR includes a variable that is efficient in the strong form of the EMH, and strong-form market efficiency is a stringent condition that is unlikely to be satisfied in the real world. We believe that our critique may apply to SVARs that include financial market variables that are likely to be efficient in the weak form or in the semi-strong form of the EMH, such as stock prices, interest rates, or possibly exchange rates.

¹⁰ For example, the expectations hypothesis holds.

Hence, one area of research where we believe that our analysis is likely to apply is the relatively large body of empirical work devoted to identifying the effects of monetary policy shocks using RSVARs. In a large strand of this literature, U.S. monetary policy shocks are identified using time series on a short-term interest rate—most often the effective federal funds rate—at monthly or lower frequencies, using a RSVAR (e.g., Christiano, Eichenbaum, and Evans, 1996, 1999).

While it is perhaps unlikely that short-term interest rates reflect *all* market information, there is considerable evidence to suggest that they reflect all publicly available information rather quickly. That is, short-term interest rates (and interest rates, more generally) are likely to satisfy the conditions for the semi-strong form of market efficiency. Further, these markets are dominated largely by public information, with private information playing a limited role relative to, for instance, the stock market. Indeed, the evidence suggests interest rates respond quickly to information that market participants believe is important for determining the stance of monetary policy. For example, interest rates responded quickly to unexpected changes in the stock of money during the period when the Fed was implementing monetary policy by targeting M1 from October 1979 through October 1982 (e.g., Cornell, 1982, 1983; Roley and Walsh, 1985; and Thornton, 1989). There is also a large body of literature showing that interest rates respond rapidly to a variety of macroeconomic information, albeit different information at different times (see Fleming and Remolona, 1997, for a summary of this literature), and respond intra-day to a number of macroeconomic announcements (e.g., Fleming and Remolona, 1999). To the extent that shocks to macroeconomic variables also reflect such information, the identifying restrictions imposed in the RSVAR will be violated.

The longer the period of time over which interest rates are averaged, the more likely it is that all rates will reflect information that was initially reflected in only one rate. That is, it becomes more likely that interest rates will reflect information that is not publicly known. Hence, the covariance restrictions frequently imposed for identification are more problematic the longer the period of time over which interest rates are averaged.

Some analysts might argue that our conclusion that short-term nominal interest rates are likely to satisfy the EMH runs counter to the treatment of short-term rates in many monetary policy analyses,

where the short-term interest rate is treated as a choice variable of the central bank. In this case, the short-term rate need not be an EMV because changes in it are made entirely in response to past information. However, if one takes seriously the evidence that interest rate rules are forward looking (e.g., Clarida, Gali, and Gertler, 2000), it is plausible that the short-term interest rate is consistent with the EMH even if it is determined solely by decisions of the central bank. Consequently, regardless of whether the short-term interest rate is determined by the market or determined by the central bank, it seems possible that the interest rate behaves in a manner consistent with the EMH under certain assumptions.^{11,12}

To investigate the significance of our critique for applied work, we estimate a seven-variable VAR similar to that estimated by Christiano, Eichenbaum, and Evans (1999). The variables used are industrial production, *Y*; the price level as measured by the consumer price index, *CPI*; the *Journal of Commerce* commodity price index, *CP*; the effective federal funds rate, *FF*; nonborrowed reserves, *NBR*; total reserves, *TR*; and the broad monetary aggregate, *M2*. With the exception of *CP* and *NBR*, the variables are identical to those used by Christiano, Eichenbaum, and Evans (1999). All of the variables except the funds rate are in natural logs. The data are monthly for the period 1959:01 to 2001:07. Following Christiano, Eichenbaum, and Evans (1999), the lag order is 12.¹³

Christiano, Eichenbaum, and Evans employ the Choleski factorization with the ordering $\{Y, P, CP, FF, NBR, TR, M2\}$. Our analysis suggests, however, that if *FF* is an EMV, it should come last in the

¹¹ Applying the EMH only to financial prices is at odds with some recent New Keynesian literature. Some of the models in this literature employ forward-looking relations for output (and inflation) that resemble asset-pricing conditions (e.g., Estrella and Fuhrer, 2002). Of course, extending our analysis to such variables would make the identification problems discussed here even more severe because EMH-type behavior might be expected of other variables in the VAR besides financial prices or interest rates.

¹² If the above argument applies, it applies primarily to the federal funds rate and then only over periods when the Fed explicitly targeted the funds rate. Moreover, Sims (1998) found that the qualitative results were unaffected by using either the Fed's discount rate or the commercial paper rate. We confirmed his finding; however, these results are not reported. The impulse response functions obtained using the federal funds rate were very similar to those obtained using a variety of other short-term rates.

¹³ Qualitatively and quantitatively similar results are obtained with shorter lag lengths.

Figure 1

Impulse Response Functions for *Y* with *FF* Fourth and Last in the Cholesky Ordering

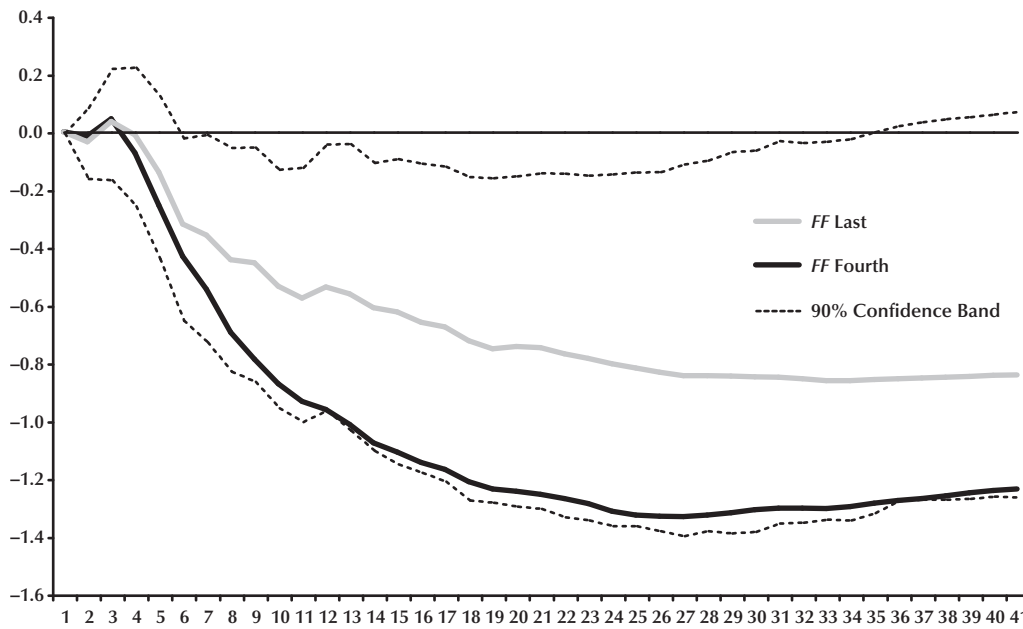


Figure 2

Impulse Response Functions for *CPI* with *FF* Fourth and Last in the Cholesky Ordering

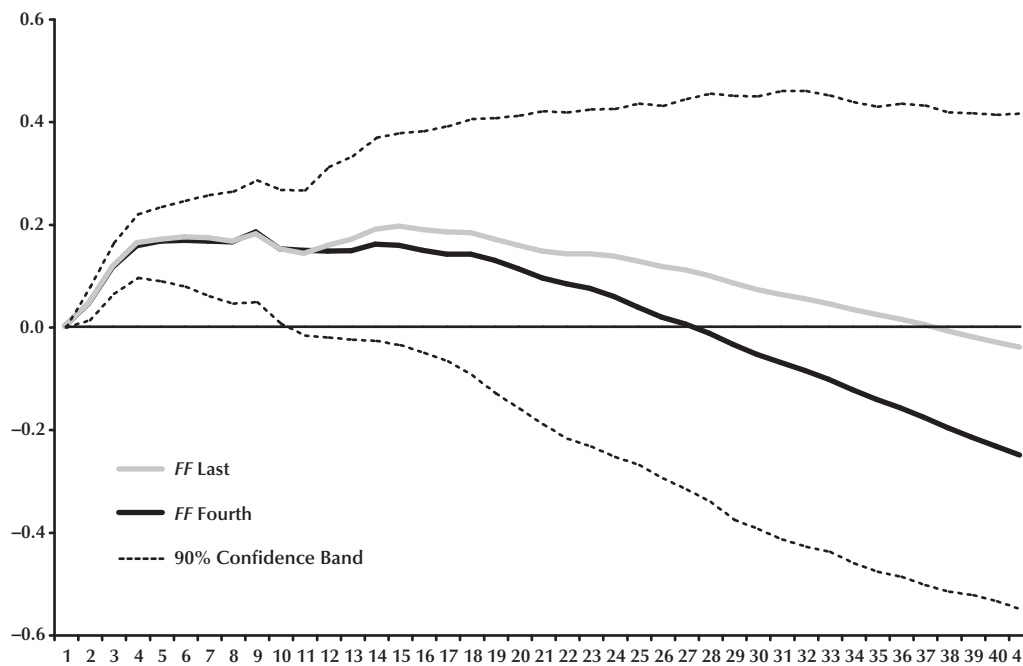


Figure 3

Impulse Response Functions for CP with FF Fourth and Last in the Cholesky Ordering

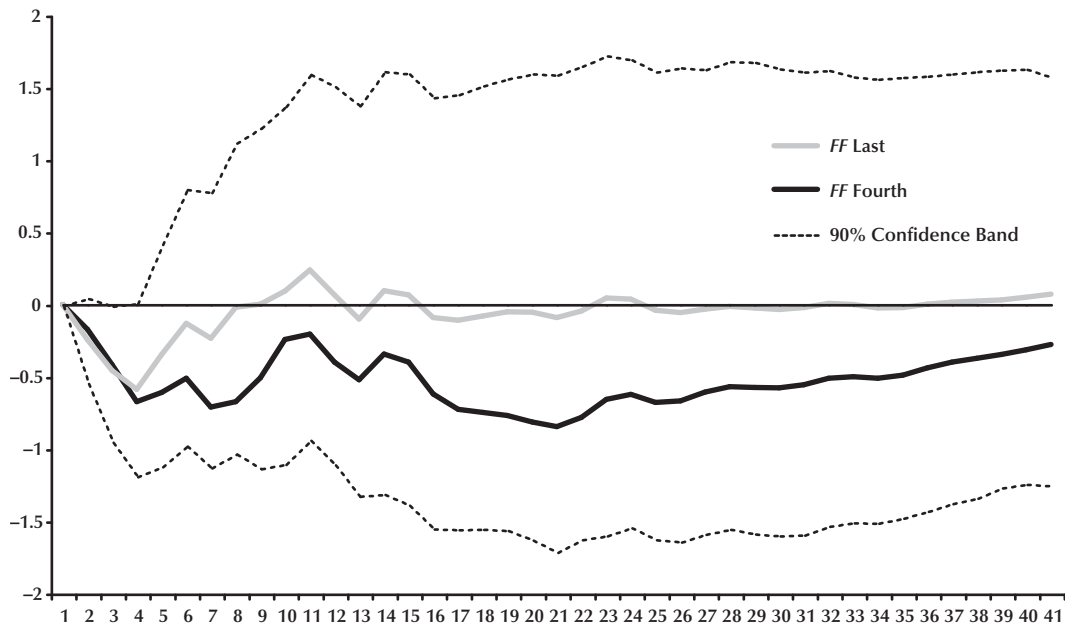


Figure 4

Impulse Response Functions for NBR with FF Fourth and Last in the Cholesky Ordering

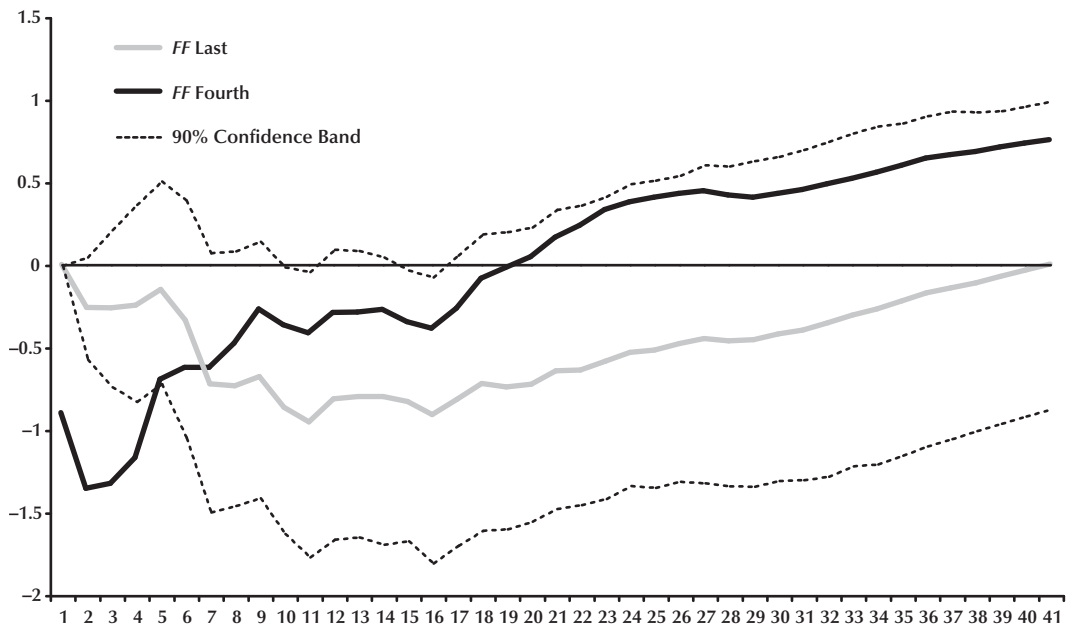


Figure 5

Impulse Response Functions for TR with FF Fourth and Last in the Cholesky Ordering

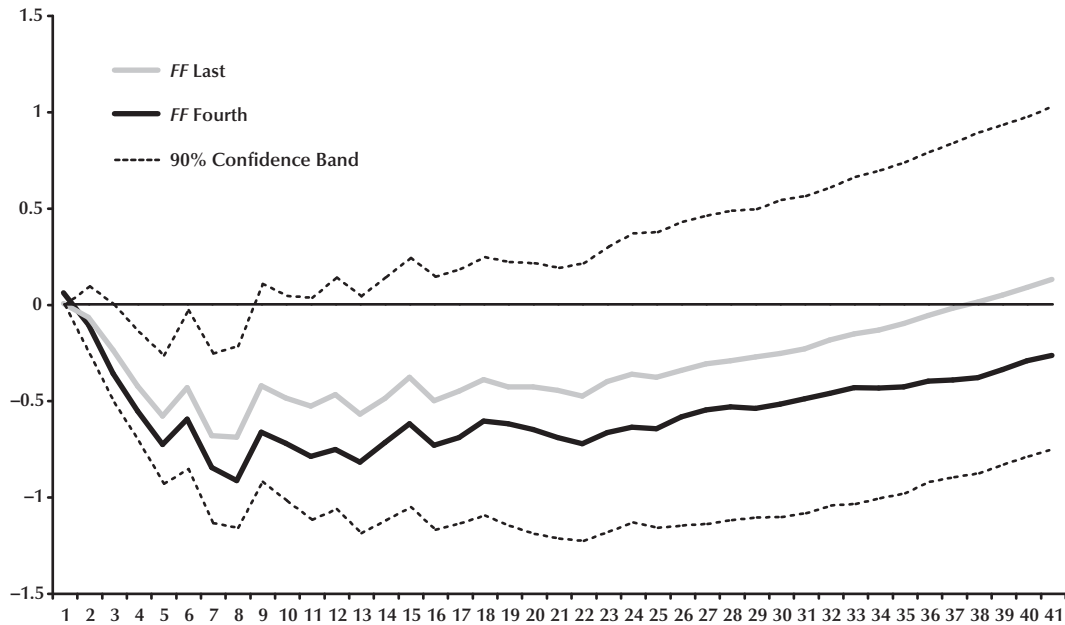


Figure 6

Impulse Response Functions for M2 with FF Fourth and Last in the Cholesky Ordering

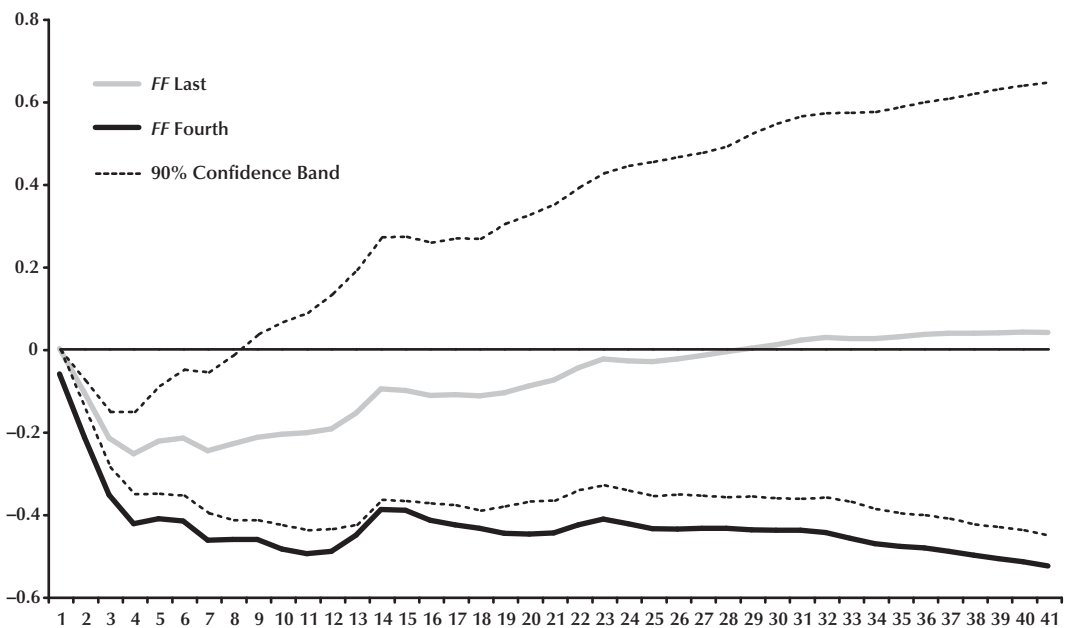
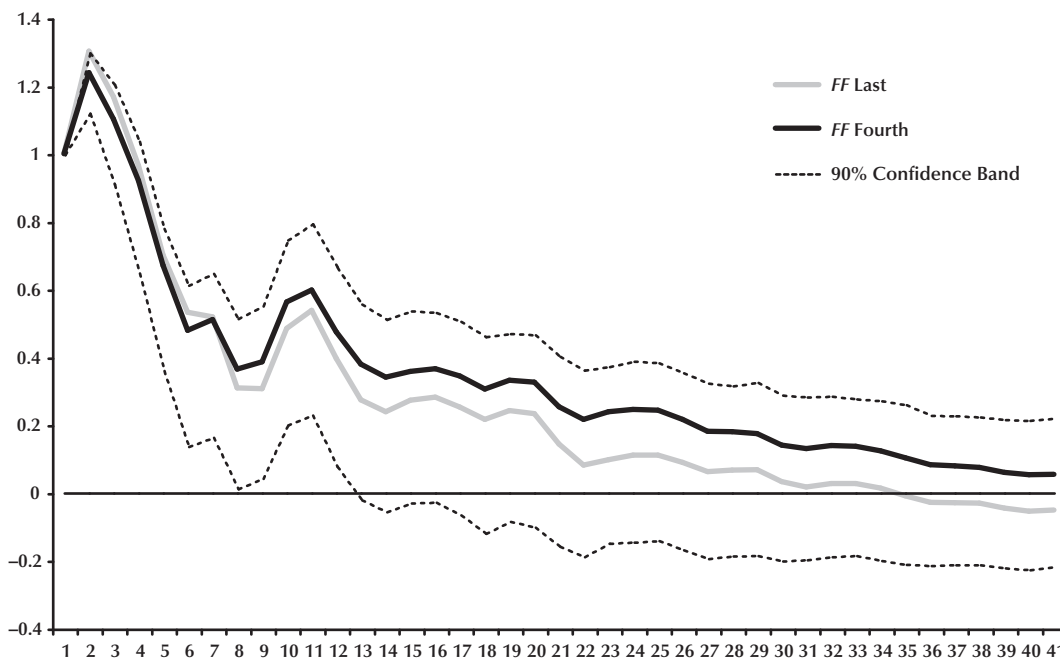


Figure 7

Impulse Response Functions for *FF* with *FF* Fourth and Last in the Cholesky Ordering



Choleski ordering. Hence, we compare the results with two orderings: $\{Y, P, CP, FF, NBR, TR, M2\}$ and $\{Y, P, CP, NBR, TR, M2, FF\}$.

Figures 1 through 7 show the impulse response functions of each of the variables to a one-unit shock to the funds rate when the funds rate is fourth in the ordering (the solid black line) and when the funds rate is last in the ordering (the solid gray line); they also show the 90 percent confidence interval (the dashed lines) for the impulse response functions obtained when *FF* comes last in the Choleski ordering. The confidence intervals are obtained by bootstrapping the model using 500 iterations. The effect of placing the funds rate in the middle rather than last in the recursive ordering is sometimes large—particularly for a funds rate shock on output, where the effect with the funds rate in the middle drifts to the lower bound of the 90 percent confidence interval.

The effect is also large for *NBR*. This is not surprising because there is a strong and contemporaneous link between *NBR* and the funds rate (Pagan and Robertson, 1995; and Thornton, 2001). Thornton (2001) has shown that this relationship is due to the

Fed's operating procedure, which caused *NBR* to respond contemporaneously and endogenously to changes in the funds rate over much of this period. In any event, when the funds rate is last in the ordering, the contemporaneous relationship between the funds rate and *NBR* is accounted for in the funds rate equation. With the contemporaneous relationship between the funds rate and *NBR* accounted for in the funds rate equation, shocks to the funds rate have no significant effect on *NBR*. Moreover, consistent with Thornton's (2001) analysis of the Fed's operating procedure, the effect of shocks to the funds rate on *NBR* and *TR* is similar.

It is well known that the response to a shock may vary with the Choleski ordering. In this respect, these results are perhaps not surprising. We have provided a rationale for why the response is likely to change with the recursive ordering in some cases. Hence, these results raise doubts about the implications obtained from RSVARs. Note that while we obtained different results by placing the potential EMV last in the Choleski ordering, we are not advocating this as a "solution" to the problem of identification when RSVARs contain a potential EMV. We are only suggesting that these results are consistent

with our overall conclusion that special care should be taken when identifying SVARs that include an EMV.

Of course, if the VAR includes two or more financial market variables, such as interest rates, stock prices, or exchange rates, identification is even more complicated.¹⁴ Such variables may be efficient at least in the semi-strong form of the EMH and, hence, will quickly reflect publicly known information. For example, a policy or other announcement that affects interest rates is likely also to affect stock prices or exchange rates. Garfinkel and Thornton (1995), who investigated the relationship between the federal funds rate, the overnight repo rate, and the 3-month T-bill rate using weekly average and daily data, found that shocks to interest rates that cause a differential between the funds rate and other rates were quickly eliminated. They also found that the idiosyncratic shocks to interest rates, as they identify them, are not correlated with three measures of monetary policy actions, suggesting that monetary policy actions were quickly reflected in market interest rates, including the federal funds rate. Consistent with these results, Sarno and Thornton (2003) found that disturbances to the equilibrium between the daily funds rate and the 3-month T-bill rate dissipate very rapidly.

The problem is that if economic variables contemporaneously reflect the same information, structural identifying assumptions that impose the condition that shocks do not affect such variables contemporaneously will be violated. While the importance of this critique for applied work is an empirical question, the empirical analysis presented here supports the argument that covariance restrictions imposed in the SVAR literature may be inappropriate and that greater caution should be exercised in choosing the identifying restrictions in such models.

SUMMARY AND CONCLUSIONS

The SVAR methodology identification is frequently applied by imposing restrictions that prevent economic variables from responding contemporaneously to one or more structural shocks. This paper shows that such restrictions are not applicable if the variable is efficient in the strong form of the

EMH because EMVs, as we term them, respond to all information.

While, strictly speaking, our analysis applies only to variables that are efficient in the strong form of the EMH, the longer the period of time over which the data are measured, the more likely it is that the variables that are efficient in the semi-strong or weak forms of the EMH will reflect information that was initially known only to relatively few market participants. Hence, our analysis is likely to have implications for empirical analyses that use variables that are efficient in the semi-strong or weak forms of the EMH, especially when data are measured at monthly and quarterly frequencies and for markets where public (as opposed to private) information is dominant.

We illustrate the potential importance of our analysis by estimating a RSVAR often used to identify the effects of monetary policy shocks on the economy. Our results suggest that some of the effects of monetary policy shocks, so identified, are sensitive to whether the interest rate is ordered in the middle of the VAR, as is most often the case, or at the end (which avoids the problem in RSVARs that include only one EMV). This does not imply that one can simply overcome the problem by putting the EMV last in the Choleski ordering. It does, however, support our conclusion that researchers need to be extremely careful when using the standard contemporaneous identifying restrictions employed in the SVAR methodology when the VAR includes one or more variables that may satisfy some form of the EMH. Caution is particularly required when the data employed are at the monthly or quarterly frequency, as is often the case in applied macroeconomics and monetary economics.

REFERENCES

- Bernanke, Ben S. "Alternative Explanations of the Money-Income Correlation." *Carnegie-Rochester Conference Series on Public Policy*, Autumn 1986, 25(0), pp. 49-99.
- Blanchard, Olivier Jean and Quah, Danny. "The Dynamic Effects of Aggregate Demand and Supply Disturbances." *American Economic Review*, September 1989, 79(4), pp. 655-73.
- Blanchard, Olivier Jean and Watson, Mark W. "Are Business Cycles All Alike?" in Robert J. Gordon, ed., *The American Business Cycle*. Chicago: University of Chicago Press, 1986.
- Campbell, John Y.; Lo, Andrew W. and MacKinlay, A. Craig.

¹⁴ In an effort to estimate the effect of monetary policy actions—shocks to the federal funds rate—on the yield curve, Evans and Marshall (1998) estimate a number of SVARs that include the effective federal funds rate and a "long-term rate," with maturities ranging from one month to ten years. Evans and Marshall use three alternative identifying assumptions, including a Choleski ordering.

- The Econometrics of Financial Markets*. Princeton: Princeton University Press, 1997.
- Chordia, Tarun; Roll, Richard and Subrahmanyam, Avanidhar. "Evidence on the Speed of Convergence to Market Efficiency." Unpublished manuscript, 2002.
- Christiano, Lawrence J.; Eichenbaum, Martin and Evans, Charles L. "The Effects of Monetary Policy Shocks: Evidence from the Flow of Funds." *Review of Economics and Statistics*, February 1996, 78(1), pp. 16-34.
- Christiano, Lawrence J.; Eichenbaum, Martin and Evans, Charles L. "Monetary Policy Shocks: What Have We Learned and To What End?" in John B. Taylor and Michael Woodford, eds., *Handbook of Macroeconomics*. Volume 1A. New York: Elsevier Science/North-Holland, 1999, pp. 65-148.
- Clarida, Richard H.; Gali, Jordi and Gertler, Mark. "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory." *Quarterly Journal of Economics*, February 2000, 115(1), pp. 147-80.
- Cooley, Thomas F. and LeRoy, Stephen F. "Atheoretical Macroeconomics: A Critique." *Journal of Monetary Economics*, November 1985, 16(3), pp. 283-308.
- Cornell, Bradford. "Money Supply Announcements, Interest Rates, and Foreign Exchange." *Journal of International Money and Finance*, August 1982, 1(0), pp. 201-08.
- Cornell, Bradford. "The Money Supply Announcement Puzzle: Review and Interpretation." *American Economic Review*, September 1983, 73(4), pp. 644-57.
- Estrella, Arturo and Fuhrer, Jeffrey C. "Dynamic Inconsistencies: Counterfactual Implications of a Class of Rational-Expectations Models." *American Economic Review*, September 2002, 92(4), pp. 1013-28.
- Evans, Charles L. and Marshall, David A. "Monetary Policy and the Term Structure of Nominal Interest Rates: Evidence and Theory." *Carnegie-Rochester Conference Series on Public Policy*, December 1998, 49(0), pp. 53-111.
- Fleming, Michael J. and Remolona, Eli M. "What Moves the Bond Market?" Federal Reserve Bank of New York *Economic Policy Review*, December 1997, 3(4), pp. 31-50.
- Fleming, Michael J. and Remolona, Eli M. "Price Formation and Liquidity in the U.S. Treasury Market: The Response to Public Information." *Journal of Finance*, October 1999, 54(5), pp. 1901-15.
- Garfinkel, Michelle R. and Thornton, Daniel L. "The Information Content of the Federal Funds Rate: Is It Unique?" *Journal of Money, Credit, and Banking*, August 1995, 27(3), pp. 838-47.
- Keating, John W. "Structural Approaches to Vector Autoregressions." Federal Reserve Bank of St. Louis *Review*, September/October 1992, 74(5), pp. 37-57.
- Klein, Lawrence R. *Economic Fluctuations in the United States, 1921-1941*. New York: Wiley and Sons, 1950.
- Koopmans, Tjalling C. "Identification Problems in Economic Model Construction." *Econometrica*, April 1949, 17(2), pp. 125-44.
- Malkiel, Burton G. "Efficient Market Hypothesis," in Peter Newman, Murray Milgate, and John Eatwell, eds., *The New Palgrave Dictionary of Money & Finance*. New York: MacMillan Press, 1992, pp. 739-44.
- Pagan, Adrian R. and Robertson, John C. "Resolving the Liquidity Effect." Federal Reserve Bank of St. Louis *Review*, May/June 1995, 77(3), pp. 33-54.
- Pesaran, M. Hashem. "Identification of Rational Expectations Models." *Journal of Econometrics*, August 1981, 16(3), pp. 375-98.
- Roley, V. Vance and Walsh, Carl E. "Monetary Policy Regimes, Expected Inflation and the Response of Interest Rates to Money Announcements." *Quarterly Journal of Economics*, Supplement 1985, 100(4), pp. 1011-39.
- Samuelson, Paul. "Proof that Properly Anticipated Prices Fluctuate Randomly." *Industrial Management Review*, 1965, 6, pp. 41-49.
- Sarno, Lucio and Thornton, Daniel L. "The Dynamic Relationship Between the Federal Funds Rate and the Treasury Bill Rate: An Empirical Investigation." *Journal of Banking and Finance*, June 2003, 27(6), pp. 1079-110.
- Shapiro, Matthew D. and Watson, Mark W. "Sources of Business Cycle Fluctuations," in Stanley Fischer, ed., *NBER Macroeconomic Annual 1988*. Cambridge, MA: MIT Press, 1988, pp. 111-48.

- Sims, Christopher A. "Macroeconomics and Reality." *Econometrica*, January 1980, 48(1), pp. 1-48.
- Sims, Christopher A. "Are Forecasting Models Usable for Policy Analysis?" Federal Reserve Bank of Minneapolis *Quarterly Review*, Winter 1986, 10(1), pp. 2-16.
- Sims, Christopher A. "The Role of Interest Rate Policy in the Generation and Propagation of Business Cycles: What Has Changed Since the '30s?" in Jeffery Fuhrer and Scott Schuh, eds., *Beyond Shocks: What Causes Business Cycles?* Conference Series No. 42, Federal Reserve Bank of Boston, 1998, pp. 121-60.
- Stock, James H. and Watson, Mark W. "Vector Autoregressions." *Journal of Economic Perspectives*, Fall 2001, 15(4), pp. 101-15.
- Schwert, G. William. "Anomalies and Market Efficiency." NBER Working Paper No. w9277, National Bureau of Economic Research, October 2002.
- Thornton, Daniel L. "The Federal Reserve's Operating Procedure, Nonborrowed Reserves, Borrowed Reserves and the Liquidity Effect." *Journal of Banking and Finance*, September 2001, 25(9), pp. 1717-39.
- Thornton, Daniel L. "The Effect of Unanticipated Money on the Money and Foreign Exchange Markets." *Journal of International Money and Finance*, December 1989, 8(4), pp. 573-87.
- Wallis, Kenneth F. "Econometric Implications of the Rational Expectations Hypothesis." *Econometrica*, January 1980, 48(1), pp. 49-74.
- Wold, H. "Causality and Econometrics." *Econometrica*, April 1954, 22(2), pp. 162-77.