

Lessons Learned From the Recent Business Cycle

Based on President Santomero's Hutchinson Lecture at the University of Delaware, April 12, 2005

BY ANTHONY M. SANTOMERO

The U.S. economy enjoyed a remarkable run in the 1990s. As it moved into the new century, however, the economy underwent various fits and starts before entering its current expansion phase. In this quarter's message, President Santomero shares his views on the U.S. economy and outlines some of the lessons learned from the most recent business cycle.

This quarter, I would like to share my views on the U.S. economy and some of the lessons learned from our recent business cycle. By way of perspective, it should be remembered that the U.S. economy enjoyed a remarkable run in the 1990s. Then, it stumbled as we came into the new century and struggled to find solid footing, going through numerous fits and starts early in the new millennium. Now, in 2005, the recession and recovery phases of the current cycle are behind us, and the economy has entered an expansion phase that I expect will carry us forward for some time. As the economy moves along this path of self-sustaining growth, the Federal Reserve has been steadily removing the accommodative monetary policy that has been in place over the past few years, as it moves toward a more neutral policy stance.

In reflecting on the current business cycle and the turbulent times surrounding it, I will focus on how recent events, as well as ongoing trends, have affected both the economy and the

conduct of monetary policy in this cycle. I will also address how they will influence the economy going forward and how I see the economic expansion progressing.

As most readers will appreciate, it is important that we learn from the experiences of the past. As the saying goes: "Those who cannot remember the past are condemned to repeat it." Hopefully, some of the lessons we learned from our recent past will be incorporated into the policy decisions we make in the future. Nonetheless, before we start, I must remind you that every business cycle is different. Each is the unique product of (1) a relentlessly evolving economic structure, (2) some surprising new developments, and (3) a sequence of policy actions attempting to stabilize the situation. This most recent experience is no exception.

EXAMINING THE CONTEXT

To discuss the most recent business-cycle experience, one must start at the beginning: with the revolution in information and communications

technology and its dramatic effect on the economic structure of the U.S. Cheap hardware, sophisticated software, and extensive networking capabilities — both Internet and intranet — began transforming business processes in earnest in the latter half of the 1990s. Of course, this was a worldwide phenomenon, but it clearly had profound effects on the U.S. economy.

History tells us, and our most recent experience reconfirms, that a technological revolution of this magnitude does not produce a smooth economic progression. It is, by its nature, disruptive to the existing order of things. Nonetheless, the application of new information technologies brought real economic benefits to our economy. As these technologies were introduced into organizations and infused into business processes, productivity measurably accelerated.

At the same time, however, it spawned unrealistic expectations that



Anthony M. Santomero, President,
Federal Reserve Bank of Philadelphia

were manifested in a stock market bubble and overinvestment in new capital. When the bubble burst and the investment boom deflated, aggregate demand decelerated rapidly, ultimately driving the economy into recession.

The technology revolution has also been an important contributor to globalization — a second fundamental factor of structural change driving the economy's evolution in this business cycle. By slashing communications costs, new technologies made the markets for financial assets, goods and services, and even labor, more globally integrated. Globalization was driven by other forces as well. Freer trade among nations and, even more fundamentally, the triumph of the market system over centralized planning were both movements that spurred integration.

Like the introduction of new technologies, the globalization of the marketplace has been and continues to be a good thing. It fosters greater specialization and gains from trade, affording everyone higher living standards. These benefits are genuine and worthwhile, but they do not come without some costs. The adjustment costs are significant, and in an environment of rapid change, they are ongoing.

I will say more about technology and globalization later in the article. But first, let me turn to the second ingredient of any business cycle, that is, the arrival of new developments and unexpected events.

SHOCKS TO THE ECONOMY

There were several new and surprising developments during the most recent business cycle. We often refer to these events as economic shocks. In 2000, the U.S. stock market declined precipitously and the tech bubble burst. The NASDAQ, which was valued at just under 5000 in March 2000, fell to under 2000 in April 2001. This led to a decrease in national wealth and had

a negative effect on the economy as a whole. The Dow suffered a similar, if less dramatic, decline, as well.

This was followed by certainly the most profound event affecting the course of the recent business cycle: the terrorist attacks of September 11, 2001. It goes without saying that September 11 stands as one of the most shocking and tragic episodes in our nation's history.

Like the introduction of new technologies, the globalization of the marketplace has been and continues to be a good thing...but the adjustment costs are significant.

The physical effects of September 11 were readily apparent. We saw the great loss of life, the horrific sights of the collapsing twin towers in New York, the damaged Pentagon in Washington, and the smoldering wreckage of a jet in western Pennsylvania. Yet, in purely economic terms, the immediate impact on the productive capacity of the U.S. was relatively small when measured against our collective resources — our labor force and our capital infrastructure. Longer term, there have been productivity losses that are more difficult to quantify, namely, those created by enhanced security procedures in airports, office buildings, and mailrooms.

In any case, the events of September 11 had an immediate and profound contractionary effect on the demand side of the economy. At first, shock, fear, and uncertainty paralyzed everyone. We were absorbed by what happened, and we tried to figure out what it meant for our country and ourselves personally. Meanwhile, we cancelled air travel and hotel reservations and put all but essential spending on hold.

All things considered, consumer spending came back relatively quickly.

But for businesses, it was a much different story. Already left with an overhang of equipment from the investment boom of the late 1990s, businesses confronted these new uncertainties about the future and saw new reasons to defer and delay investment spending.

The events that followed in the aftermath of September 11 — the anthrax attacks and then the wars in

Afghanistan and Iraq — only served to heighten these uncertainties. In the case of Iraq, the uncertainties were extended and indeed to some extent still remain. First, there was uncertainty about whether war with Iraq would come, then about how the war would go, and now about whether we can secure the ultimate objective there — a politically stable and economically successful nation.

Meanwhile, as the U.S. economy began on its path to a slow recovery, accounting scandals and corporate governance issues created new uncertainties, and what some referred to as another “soft spot” in the economy. Scandals surrounding Enron and Worldcom, just to name two of the largest, undermined confidence and created mistrust of large corporations in the U.S. psyche. This further heightened investor uncertainty and weakened both households' and businesses' willingness to spend. For businesses, this rise in investor skepticism increased risk spreads in credit markets, raising the cost of capital faced by firms at least for a time.

Beyond the financial markets' reaction, these revelations also triggered

reforms legislated under the Sarbanes-Oxley Act. The act was designed to boost investor confidence in corporate America by improving the quality of corporate disclosure and financial reporting and increasing the role and responsibility of corporate officers and directors. Compliance with Sarbanes-Oxley focused companies' attention and resources on their audit, accounting, and governance processes, and it remains a topic of conversation in the corporate suites and boardrooms around our nation. While this may have been appropriate and necessary, it also has diverted companies' attention from new investment projects and slowed plans for future expansion.

Completing the list of disturbances buffeting our economy is one more major shock that hit the economy in 2004: a sharp increase in both the price and the volatility of the price of oil. The international benchmark jumped from \$20 per barrel in early 2002 to over \$50 per barrel in late 2004. It has been oscillating around this higher figure since late last year.

POLICY DURING THE CYCLE

Thus far, I have talked about the structural changes and surprising developments affecting the shape of the current business cycle. But how has the third factor, namely, policymakers' actions, affected economic dynamics over the past few years?

Here, I would contend that remarkably aggressive policy action was a defining characteristic of this business cycle. Indeed, monetary and fiscal policy worked together particularly well this time around to provide ample and rapid stimulus during the economic downturn.

The National Bureau of Economic Research has determined that the U.S. economy fell into recession in March 2001. On the monetary policy side, the Fed had begun reducing the fed

funds rate two months earlier, in January 2001, and had dropped it 300 basis points by August. On the fiscal policy side, the Bush administration's first round of tax cuts was enacted in the spring of 2001, and the first tax rebate checks were in the mail by July. With the benefit of hindsight, the timing of this fiscal stimulus was quite fortuitous.

I think a case can be made that, had it not been for September 11, this double dose of strong stimulus might have averted a recession by countering the existing weakness and giving the economy the push it needed to return to a positive growth path. I said so then and remain of that opinion.

In any event, the recession oc-

Looking forward, the economy appears to be on course for a sustained period of solid expansion.

curred, and the recovery was attenuated in its aftermath. In response, both monetary and fiscal policymakers reacted by providing yet additional rounds of stimulus. These policy actions may not have succeeded in turning business investment spending around very quickly, but they certainly helped buoy consumer spending. This kept the economy growing while businesses positioned themselves to re-engage.

In 2004, the U.S. economy had a fairly good year. Output growth of nearly 4 percent and the creation of over 2 million net new jobs lend credence to the argument that the economy has regained its balance and is now on a path of sustained expansion. And this occurred without a noticeable acceleration in core inflation.

Looking forward, the economy

appears to be on course for a sustained period of solid expansion. I expect real GDP to grow at an annual rate of around 4 percent this year and next, with payroll employment increasing by 150,000 to 200,000 jobs per month.

On the demand side, consumers will continue to spend at a good pace. As I stated earlier, during this most recent recession and recovery, consumer spending held up unusually well, continuously expanding throughout the cycle. Looking forward, steady job growth and rising household incomes will fuel continued growth in consumer spending, replacing the stimulative effects of low interest rates and tax rate reductions, which were key to the earlier period of continued consumption growth.

Going forward, the expansion will be driven by business spending. Firms have ample cash flow and have had significant profit growth. They are now well positioned for greater efficiency and will see the need for greater productive capacity as the expansion continues. For all these reasons, I anticipate that the robust growth in business investment spending we have been experiencing will continue for the foreseeable future. Add to this pattern of private-sector spending moderate growth in government spending on goods and services, and you have solid growth in domestic final sales.

One potential constraint on demand growth that has re-emerged recently is rising oil prices. As I mentioned, we saw oil prices reach over \$50 a barrel in late 2004. Subsequently, they fell back a bit, but now the U.S. economy is faced with oil prices in excess of \$50 a barrel once again. With gasoline prices rising to substantially over \$2 a gallon, consumers may find that growth in their discretionary spending must slow in order to accommodate the increased cost of filling their gas tanks. Similarly, rising

energy costs could curtail businesses' capacity to increase their investment spending. The bottom line is that oil prices persistently in the \$50-per-barrel-plus range could slow the pace of domestic demand growth this year, though they should not jeopardize the expansion itself.

Of course, as we have all become aware, just how much of that domestic *demand* translates into domestic *production* depends on what happens to our international trade balance. Over the past decade, a strong dollar and a relatively strong U.S. economy drove the current account to unprecedented heights. It now represents a sizable percentage of U.S. GDP. In fact, in 2004, the widening trade gap or current account deficit — take your pick — drained more than 1.5 percent from domestic output growth.

Over the past year or so, at least partially in response to the large trade deficits, the dollar has steadily depreciated. A lower dollar should eventually help stabilize our net export position. Though economic growth has been somewhat uneven among our trading partners of late, continued global economic expansion should help as well. As the trade deficit begins to stabilize, solid growth in spending by U.S. consumers and businesses will translate directly into solid growth in real GDP for the U.S.

Having emphasized the output growth in the current expansion, I want to turn to another development that has received considerable attention over this entire cycle and, more recently, as the economy has moved from recovery to expansion. This is the issue of the dynamics of inflation and the potential for price pressures developing as the economy moves along its path of continued growth.

As an economist, I recognize that price pressures are an inevitable part of any business expansion. I think we

all recognize that as the economy continues on its path of expansion, price dynamics are prone to shift. As productivity growth returns to trend, unit labor costs will probably start to rise, potentially putting pressure on prices. We already saw some indications of a shift down toward long-run productivity growth at the end of last year. In addition, higher prices for oil and other commodities may lead producers to try

It is incumbent upon the Fed to make every effort to keep price pressures well contained.

to pass on some of their higher input costs, potentially igniting or exacerbating latent price pressures. Moreover, the recent decline in the value of the dollar may lessen the competitive pressure on domestic producers that has until now limited their pricing power. Recently, I have been hearing from my contacts around the District that price pressures are building and there has been some evidence of firms passing on higher costs in final product prices.

It is incumbent upon the Fed to make every effort to keep price pressures well contained. As long as the public remains confident in the Fed's commitment to essential price stability — and the Fed conducts its policy in a manner consistent with that commitment — transitory adjustments in prices will not generate persistently higher inflation.

The Federal Reserve has already begun the transition from an accommodative policy stance to a neutral one, more consistent with sustained noninflationary economic growth. If the economy evolves as I have suggested here, then I expect we will continue on our present course of moving

the federal funds rate toward neutrality. However, the precise course we take depends on the precise course the economy takes. If signs of heightened price pressure emerge on a consistent basis, we will need to consider quickening the pace at which we move toward policy neutrality.

LESSONS LEARNED

To summarize, the U.S. economy experienced a period of extraordinary growth over the decade of the 1990s, followed by a sharp slowdown in spending on new information and computer technology. Then it was pushed into recession and a tenuous recovery by the September 11 attacks and their aftermath, as well as a series of corporate scandals and other events. Now with these events behind us, I believe the economy is on a course for steady growth at a sustainable pace. This pattern of growth should foster continued job growth and a relatively stable price environment. All in all, economic prospects are reasonably good in the U.S.

Having said that, now is probably a good time to look back at the past four years and try to extract some lessons that policymakers can carry forward to the next business cycle, whenever it may come. In that spirit, I will outline five distinct lessons that I garnered from the experiences of the recent past.

LESSON 1: TECHNOLOGICAL INNOVATION CAN DRIVE A CYCLE

The first lesson that I take away from an examination of our most recent economic episode is that new technologies and investment in new technologies can be powerful drivers of business cycle dynamics. The most recent business cycle, from the historic 10-year expansion to the recession of 2001 and the subsequent recovery, was

an investment-driven one. Growth in investment spending strengthened and sustained the expansion of the 1990s. Then the collapse in business investment spending generated the recession and attenuated the recovery. Finally, the return of business investment spending ushered in the broader economic recovery beginning in 2003.

At the same time, the increased productivity experienced in the late 1990s, owing to the large investment in information and communication technology (ICT), allowed the U.S. economy to produce high levels of output while not experiencing inflationary pressures.

The dynamic at work was that the new, profitable investments being offered in ICT created an increase in productivity, which translated into increased profits, and thus more investing and consuming. At the same time, the increase in productivity growth helped keep down unit labor costs and prices. This led to a period of strong growth and low inflation.

In retrospect, business technology spending in the late 1990s represented a mix of both good and bad business judgments. Some of the ICT spending turned out to be wise and even prescient investment in productive new capital. Some of it was just investment pulled forward for fear that legacy equipment would malfunction in Y2K. And some of it — often associated with ill-conceived “dot-com” business plans — reflected “irrational exuberance” about the viability of new business models.

However, much of this overinvestment can be explained by rational behavior. It may be that in the 1990s, firms were rationally forecasting huge gains in productivity due to the ICT revolution. Firms were very optimistic about the future, so they built up large amounts of capital. This led to increases in output, employment, and invest-

ment. However, when these expectations were not fully met, and it became evident there was an over-buildup in capital, firms stopped investing.

In any case, it took the business sector three years, from 2000 through 2002, to digest those investments. From an accounting perspective, it took three years to depreciate accumulated stock of hardware and software. From an economic perspective, it took three years to put existing capital to its most productive use by reallocating it across firms and fully exploiting its capabilities to boost productivity and cut costs within firms.

The time it took for firms to begin investing again may have been amplified by the large negative shocks

LESSON 2: GLOBALIZATION IS AN IMPORTANT FACTOR IN ECONOMIC DYNAMICS AND INFLATION

A second lesson this most recent business cycle brought into focus is that global dynamics play an important role in the path our domestic economy will follow. There has been considerable discussion concerning the increased role of globalization and its effect on developed economies. This cycle has spotlighted three distinct but interrelated effects the global economy has had on our domestic economy.

The first is the traditional one that focuses on the competitive pressures that globalization has brought to the market for goods and services.

Firms are again investing in everything from high-tech equipment and software to warehouses and equipment, positioning themselves for greater efficiency and greater productive capacity.

I spoke of earlier, and businesses may have been reluctant to increase investment in this environment of uncertainty. But whatever the cause, variation in business spending caused variation in economic activity.

Now, the forces are aligned for strong growth in business investment spending. Firms have had time to fully digest their previous acquisitions of capital. Profits have been strong. The economic outlook is positive, and some of the previous risks and uncertainties are dissipating. Indeed, firms are again investing in everything from high-tech equipment and software to warehouses and equipment, positioning themselves for greater efficiency and greater productive capacity going forward. The U.S. economy is again on a path of sustained expansion.

Here, the impact of the current account on domestic production has been an essential ingredient of the dynamics of the U.S. economy.

In this cycle, the debate expanded to a second area, the labor market, to include the “outsourcing” or “off-shoring” of labor services, a trend tied to the technology revolution. Improvements in information and communications technology are creating a globally integrated marketplace — not only for goods and services but also for labor. Of course, such “off-shoring” has been the trend in much of the production activity associated with manufacturing for a long time. But it seemed to intensify in this cycle, particularly with the opening of several newly developing economies. It also seems to be spreading to the service sector.

Increasingly, then, U.S. firms compete with firms around the world in the markets for raw materials and final goods and services, while U.S. workers compete with workers around the world for positions in a widening array of occupations and industries. From the macroeconomic perspective, this globalization of the marketplace and the increased degree of competition it brings are powerful forces that can alter the wage and price dynamics of the U.S. economy and, indeed, have done so over this cycle, persistently dampening upward price pressures.

The third important aspect of globalization from a U.S. policymaker's perspective is the globalization of capital markets. Indeed, globalization of capital markets has substantively affected both the dynamics of trade and domestic production in this cycle.

Investors, believing the return on capital in the United States to be relatively attractive on a risk-adjusted basis, funneled a large fraction of global wealth into the U.S. capital market. Global investors purchased large quantities of dollar-denominated assets, keeping the dollar's exchange value high through the tech boom — even while the economy went into recession and the current account turned decidedly negative.

The trade-weighted exchange value of the dollar appreciated 35 percent from 1995 to 2001 and stayed strong through 2002. This had a two-prong effect on the U.S. economy. First, it drove up our trade deficit to record levels. Second, it kept a relatively tight lid on inflation by putting low-priced goods on the market in the U.S.

Now, it seems that investors are becoming less willing to channel so much of their savings into additional dollar-denominated instruments. Some have suggested that they are beginning to diversify into other currencies, such as the euro. This has caused the dollar

to depreciate against other currencies. In fact, over the past year, the trade-weighted value of the dollar has fallen about 10 percent.

Gradually, the depreciation of the dollar will translate into lower prices for exports from the U.S. and higher prices for imports into the U.S. Thus, the pattern of output and prices in the

to recover and a reason to invest into a better future. The precise channels through which monetary policy operates may vary from cycle to cycle, but its use in this cycle clearly showed its effectiveness.

Fiscal policy also played a key role in the dynamics of this cycle. Well-timed tax cuts and tax rebates helped

The countercyclical monetary policy the Fed implemented gave consumers the opportunity to borrow at relatively low interest rates, and they certainly seized it.

U.S. in this cycle has been, and will continue to be, affected by the global economy.

LESSON 3: COUNTERCYCLICAL POLICY CAN BE AN EFFECTIVE DEMAND FORCE

The shape of this business cycle was substantively affected by countercyclical government policies. Aggressive use of both monetary and fiscal policy clearly reduced the severity of the recession and accelerated the course of the recovery.

On the monetary policy side, the Federal Reserve reduced its target federal funds rate by 475 basis points — from 6.5 percent to 1.75 percent — in the recession year of 2001. When the recovery threatened to stall, the Fed once again reacted, dropping the target fed funds rate to just 1 percent, its lowest level since the 1950s.

The countercyclical monetary policy the Fed implemented gave consumers the opportunity to borrow at relatively low interest rates, and they certainly seized it. Households increased their purchases of homes and durables at record rates, dampening the breadth and depth of the past recession. They also sustained that growth, which gave business investment both time

sustain consumer spending during the recession and the early stages of the recovery. However, the application of fiscal stimulus is notoriously hard to time properly. The tax cuts enacted in this cycle had been proposed not as countercyclical measures but as part of a long-term shift in tax policy. Their timing was fortuitous.

Moreover, as we are now seeing, it is extremely difficult to remove fiscal stimulus once the economy is on the road to recovery. Indeed, it remains to be seen whether expansive fiscal policies can be reversed and the federal budget can be returned to balance as we move through the expansion phase of the cycle. As an economist, I see the value of fiscal integrity, and this requires a cyclically balanced federal budget.

LESSON 4: MONETARY POLICY WORKS BEST IN A STABLE PRICE ENVIRONMENT

The next lesson I would like to offer is that we have learned that monetary policy works best in a stable price environment. In such an environment, the central bank can reduce interest rates without the fear of increasing inflation expectations. Consumers and businesses perceive the reduction

in real interest rates as temporary and so see it as an opportune time to shift spending forward. By doing so, they dampen the recession. Then, as the recovery proceeds, the private sector can anticipate the actions of the central bank and its plan to return short-term rates to more normal levels.

This played out quite well in the recent cycle. The core PCE was within a 1.5 percent to 2 percent band heading into the recession and has remained in that range during the recovery. This was true even while the Federal Reserve reduced the fed funds target rate in the aggressive manner I have laid out. Not only did the Federal Reserve reduce rates to these historically low levels, but it sent the message that it would keep these rates low for the foreseeable future. In fact, we did, keeping the target fed funds at 1 percent for an entire year.

LESSON 5: EXPECTATIONS MATTER

This discussion brings me to my last lesson, something I have been saying for some time. Expectations matter, and they play an important role in the conduct of national monetary policy. Let me explain why.

The goal of the Federal Reserve is to create financial conditions that foster maximum sustainable economic growth. To achieve this, the Fed must make two important contributions to the economy. First, it is charged with providing essential price stability, meaning little or no inflation. Second, it attempts to offset shifts in demand that deter the economy's ability to reach its potential. These goals are compatible, but each receives different emphasis as the situation warrants.

As a central banker, I recognize that long-run price stability is always

of utmost importance. This means not only a stable price level in the near term but also the expectation of stable prices over the long term. This implies that optimal monetary policy is not simply a matter of establishing a stable price level *today* but of ensuring stable prices — and expectations of price stability — into the *future*. Only then can consumers and investors be confident in the environment in which they must make decisions that have implications far into the future. For this reason, central bankers often talk about the need to establish credibility and the public's confidence in our long-run commitment to price stability.

The Fed can maintain the credibility of its commitment to price stability and avoid sharp changes in public expectations about monetary policy by being as transparent as possible about its own decision-making. As a result, information about the Fed's policy goals, its assessment of the current economic situation, and its strategic direction are increasingly a part of the public record. For some time, the Federal Open Market Committee (FOMC) has released statements after every FOMC meeting. Very recently, the FOMC began releasing the minutes of each meeting prior to the next meeting. They not only report our decisions concerning immediate action but also our sense of the key factors driving near-term economic developments and the strategic tilt to our actions going forward.

Increasing the degree of central bank transparency is one reason I and some of my colleagues have spoken in favor of an explicit inflation-targeting program. I believe we have reached a point where institutionalizing inflation targeting simply makes good sense

from an economic perspective. In short, it is a reasonable next step in the evolution of U.S. monetary policy, and it would help secure full and lasting benefits from our current stable price environment.

Evolving to explicit inflation targeting from our current implicit target has significant potential benefits, and the costs may be minimal if we can implement it in a constructive manner. Clearly, proper implementation of inflation targeting is crucial to its success. That, in turn, requires more research and analysis about how and when to introduce it. But while it requires more public debate and discussion, it may be an idea whose time is approaching.

CONCLUSION

I hope I have convinced you that there are useful lessons to be learned from the dynamics of the recent business cycle in the U.S. While every cycle is unique, each also highlights some enduring realities that bear remembering. Indeed, it is careful attention to both aspects of our experience that moves forward both the science of economics and the art of economic policymaking. If we keep learning, perhaps both the practice of macroeconomic policy and the theory of central banking taught at great universities will advance.

I recognize that no matter how much we learn, the central bank's power will always be limited. I do not think we will ever reach a point where we will eliminate the business cycle! But we may be able to move closer to conducting optimal monetary policy in a world where change is relentless and surprising new developments continue to unfold. ☎

The Relationship Between Capacity Utilization and Inflation

BY MICHAEL DOTSEY AND THOMAS STARK

A

common belief is that when there's slack in the economy — that is, when labor and capital are not fully employed — the economy can expand without an increase in inflation.

One measure of the intensity with which labor and capital are used in producing output is the capacity utilization rate. According to some economists, when capacity utilization is low, firms can increase employment and their use of capital without incurring large increases in the costs of production. So firms will not be forced to raise prices in order to make profits on additional output. But this theory is not universally accepted. In this article, Mike Dotsey and Tom Stark investigate some of the problems with what, at first glance, seems a compelling story.

A commonly held view in economics is that when there is slack in the economy — that is, labor and capital are not fully employed — the economy can expand without an increase in inflation. This idea has a long history in economic theory, with its earliest clear exposition dating back to John Maynard Keynes. There is also recent support for this view. For example, earlier this year Goldman Sachs noted in its newsletter that “core infla-

tion has fallen by about one percentage point over the past year... This disinflation is consistent with the view that resource utilization is indeed too low.”¹ Likewise, in its February 2004 forecast, Macroeconomic Advisers stated that “over the near term, inflation will be held in check by recently exceptional growth in productivity, slack conditions in labor markets, and global excess capacity in many goods markets.”

One measure of the intensity with which labor and capital are used in the production of output is the capacity utilization rate.² When the capacity

¹ Goldman Sachs *Global Economic Research* (newsletter), February 6, 2004.

² The capacity utilization rate is not the only measure that conveys whether resources are

utilization rate is low, implying that there are unemployed workers and idle plant and equipment, it is assumed that firms can increase employment and their use of capital without incurring large increases in the costs of production. Hence, some theories accord with what seems like a very intuitive notion, namely, that firms will not be forced to raise prices in order to make profits on additional output. In that case, output can increase with very little inflation.

However, the above story is not universally accepted, and we shall investigate some of the problems with what, at first glance, seems a compelling story.³ Further, even if the relationship between capacity utilization and inflation were theoretically sound, the strength of the relationship and its usefulness for monetary policy purposes is an empirical matter.

Our empirical research suggests that up to the mid-1980s, capacity utilization is modestly useful in

underutilized. Other common measures are the output gap (which measures the difference between the level of GDP and the level of potential GDP (that is, the level of maximum sustainable GDP), the NAIRU (which is the unemployment rate consistent with stable inflation), and the help-wanted index.

³ An excellent example of a contrary view is given in the 1996 article by Mary Finn.



Mike Dotsey is a vice president and senior economic policy advisor in the Research Department of the Philadelphia Fed.



Tom Stark is the macroeconomic database and policy support manager in the Research Department of the Philadelphia Fed.

helping to explain the behavior of inflation. However, the relationship between utilization and inflation is not a stable one. As the sample period is extended into the mid-1990s, capacity utilization's predictive power wanes or becomes nonexistent. Further, although the economic theory that underpins the intuition discussed above also indicates that the relationship between capacity utilization and inflation would vary with the rate of capacity utilization — with inflation rising more rapidly as capacity utilization increases — we find no evidence that this is the case.

A FIRST LOOK AT THE DATA

The capacity indexes computed by the Federal Reserve Board attempt to measure the ratio of the actual level of output to sustainable maximum or capacity output. The Board defines sustainable maximum output as “the greatest level of output a plant can maintain within the framework of a realistic work schedule, after factoring in normal downtime and assuming sufficient availability of inputs to operate the capital in place.”⁴ Thus, it measures output relative to what could reasonably be called normal output when the plant is employing the usual number of workers and using its machinery at a typical intensity. The capacity level of production is estimated from annual surveys of manufacturing capacity utilization conducted by the Bureau of the Census along with data supplied by other government and private-industry sources. The staff at the Board of Governors use this information to construct estimates of capacity and capacity utilization for industries in manufacturing, mining, and electric

and gas utilities.⁵ Because the survey is yearly, changes in the capacity utilization rate largely reflect actual movements in production.⁶

We begin our investigation of the relationship between capacity utilization and inflation by plotting the two series over the period 1959 to 2003.⁷ Examining the relationship between capacity utilization and inflation, we see that there are periods when utilization and inflation move in the

same direction and even when the movements in utilization precede movements in inflation (Figure 1). For example, in 1972 manufacturing capacity utilization increased from roughly 77 percent to 88 percent and was followed by an increase in annual inflation of 8 percentage points. Likewise in 1976, manufacturing capacity utilization increased a dramatic 14 percentage points and was followed by an increase in the inflation rate of 4 percentage points. Moreover, the relationship between utilization and inflation has not just involved positive responses. In 1974, utilization declined 16 percentage points, and inflation soon decreased 5 percentage points. On the other hand, we see large increases as well as high levels of utilization throughout the 1990s, and inflation steadily declined during that period. The same overall pattern of behavior is observed in the early 1960s. Thus, from looking at the raw data, we cannot easily discern the presence

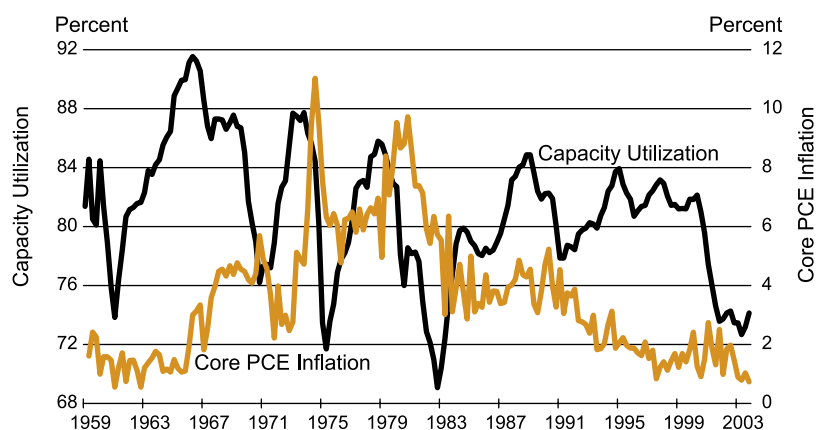
⁵ On the basis of these surveys, the Board staff also makes monthly estimates of capacity by assuming that capacity follows a linear trend within the year.

⁶ For a more thorough discussion of how capacity utilization is constructed, see the articles by Norman Morin and John Stevens, Carol Corrado and Joe Matthey, and Zolton Kenessey.

⁷ To measure capacity utilization, we use the capacity utilization rate in manufacturing. Our measure of inflation is the annualized quarterly change in the price index for personal consumption expenditures less food and energy (core PCE).

FIGURE 1

Core PCE Inflation and Capacity Utilization



Core PCE Inflation is measured as the annualized one-quarter percent change in the core price index for personal consumption expenditures.

Capacity utilization is capacity utilization in manufacturing.

⁴ See the explanatory notes for the Industrial Production and Capacity Utilization G.17 Federal Reserve Statistical Release at www.federalreserve.gov/releases/G17/cap_notes.htm.

of a significant statistical or predictive link between capacity utilization and inflation.

But can we find a more exact relationship by concentrating on the link between capacity utilization and inflation over the business cycle? Capacity utilization is highly cyclical, and it may be that its primary influence on inflation is over the business cycle as well. Our first empirical examination of the link between the capacity utilization rate and inflation is to look at their correlations once we have removed both the trends and the very short-term noise in the series (Figure 2).⁸ As seen in the figure, current capacity utilization is highly positively

⁸ To do this, we first used a band-pass filter to filter out long-run and very short-term components of the two series. We then computed the correlation between the two series.

correlated with future inflation, indicating that when capacity utilization is high, inflation in the future will also be high. Similarly, if capacity utilization is currently low, inflation will be low in the future as well. The current capacity utilization rate shows its highest correlation with inflation five quarters in the future. Thus, over the business cycle, it looks like capacity utilization rates lead inflation.

A SKETCH OF SOME THEORIES

Effects of Increases in Demand Induced by Monetary Policy. The clearest early exposition of the relationship between the intensity with which resources are used in production and changes in the price level is provided in John Maynard Keynes' *General Theory of Employment, Interest, and Money*. In his treatise, Keynes postulated that the price level was tied directly to the cost of production and

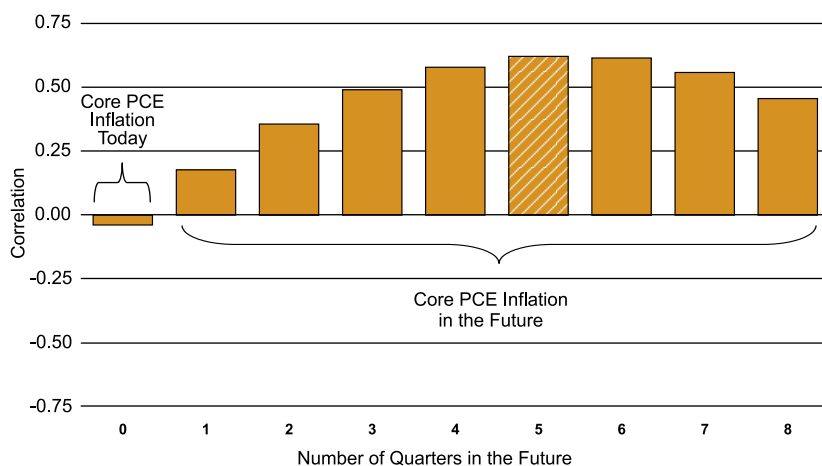
that production costs, in turn, were linked to the intensity with which factors of production — labor and capital — were used. For example, if employment was well below full employment, Keynes assumed that a monetary-policy-induced increase in aggregate demand would not cause an increase in wages. Additional labor would be readily supplied at the going wage rate. As a result, the cost of producing more output did not require any increase in prices. Thus, when employment was below full employment, monetary policy could stimulate output with very little increase in the price level — that is, the general level of prices in the economy.

He also considered how intensively capital was being used when thinking about how much prices would need to adjust when demand increased. He postulated that all factors of production would generally not reach their full employment levels simultaneously, nor would all industries simultaneously reach full production. As demand increased, more and more industries would find themselves at full employment, and any further increase in demand would merely cause an increase in the prices they charged. Thus, as the economy as a whole got closer to fully employing labor and capital, prices would increase at an accelerated pace as aggregate demand increased. In other words, higher levels of capacity utilization would imply an increasingly higher price level.

Although the original theory was postulated as a relationship between the price level and utilization, the modern view links inflation with utilization. This theory suggests that prices increase at a faster rate when utilization rates are high and that we should, therefore, see a stronger relationship between inflation and utilization when utilization rates are high. Importantly, the rate of utilization will influence the inflationary consequences of monetary

FIGURE 2

Business-Cycle Correlations Between Capacity Utilization Today and Core PCE Inflation Today and in the Future



A correlation of one indicates that the series move together perfectly, while a correlation of zero indicates that the two series are unrelated. A correlation of minus one indicates that the series moves in opposite directions perfectly.

policy. For example, accommodative policy might be more inflationary when capacity utilization is high.

Long-Run Implications. Keynes' theory, like many modern macroeconomic theories, implies that monetary policy can affect economic activity in the short run. However, unlike any respectable modern theory, his theory also implied that output was affected in the long run as well. An increase in output back to its capacity level, which was caused by a monetary-policy-induced increase in demand, was permanent. In modern models, monetary policy's only long-run effect is on prices.

Thus, according to the modern view, an increase in demand induced by monetary policy will initially cause output and utilization rates to rise. But as time passes, prices will begin to adjust and inflation will increase. As a consequence of rising prices, output and utilization rates will fall back to their initial levels. In this case, inflation and utilization rates might be negatively correlated, depending on the specific path of inflation and utilization. For example, typically, in response to expansionary monetary policy, inflation rises quite slowly at first, then picks up steam, and finally reverts to its average rate. Measured capacity utilization, on the other hand, rises quite quickly and declines much more quickly than inflation. Thus, along part of their joint trajectory — when inflation is still rising but capacity utilization rates have already begun to decline — the two series are negatively correlated.⁹ The dynamic relationship between these two variables is entirely missing from the basic Keynesian theory.

Including the Effects of Other

⁹ The description of the behavior of capacity utilization and inflation is based on the empirical work of David Altig, Lawrence Christiano, Martin Eichenbaum, and Jesper Linde.

Types of Shocks. Up to this point, we have focused on changes in demand primarily induced by monetary policy. However, changes in monetary policy account for only a part of the disturbances that affect economic activity. Changes in productivity (i.e., the output produced by an hour of work) are also a primary source of economic fluctuations, and the early Keynesian theory offers little in the way of understanding how changes in productivity affect both utilization rates and inflation. Increases in productivity lead to increases in output, but they also lead to an increase in the level of capacity;¹⁰ that is, the economy is simply capable of producing more goods. So,

The relationship between inflation and changes in capacity utilization brought about by changes in productivity could vary over time, depending on how monetary policy responds to the increase in productivity.

at first glance, productivity's effect on capacity utilization is ambiguous.

But it takes time for firms to add new capacity. Initially, firms will use their more productive workers more intensively, thereby increasing output. Thus, in the short run, increases in productivity should lead to increases in capacity utilization. In the long run, additional capital will be built up through increased investment, and capacity output and actual output will move one-for-one.

Thus, increases in productivity can lead to a short-run increase in capacity utilization. However, it is the way in which monetary policy reacts to the increase in productivity that determines whether the increase in utilization will be associated with an increase

¹⁰ This effect would be picked up in the Federal Reserve's survey-based measure of capacity.

or decrease in inflation.¹¹ Therefore, the relationship between inflation and changes in capacity utilization brought about by changes in productivity could vary over time, depending on how monetary policy responds to the increase in productivity.

CONFRONTING THE THEORY WITH THE DATA

The preceding discussion suggests that inflation could be influenced by capacity utilization rates, but at the same time, it indicated that the relationship might not be very exact. The simple Keynesian theory suggested a strong relationship between changes in capacity utilization and inflation when

these changes were demand driven, while long-run considerations and the consideration of other types of disturbances indicated that the link might not be very strong at all.

To shed light on the theoretical uncertainty, we now explore the statistical relationship between capacity utilization and inflation along a number of dimensions.¹² First, how well does capacity utilization predict inflation?

¹¹ For a more complete explanation of the role monetary policy plays in how productivity improvements affect the economy, see Mike Dotsey's previous *Business Review* article.

¹² We investigate a particular measure of inflation, inflation in the core PCE; a particular measure of resource utilization, the capacity utilization rate; and a particular simple specification of the relationship between the two, one that doesn't include other variables that might influence the relationship, e.g., the unemployment rate or productivity growth. A more thorough analysis would include more complicated specifications and other measures of inflation and resource utilization.

In the simple theories outlined above, it is possible that utilization will begin to change before inflation changes, and we wish to see if we can confirm this behavior. So we will test whether the past and current behavior of utilization rates helps predict future rates of inflation. Second, when the capacity utilization rate is low, some theories predict that inflation may not be very responsive to an increase in demand. At the same time, when utilization rates are high, inflation will be very responsive to demand. Thus, utilization's effect on inflation may vary with the level of utilization, and we will test to see if this is the case as well.

In particular, we want to see if utilization rates can tell us anything more about the behavior of inflation than we could learn just by looking at the behavior of inflation itself.¹³ For instance, our look at simple correlations indicated that past utilization rates are positively correlated with current inflation. We would like to know, however, if utilization rates help to predict future inflation over the period 1959-2003, taking into account the behavior of current and past inflation.

To test whether capacity utilization aids our ability to predict core PCE inflation over and above what we could have done by just using inflation itself, we ran two regressions: a regression of average inflation over the past year on a constant, past capacity utilization, and on past quarterly inflation rates, and a regression of average inflation over the past year on past quarterly inflation rates alone (see *Empirical Specification*).

The top panel of Figure 3 shows the actual year-over-year inflation rates (blue line) for the core PCE and

Empirical Specification

Our basic regression is

$$100[P(t) - P(t-4)] = a + b_0*[400(P(t-4)-P(t-5))] + b_1*[400(P(t-5)-P(t-6))] + \dots + b_n*[400(P(t-4-n)-P(t-5-n))] + c_0*CU(t-4) + \dots + c_m*CU(t-4-m) + e(t),$$

where $P(t)$ is the log of the quarterly average of the monthly chain-weighted price index for core personal consumption expenditures at time t and $CU(t)$ is the rate of capacity utilization in manufacturing at time t . The number of lags was chosen by minimizing the Bayesian information criteria, and standard errors are corrected for heteroscedasticity and serial correlation using the methodology of Newey and West. For the sample period covering 1959:Q1 to 2003:Q4, our Granger-causality results are based on the parameter estimates in the table below. The coefficient, c_0 , on capacity utilization is significant at the 1 percent level, indicating that capacity utilization helps forecast core PCE inflation over the entire sample.

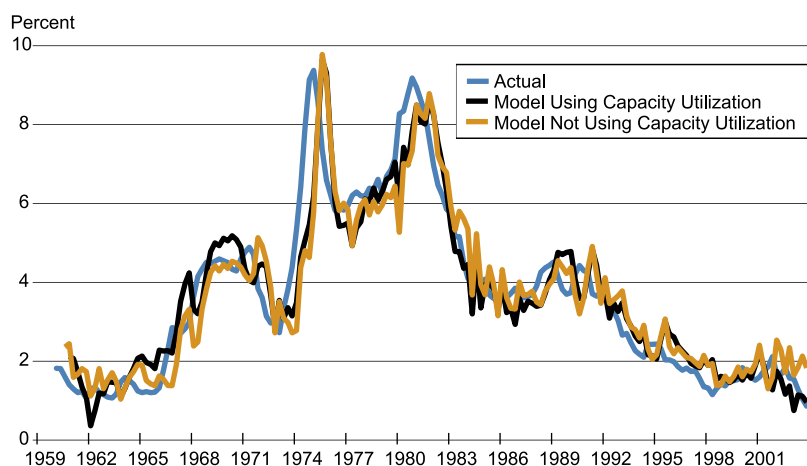
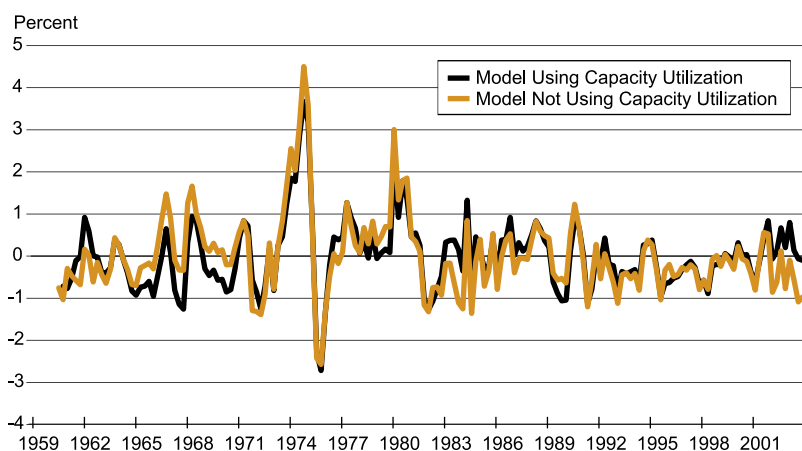
Coefficient	Estimate	HAC Standard Error
a	-8.455	2.032
b_0	0.516	0.074
b_1	0.219	0.105
b_2	0.200	0.069
c_0	0.107	0.025
R^2	0.85	
SEE	0.84	

the predicted values of inflation from the two regressions. The predictive values that use capacity utilization are shown by the black line and those that use only past inflation are shown by the orange line. For our entire sample period covering 1959 to 2003, we find that past rates of capacity utilization are statistically significant — that is, they help predict future inflation — but that their effect on the actual forecast is quite small.¹⁴ The predictions of inflation do not appear to be very different whether we include capacity

utilization or not — the orange line tracks the blue line about as well as the black line does. This is seen more clearly in the bottom panel when we look at the difference between the predicted values and actual values (called forecast errors). The average absolute value of the forecast error falls from 0.66 percent when capacity utilization is not included to 0.60 percent when capacity utilization is included. Moreover, the ability of capacity utilization to forecast inflation has fallen over time. Over the period 1984-2003 our estimations indicate that capacity utilization no longer statistically helps predict inflation. This result is consistent with the graphs in Figure 1, which suggest that the relationship between capacity utilization and inflation is less strong over the latter half of the sample period. For example,

¹³ The statistical name for this procedure is a Granger causality test. In all of the regressions, we chose the number of lags that gave the best specification as determined by that which minimized the Bayesian information criterion.

¹⁴ Specifically, our results are significant at the 1 percent level. A 1-percentage-point increase in the utilization rate leads to an increase in yearly inflation of only 0.107 percentage point. These results are consistent with those reported in the paper by Stephen Cecchetti and the one by Kenneth Emery and Chih-Ping Chang.

FIGURE 3**Actual and Predicted Core PCE Inflation:
In-Sample****Difference Between Actual Value and Predicted
Value of Core PCE Inflation**

The mean absolute error is 0.66 percent in the model not using capacity utilization and 0.60 percent in the model using capacity utilization.

capacity utilization rates are moving up throughout most of the 1990s while core PCE inflation is falling.¹⁵

¹⁵ The vanishing predictive content of utilization found here matches results reported in Emery and Chang (1997). This means that over the later sample, past capacity utilization has no statistically significant independent effect on inflation other than its possible effect on past inflation rates themselves.

Explaining the Empirical Findings. Why might the relationship be significant in some periods and not in others? One possible explanation may be related to the different types of shocks that have hit the economy over the sample period and the different responses that utilization and inflation have to these shocks.

Another explanation revolves

around the changing nature of monetary policy itself. Recall that the theoretical link between capacity utilization and inflation is most precise when the predominant economic disturbances are shocks to demand brought about by changes in monetary policy. Expansionary monetary policy in the presence of economic slack leads to increases in output with little upward pressure on inflation. During times when labor and capital markets are tight, it leads mostly to rising prices and inflation.

With respect to productivity disturbances, the implications are less clear. Depending on how monetary policy reacts, there could be little relationship between utilization and inflation. Indeed, recent theoretical work indicates that it is optimal for monetary policy to insulate the price level and inflation from productivity disturbances.¹⁶ Doing so maximizes the economy's ability to react efficiently to changes in productivity. If we look at the data over the 1990s, monetary policy appears to have done that. So if much of the economic activity in the 1990s was driven by changes in productivity, and if the central bank was operating in an optimal manner, we would not expect to see a strong link between inflation and capacity utilization rates over this sample period.

Does Utilization's Effect Vary with Its Level? Another reason that capacity utilization's effect on inflation might vary over time is that its effect may depend on its level. This would be the case if, as suggested by basic Keynesian theory, the weakest link between capacity utilization and inflation occurred at very low utilization rates,

¹⁶ The intuition for this result is discussed more fully in Mike Dotsey's previous *Business Review* article. More detailed theoretical analysis can be found in the papers by Robert King and Alexander Wolman; Aubhik Khan, Robert King, and Alexander Wolman; and Michael Woodford.

while the strongest link occurred at very high utilization rates. For the former, we would expect that when utilization was below some threshold, utilization rates would rise with no change in inflation. For the latter, we would expect that when utilization rates were above some threshold, changes in aggregate demand would bring about big changes in inflation.

To test this implication, we ran a regression where we separately considered the effects of very high utilization rates, average utilization rates, and very low utilization rates.¹⁷ We found that the relationship between utilization rates and core PCE inflation does not vary with the level of utilization. This result rejects one of the implications of the Keynesian theory¹⁸ and indicates that, in our specification, changes in utilization, whether starting from a level of slack or a level of tightness, imply the same future effect on core PCE inflation, namely, a 1-percentage-point increase in manufacturing capacity utilization implies a 0.107-percentage-point increase in core PCE inflation.

¹⁷ We do this by dividing the utilization rates into three roughly equal portions: u-low, u-middle, and u-high. For a normally distributed variable the boundaries determining u-middle are the mean of $u \pm 0.43$ times the standard deviation of u . Thus, the groups are formed by defining $u\text{-low} = u$ if u is less than the mean of u minus 0.43 times the standard deviation of u and zero otherwise. Similarly, $u\text{-high} = u$ if u is greater than the mean of u plus 0.43 times the standard deviation of u and zero otherwise. $U\text{-middle} = u$ if u falls in between these two bounds and zero otherwise. We find that it works well and that it approximately divides the utilization series into three equally represented orthogonal components. We computed 56 nonzero observations that fall into the u-high category, 60 in the u-low category, and 64 in the u-middle category for the period 1959 to 2003. The mean of the nonzero observations falling into u-high is 86.05, 76.08 for u-low, and 81.32 for u-middle.

¹⁸ These results are consistent with those reported in Mary Finn's 1995 article. Finn uses a slightly different specification over a different sample period.

FORECASTING USING ONLY SOME OF THE AVAILABLE DATA

If a policymaker were to rely on the relationship between capacity utilization and inflation when setting policy, he could only use available data. A policymaker in 1983 would have had no knowledge of the statistical relationship between these two variables in the 1990s because that data had not yet been generated. Further, it is not clear that the policymaker would even want to use all the data available to him at the moment. We just discussed

Over some periods including utilization helps to predict core PCE inflation, but at other times, including it actually makes the forecasts worse.

our analysis of the statistical relationship between capacity utilization and core PCE inflation over the entire sample period, which is the correct procedure if the statistical relationship is stable. However, the relationship may not be stable, implying that it is different in different periods. For example, if the relationship between capacity utilization and inflation differs between the 1960s and the 1980s, we would not want to use data from the 1960s to help us predict inflation in the 1980s. To address this issue, we would need to look at so-called out-of-sample prediction, that is, predicting future inflation at any point in time by using only data that were available at that time, and perhaps only some portion of the available data.¹⁹

¹⁹ We do, however, use final revised data rather than real-time data in this exercise.

Our statistical analysis (discussed in *The Changing Relationship Between Inflation and Utilization Rates*) suggests that the relationship between core PCE inflation and capacity utilization is not stable, implying that additional tests for analyzing whether capacity utilization helps predict inflation are required. Therefore, we re-estimated our model using only the most recent 60 quarters of data, starting from the first quarter of 1961 through the fourth quarter of 1975, and then successively updating our 60-quarter sample. For example, the prediction of inflation for 1983 is based on data over the sample 1968-1982.

Figure 4 is similar to Figure 3 in showing both the predicted inflation from these rolling regressions when capacity utilization is either included or excluded and the resulting forecast errors of the two specifications. Our results indicate that up to about 1990, it matters whether utilization rates are included. Over some periods — for example, during the early 1980s — including utilization helps to predict core PCE inflation, but at other times, such as the late 1980s, including it actually makes the forecasts worse. The forecast errors actually become larger when capacity utilization is included. Over the entire period, we find virtually no difference in forecast accuracy. As the sample progresses, capacity utilization neither hurts nor helps our ability to forecast core PCE inflation, reflecting the fact that over the past 13 years, capacity utilization has not proven very useful for forecasting core PCE inflation.²⁰

²⁰ The waning usefulness of capacity utilization as a predictor of core PCE inflation is consistent with recent work by Stephen Cecchetti, Rita Chu, and Charles Steindel. However, James Stock and Mark Watson find that capacity utilization continues to help predict inflation over the period 1984-1996 using a recursive forecasting method. Because we find some evidence of parameter instability, we used the alternative procedure of rolling regressions.

The Changing Relationship Between Core PCE Inflation and Capacity Utilization Rates

A

monetary policy-maker who wanted to formulate policy relying on the relationship between capacity utilization

and inflation would need to know if that relationship would continue to hold. But how stable is the empirical relationship between capacity utilization and inflation?

To explore the stability of the relationship between capacity utilization rates and core PCE inflation, we looked at the behavior of the estimated regression coefficients over time. To do this, we ran a number of regressions, each on 60 quarters of data. We started with a sample period beginning in the first quarter of 1961 and ending in the fourth quarter of 1975 and then updated the starting and ending dates by one quarter. Our last regression covered the period from the first quarter of 1989 through the fourth quarter of 2003. For each of these rolling regressions, the top and bottom panels of the figure show the coefficients on the first lag of inflation and the first lag of capacity utilization as well as the 95 percent confidence interval for each of the coefficient estimates. These confidence intervals indicate that the true value of the coefficient lies within the range with 95 percent probability. When the interval includes zero, the coefficient is not statistically different from zero.

It is easy to see that the coefficients describing the behavior of core PCE inflation (i.e., the coefficients on $(P_{t-4} - P_{t-5})$ and CU_{t-4}) are changing over time. The coefficient on capacity utilization is positive and generally significantly different from zero over the

early part of the sample. As time goes forward, however, it becomes insignificantly different from zero. This experiment gives further credence to the assertion that the relationship between capacity utilization rates and inflation has changed over time.

FIGURE

Rolling Coefficient Estimates for Core Inflation

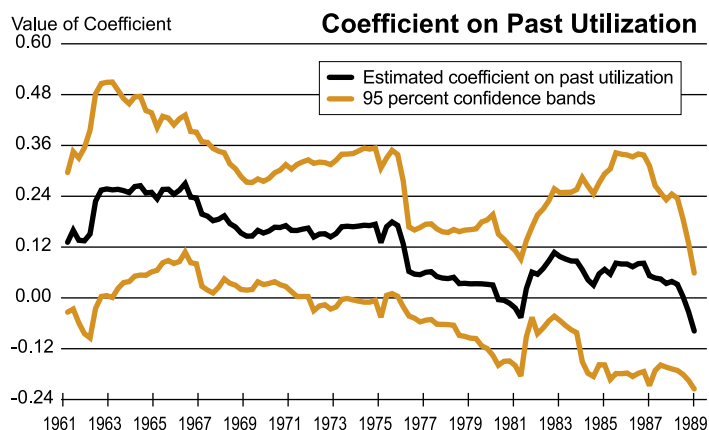
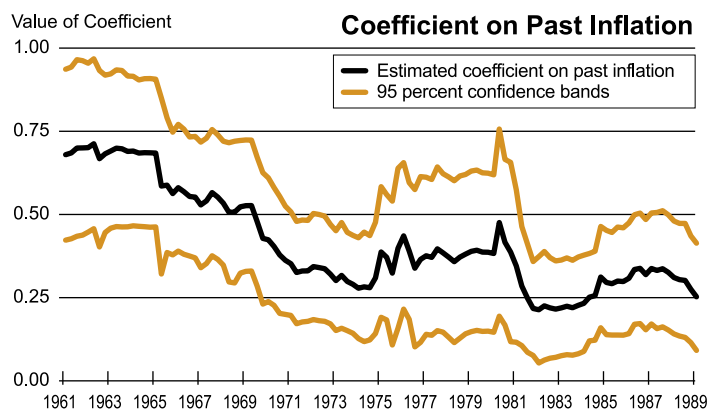
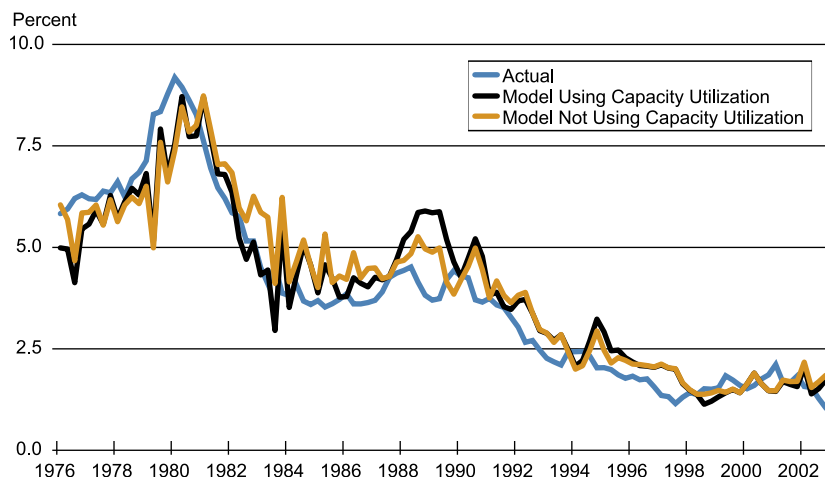
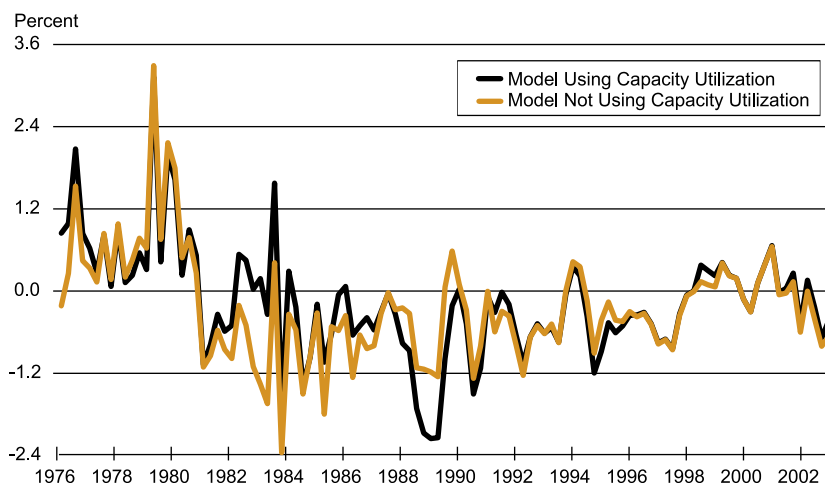


FIGURE 4

Actual and Predicted Core PCE Inflation: Out-of Sample



Difference Between Actual Value and Predicted Value of Core PCE Inflation



CONCLUSION


Various theories suggest that the intensity of resource use could be an important determinant of inflation. At first glance, it appeared that an economy with lots of spare capacity was less likely to experience an increase in inflation than one that was fully employing all of its resources.

However, the theories describing the causal relationship between utilization and inflation are not

universally accepted, and it is quite possible that both inflation and capacity utilization are driven by more fundamental factors, such as changes in productivity or monetary policy. Moreover, the relationship between utilization and inflation could be sensitive to which fundamental factor is driving the economy and the way in which monetary policy responds to those fundamentals, making the relationship quite

complex and conditional on economic circumstances. Therefore, drawing inferences about how capacity utilization will affect inflation is a bit tricky. It depends on both the types of shocks hitting the economy and the central bank's response to those shocks. Thus, the joint behavior of utilization and inflation could vary over time for a number of reasons.

Our empirical investigation of one specification of the statistical relationship between capacity utilization and core PCE inflation suggests that the relationship is not robust. Over different sample periods, capacity utilization's ability to help explain or predict the behavior of core PCE inflation varies quite a bit. Sometimes utilization rates are modestly useful, and at other times, especially over the past 15 years or so, they have been unhelpful.

This lack of robustness could be due to changing policy responses to productivity shocks. A well-run monetary policy will allow changes in productivity to influence economic activity without changing inflation. If changes in productivity have been the prevailing driving force behind the economic activity of the last 15 years, and if monetary policy has been conducted in an optimal manner,²¹ changes in utilization should not be correlated with changes in inflation. That evidence would not necessarily imply that in response to some other type of economic disturbance, the utilization rate would be uninformative about the likely path of inflation. But our empirical results, using linear forecasting equations, suggest that one should be cautious in predicting core PCE inflation using a simple model of capacity utilization rates. 

²¹ See Mike Dotsey's previous *Business Review* article for suggestive evidence that this has indeed been the case.

REFERENCES

Altig, David, Lawrence J. Christiano, Martin Eichenbaum, and Jesper Linde. "The Operating Characteristics of Alternative Monetary Policy Rules," paper, November 2002.

Cecchetti, Stephen G. "Inflation Indicators and Inflation Policy," *NBER Macroeconomics Annual* 1995, pp. 189-219.

Cecchetti, Stephen G., Rita S. Chu, and Charles Steindel. "The Unreliability of Inflation Indicators," *Federal Reserve Bank of New York Current Issues in Economics and Finance*, 6, 4, April 2000.

Corrado, Carol, and Joe Matthey. "Capacity Utilization," *Journal of Economic Perspectives*, 11, 1, Winter 1997, pp. 151-67.

Dotsey, Michael. "How the Fed Affects the Economy: A Look at Systematic Monetary Policy," *Federal Reserve Bank of Philadelphia Business Review*, First Quarter 2004, pp. 6-15.

Emery, Kenneth M., and Chih-Ping Chang. "Is There a Stable Relationship Between Capacity Utilization and Inflation?" *Federal Reserve Bank of Dallas Economic Review*, First Quarter 1997, pp. 14-20.

Finn, Mary G. "Is 'High' Capacity Utilization Inflationary?" *Federal Reserve Bank of Richmond Economic Quarterly*, 81, 1, Winter 1995, pp. 1-16.

Finn, Mary G. "A Theory of the Capacity Utilization/Inflation Relationship," *Federal Reserve Bank of Richmond Economic Quarterly*, 82, 3, Summer 1996, pp. 67-80.

Kenessey, Zoltan E. "Industrial Production and Capacity Utilization," in Roy H. Webb (ed.), *Macroeconomic Data: A User's Guide*. Federal Reserve Bank of Richmond, December 1991, pp. 14-16.

Keynes, John Maynard. *The General Theory of Employment, Interest, and Money*. (New York: Harcourt, Brace and Company), 1936.

Khan, Aubhik, Robert G. King, and Alexander L. Wolman. "Optimal Monetary Policy," *Review of Economic Studies*, October 2003, pp. 825-60.

King, Robert G., and Alexander L. Wolman. "What Should the Monetary Authority Do When Prices Are Sticky?" in John B. Taylor (ed.), *Monetary Policy Rules*. (Chicago: University of Chicago Press), 1999, pp. 349-98.

Morin, Norman, and John Stevens. "Estimating Capacity Utilization from Survey Data," Working Paper, Finance and Economic Discussion Series 2004-49, Federal Reserve Board, 2004.

Staiger, Douglas, James H. Stock, and Mark W. Watson. "The NAIRU, Unemployment, and Monetary Policy," *Journal of Economic Perspectives*, 11, 1, Winter 1997, pp. 33-50.

Stock, James H., and Mark W. Watson. "Forecasting Inflation," *Journal of Monetary Economics*, 44, 2, October 1999, pp. 293-335.

Woodford, Michael. *Interest and Prices: Foundations of a Theory of Monetary Policy*. (Princeton: Princeton University Press), 2003.

International Risk-Sharing: Globalization Is Weaker Than You Think

BY SYLVAIN LEDUC

W

ith the development of international financial markets, households should be better equipped to pool their resources so that their level of consumption varies less from year to year. Yet the extent of international risk-sharing remains surprisingly small. In this article, Sylvain Leduc digs a little further into the data to uncover why, in spite of recent trends, financial globalization remains weaker than you think.

From 1980 to 2004, world trade in goods and services increased from 36 percent to 50 percent of world GDP. As the world experienced a surge in the trade of goods and services, it also saw a substantial rise in the trade of financial assets. The share of foreign equities in U.S. investors' portfolios, for instance, increased from about 1 percent in the early 1980s to 12 percent in 2000.¹ On that dimension, the impression that we are living in a more integrated world is borne out in the data. But if we dig in a little further,

¹ See Francis Warnock's article.



Sylvain Leduc is an economist with the Board of Governors. When he wrote this article, he was a senior economist at the Philadelphia Fed.

we will find that, notwithstanding the trend toward globalization, the world's economies remain strikingly insular along many dimensions.

With the developments of international financial markets, households should be better equipped to diversify their portfolios and protect their investments against unforeseen events, which ultimately should result in more sharing of consumption risk across countries. That is, households would effectively pool their resources so that their level of consumption varies less from year to year. Yet, the extent of international risk-sharing remains surprisingly small and is one key reason that globalization is weaker than you think.

Standard macroeconomic models offer predictions regarding the extent of international risk-sharing. If consumers are diversifying internationally, we should see consumers in one country consuming more than those in another country when the price of doing so is lower than in the other

country. This relative price is the real exchange rate, that is, the exchange rate between the countries' currencies adjusted for the rate of inflation in the two countries. One reason for the lack of international risk-sharing is that, empirically, real exchange rates often move in a way that hinders the risk-sharing process. As a result, full globalization remains far away, at least along this important dimension.

INTRODUCING RISK-SHARING

At the base of the concept of risk-sharing is the idea that most people would prefer to keep a relatively stable pattern of consumption instead of a highly variable one. The challenge is to achieve this smooth consumption pattern even though income may vary a lot from year to year. For instance, many workers are, at times, temporarily laid off because of a slowdown in their particular line of business. Or people may have to temporarily quit their jobs for health reasons. Depending on the frequency of such events, incomes can vary quite a bit in any given year.

If households do not save or borrow, their level of consumption will follow their variable level of income. For instance, imagine a simple economy composed of two households, the Greens and the Verdis, that have fluctuating incomes from year to year.² Suppose we look at how much money these households made over the last two years and we find that the Greens had an after-tax income of \$10,000 in year 1 and \$30,000 in year 2. For

² See also Keith Sill's *Business Review* article for a discussion of risk-sharing.

simplicity, imagine that the opposite is true for the Verdis: in year 1, the Verdi household took home \$30,000, while it earned \$10,000 in year 2.

First, to keep the argument simple, assume that both households use their income to consume the same basket of goods and that they pay the same price for one unit of those goods, \$1. This is an important assumption that I will relax in the next section. If the households do not save or borrow, their level of consumption will follow their level of income. That is, in year 1 the Greens consume 10,000 units of goods and the Verdis 30,000 units of goods, and vice versa in year 2.

How could the Greens and the Verdis achieve a relatively more stable consumption pattern? It could be simply achieved if we let the households pool their income each year and divide the total equally between them. Both households could therefore keep a constant consumption level of 20,000 units of goods per year. Notice that, in this example, one implication is that risk-sharing equalizes consumption across the two households. That is, by pooling their resources, households are able to “share” the risks of their fluctuating incomes and therefore eliminate or “insure” against their consumption risk.

However, it might be quite difficult to find another household that will agree to pool its income with yours. In practice, this risk-sharing process is instead carried out through financial markets. For instance, households can save by buying stocks of firms or government bonds when their income is unexpectedly high, or they can buy goods with credit when their income is unexpectedly low and repay their debt in more prosperous times. Through borrowing and lending in financial markets, households can smooth out the bumps in their income streams and

achieve a more stable consumption pattern. As long as households keep a well-diversified investment portfolio, they are better equipped to smooth out their consumption risk. Indeed, one of the tenets of modern finance is that households should hold a well-diversified investment portfolio so that the portfolio’s overall risk is less subject to the vagaries of one particular sector or one particular stock.

In the above example, note that I did not mention the country of residence of the two households. In fact, the argument does not depend on the households’ locations. As long as household incomes do not move in the same direction — up or down — at the same time, there is scope for sharing consumption risk, be it within or between nations. Since world economies are not always in sync, and some countries fall into recession while others continue to expand, household incomes in different countries do not

both the Greens and the Verdis to be temporarily laid off at the same time. In this case, there is no scope for mutually beneficial trade by which to insure against consumption risk.

Global risks will necessarily trigger movements in consumption. But every household’s consumption will be moving in the *same* way. Therefore, in a world in which households can use financial markets to insure against all possible *idiosyncratic* risks to their income and in which households consume the same basket of goods and pay the same price for those goods, theory predicts that consumption should move in the same direction across countries.

INTERNATIONAL RISK-SHARING AND RELATIVE PRICES

Obviously, this prediction is derived under relatively strong conditions. For instance, it is unlikely that households consume the same basket of goods and services. There

As long as household incomes do not move in the same direction — up or down — at the same time, there is scope for sharing consumption risk, be it within or between nations.

always move together. So there is potential for sharing consumption risk across countries.

However, households cannot insure against every type of risk. For instance, global risk (as opposed to idiosyncratic risk) is not insurable, since it affects everyone in the same manner, at the same time.³ In terms of our previous example, global risk could include a recession that leads

is also ample evidence that different consumers do not pay the same price for the same goods, especially when these consumers live in different countries (see *Where You Are Affects How Much You Pay*). Once we relax those assumptions, we obtain a more general prediction about sharing consumption risk. In this case, efficient risk-sharing dictates that the household facing the lower relative price consume more.

To see that, let’s look again at our previous example. Suppose that the Greens’ and the Verdis’ income patterns in year 1 and year 2 continue to

³ Contrary to global risk, which affects everybody in the economy, idiosyncratic risks affect only particular individuals.

Where You Are Affects How Much You Pay

I

n the early 1980s, total trade in goods accounted for 36 percent of world GDP; 23 years later, that ratio surged to 50 percent. The fall in trade barriers, initiated after World War II under the General Accord on Tariffs and Trade (GATT), in large part trig-

gered the rise in the trade of goods. As more goods are traded, you might expect the prices of these goods in different parts of the world to converge. That is, what economists called the *law of one price* would hold: A product would sell for the same price (expressed in the same units of currency) in different locations, absent natural or government-imposed trade barriers.

Imagine that you can freely trade cars between the U.S. and Canada and you notice that a Ford Explorer sells for \$5,000 more in Montreal than in Detroit, once you convert the price of a Ford Explorer from Canadian dollars into U.S. dollars using the exchange rate. A profitable business opportunity, called arbitrage, would be to buy Ford Explorers in Detroit at the cheaper price and sell them in Montreal for a profit of \$5,000. As long as prices (expressed in a common currency) of Ford Explorers differ between these two markets, there is an opportunity for arbitraging the price difference. Obviously, it is not costless to trade goods, since businesses have to pay transportation costs, tariffs, or the costs associated with different regulations in different locations. The presence of these costs will allow prices to differ across locations. However, as long as goods can be freely traded, prices of goods should be equalized across countries. In this case, prices would obey the law of one price.^a

You can arbitrage price differentials not only in markets in different countries but also in markets located in the same country.^b Arbitrage opportunities should tend to equalize prices in

different locations. However, it appears that price differentials are much larger across countries than across locations in a given country. For instance, in a widely cited article, economists Charles Engel and John Rogers documented that prices vary much more between Toronto and New York, say, than between Detroit and New York. This implies that price differentials across countries are not solely the result of transportation costs, since the distance between Toronto and New York is about the same as that between Detroit and New York. Rather, there seems to be something special about crossing borders.

Prices can indeed differ widely across countries.^c Mario Crucini, Chris Telmer, and Mario Zachariadis documented the price differentials for selected traded goods in different European countries. They found that price differentials are indeed large, once prices are converted in common currency units. For instance, they found that Austrians pay twice the amount Belgians pay for one pound of long-grain rice. Washing detergent is twice as expensive in Greece as it is in Germany. And two pounds of coffee is 40 percent cheaper in France than in Italy.

Moreover, it appears that deviations from the law of one price are fairly stable through time. In a National Bureau of Economic Research paper, economists Kenneth Froot, Michael Kim, and Kenneth Rogoff showed that for many commodities (for instance, barley, butter, and silver), the deviations from the law of one price are not just a property of modern economies; they were present as far back as the 13th century.

In a nutshell, the law of one price fails dramatically, and this failure provides another example that globalization is weaker than you think.

^a When the law of one price holds for every good in the economy, exchange rates will be determined according to what economists call purchasing power parity, or PPP. PPP states that nominal exchange rates should move to offset differences in inflation across countries, leaving real exchange rates constant over time. Notice that this simple approach to exchange-rate determination cannot explain the high volatility of real exchange rates.

^b See Leonard Nakamura's *Business Review* article for a discussion of the failure of the law of one price across U.S. retailers and its impact on the measurement of inflation.

^c Kenneth Rogoff's article provides a survey of the large empirical literature documenting the failure of the law of one price.

be the same as before: the Greens have an after-tax income of \$10,000 in year 1 and \$30,000 in year 2. Further suppose that the opposite is true for the Verdis. However, let's now assume that the two households do not pay the same price for the goods. Suppose that in year 1, the Greens continue to pay \$1, but the Verdis now must spend \$2 to obtain the same goods and that the

reverse is true in year 2.

If the households do not pool their resources, the Greens will consume 10,000 units of goods the first year and 15,000 units in the following year, since it must then pay \$2 for the goods. For the same reasons, the Verdis' consumption will fluctuate between 15,000 and 10,000 units between year one and year two. In

this case, the household that faces the cheaper price does not consume more. For instance, even though the Greens pay half the price as the Verdis in year 1, they consume 5,000 fewer units.

By pooling their income (\$40,000 in each year) and dividing the total equally between them (\$20,000 per household in each year), the Greens and the Verdis can take advantage of

the price differentials and achieve a more efficient consumption pattern. In year 1, the Greens would consume twice as much as the Verdis (20,000 versus 10,000 units of goods), since it must pay half the price the Verdis pay for the same goods (\$1 versus \$2). Since, in the second year, the Verdis face a lower price than the Greens (\$2 versus \$1), they will consume more (20,000 versus 10,000 units).

Note that when households face different prices, efficient risk-sharing does not state that consumption should move together across households. Rather, efficient risk-sharing dictates that the household facing the lower relative price should consume more. Intuitively, this criterion makes sense, since the world economy should channel more consumption to places where it is relatively cheap to consume.⁴

Once again, it is immaterial whether these two households live in the same country. The only difference is that when households live in different countries, the relative price of goods has a particular name: the real exchange rate.

⁴ Another way to think about optimal risk-sharing is to think in terms of costs and benefits. Optimal risk-sharing occurs when the benefit of transferring one extra dollar from the Verdis to the Greens (or vice versa) equals the cost. As long as the marginal benefit of the transfer exceeds the marginal cost, it is beneficial to transfer resources from the Greens to the Verdis. For instance, in year 1 the benefit of transferring one extra dollar from the Verdis to the Greens is that the Greens now consume one more unit. However, such a transfer has a cost. To transfer one extra dollar to the Greens, the Verdis have to lower their consumption by half a unit, since the Verdis pay twice as much as the Greens for the same basket of goods. Therefore, the cost of the transfer is the relative price, 2, times 0.5 units of consumption, which is 1 unit of consumption. Therefore, optimal risk-sharing occurs because the marginal benefit of transferring one extra dollar from the Verdis to the Greens exactly equals the marginal cost.

INTERNATIONAL RELATIVE PRICES: REAL EXCHANGE RATES

People usually think about nominal exchange rates, which denote the price of one currency in terms of another. For instance, in the first quarter of 2003, one British pound was worth 1.60 U.S. dollars. One year later, the British pound traded for 1.84 U.S. dollars. Therefore, the U.S. dollar lost 15 percent of its value against the British pound over that year.⁵

The real exchange rate, on the other hand, is the nominal exchange rate multiplied by the ratio of price levels in the two countries, as measured, for instance, by the consumer price index.⁶ A change in the real exchange rate, therefore, represents a change in the relative price of two countries' goods, controlling for inflation.

For instance, in the first quarter of 2004, the consumer price index in the United States was 121.4, and the consumer price index in the U.K. was 179.2, implying a real exchange rate of 2.36: the nominal exchange rate of 1.60 U.S. dollar per British pound times the ratio of U.K. to U.S. price indices. By the first quarter of 2003, however, the U.S. consumer price index had risen to 123.4, while the

⁵ Throughout this article I will denote the exchange rate in foreign currency units, i.e., how many U.S. dollars one unit of foreign currency (in the above example, a British pound) is worth. In this case, an upward movement in the exchange rate implies a depreciation of the U.S. dollar.

⁶ The consumer price index, or CPI, measures the cost of living for a typical urban family. The index shows how the price of a typical basket of goods changes from year to year. So the real exchange rate between the U.K. and the U.S. equals the number of dollars per British pound times the ratio of prices in the U.K. relative to the U.S.: $\left(\frac{\text{dollar price level in the UK}}{\text{pound price level in the US}} \right)$. Again, notice that a rise in the real exchange rate implies a depreciation of the U.S. dollar in real terms.

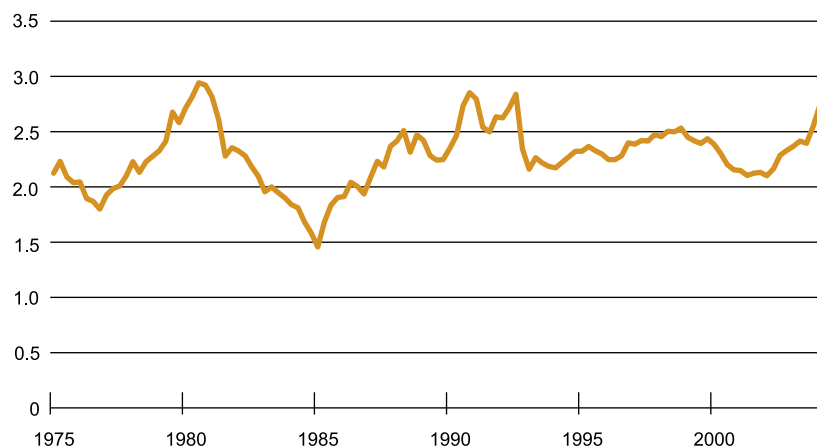
U.K.'s had increased to 183.8; thus, the real exchange rate rose to 2.74. So the real exchange rate increased 16.1 percent from the first quarter of 2003 to the first quarter of 2004. In other words, while \$1 would buy 15 percent fewer pounds in the first quarter of 2004 compared with one year earlier, \$1 of U.S. goods could be traded for 16.1 percent fewer British goods in the first quarter of 2004 than in the first quarter of the previous year.

The real exchange rate is the nominal exchange rate multiplied by the ratio of price levels in the two countries, as measured by the consumer price index.

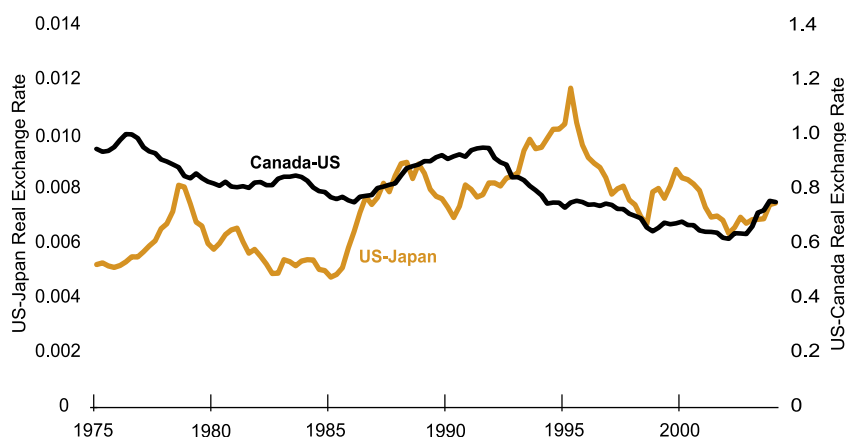
The variations in the U.S.-U.K. real exchange rate between 2003 and 2004 are not unusual. In fact, the real exchange rate has been varying widely over time (Figure 1). Moreover, other currencies, such as the Canadian dollar or the Japanese yen, have experienced similarly large fluctuations (Figure 2). The reasons for those large swings in real exchange rates have intrigued and puzzled international economists for quite a while.

What underlies the large fluctuations in real exchange rates? John Rogers and Michael Jenkins found that the source of movements in real exchange rates is the failure of the law of one price (see *Where You Are Affects What You Pay*).⁷ In fact, they found that 81 percent of the movements in real exchange rates

⁷ Under the law of one price, a good should sell for the same price in different locations, once the prices of the good are expressed in the same currency units and if there are no transport or trade-related costs.

FIGURE 1**U.S.-U.K. Real Exchange Rate**

The real exchange rate is constructed using CPI indices in the U.S. and the U.K. The exchange rates are number of U.S. dollars per unit of British pound.

FIGURE 2**U.S.-Japan and U.S.-Canada Real Exchange Rates**

The real exchange rates were constructed using CPI indices in Canada, Japan, and the U.S. The exchange rates are number of U.S. dollars per unit of Canadian dollar or Japanese yen.

occur because traded goods do not sell for the same price in different countries, once those prices are expressed in common currency units. Using a different methodology, Charles Engel showed that over 95 percent of the variations in real exchange rates are the result of deviations from the law of one price.

As we saw in the previous section, when households do not face the same price for the same goods, risk-sharing has to be modified to take into account the movements in relative prices. For households located in different countries, efficient risk-sharing dictates that consumption should be higher in the country where the rela-

tive price of consumption (that is, the real exchange rate) is lower. In other words, when the U.S. experiences a fall in the price of its consumption basket relative to that in Europe (a *depreciation* of its real exchange rate), it should also be consuming more. However, this does not appear to be the case.

THE LACK OF INTERNATIONAL RISK-SHARING

A simple way to look at the extent of consumption risk-sharing is to look at the correlation between the real exchange rate and the ratio of consumption between different countries. Here we focus on this correlation for the U.S. vis-à-vis other OECD countries (Table).⁸ The correlation captures how these two variables move over time. For instance, a positive correlation implies that when the real exchange rate increases (a depreciation of the U.S. dollar in real terms),⁹ consumption in the U.S. should rise relative to that in the foreign country. (I will call *relative consumption* the movement in U.S. consumption vis-à-vis that of the foreign country.) On the other hand, if the real exchange rate rises as relative consumption falls, the correlation would be negative.

Under efficient risk-sharing, consumption should be higher when its relative price is lower. This implies that the correlation between relative consumption and the real exchange rate should be positive.¹⁰ When the

⁸ The Organization for Economic Cooperation and Development (OECD) is a group of 30 countries that share a commitment to democratic government and the market economy.

⁹ Remember that the exchange rates are U.S. dollars per unit of foreign currency, so that an increase in the real exchange rate implies a fall in the relative value of the dollar in real terms.

¹⁰ It can be shown that, under certain conditions, the correlation between the real exchange rate and relative consumption should be exactly one.

TABLE

Correlations Between Real Exchange Rates and Relative Consumption*

Country	Correlation with U.S.
Australia	-0.01
Austria	-0.35
Belgium	-0.12
Canada	-0.41
Denmark	-0.16
E.U.	-0.30
Finland	-0.27
France	-0.18
Germany	-0.27
Italy	-0.26
Japan	0.09
South Korea	-0.73
Mexico	-0.73
Netherlands	-0.41
New Zealand	-0.25
Portugal	-0.56
Sweden	-0.52
Spain	-0.60
Switzerland	0.16
Turkey	-0.31
U.K.	-0.47

* Consumption and real exchange rate data are annual series from the OECD Main Economic Indicators data set, from 1973 to 2001.

real exchange rate increases, which implies a fall in the relative value of the dollar in real terms, consumption in the U.S. should be higher than it is abroad. The correlations reported in the table demonstrate that there is little consumption risk-sharing among the OECD countries. In fact, all of the correlations are negative, which means that consumption is higher in

the country in which the relative price of consumption is higher — the exact opposite of what efficient sharing of consumption risk predicts. Therefore, sharing of consumption risk across the different countries of the world remains small, even though over the last several decades the world has become seemingly much more integrated.

What underlies the lack of international consumption risk-sharing across countries? One reason is obviously that investors fail to hold a well-diversified portfolio. Indeed, a large literature has documented the puzzling fact that most investors hold a disproportionate share of assets of their country of residence in their portfolio, yet another sign that globalization is weaker than you think. In other words, U.S. investors hold mostly U.S. assets, while French investors' portfolios are mainly composed of French assets. For instance, Francis Warnock, an economist at the Federal Reserve Board, reports that, in 2000, the share of foreign equities in U.S. investors' equity portfolios was about 12 percent, a substantial increase from the 1 percent share in the early 1980s. Yet, U.S. investors remain far from being well diversified: Warnock estimates that, in 2000, a well-diversified U.S. portfolio would have roughly 50 percent in foreign equities. As a result, U.S. investors are exposed to specific risk originating in the U.S., for instance, a recession in the U.S. economy. To the extent that country-specific risks are not perfectly positively correlated across countries, investors could lower the risk of their portfolios by holding stocks of different countries' firms. Trying to understand why investors do not do so remains a very active area of research. Yet, even given that investors' portfolios are not well diversified, it remains puzzling that a country's consumption is higher when its exchange rate is high relative to that of other countries.

REAL EXCHANGE RATES AND RISK-SHARING

We have seen that real exchange rates exhibit large fluctuations, sometimes gaining 10 percent to 20 percent in value in a couple of years, followed by equivalent or larger losses in value. In fact, like any other prices in the economy, real exchange rates react to changes in demand and supply conditions, which can be affected by a variety of fundamental factors such as monetary and fiscal policy or technological innovations. In a recent paper, Giancarlo Corsetti, Luca Dedola, and

Under efficient risk-sharing, consumption should be higher when its relative price is lower.

I documented one reason behind the lack of risk-sharing: Real exchange rates often move in a way that hinders risk-sharing in response to technological changes (Table).

Theory predicts that as a country becomes more productive because of an improvement in technology, it should produce and consume more goods relative to other countries, and it should also experience a depreciation of its real exchange rate, i.e., the price of its goods (in real terms) relative to that in the other country should fall. With an improvement in technology, a country can produce more goods for a given level of inputs, such as the number of workers or machines in the economy. As the supply of goods increases, prices fall. Remember that the real exchange rate is the relative price of goods across countries. As the prices of the goods a country produces fall, the real exchange rate, in general, depreciates.¹¹ Moreover, as a country becomes more productive, it also becomes richer, and its level of consumption should therefore rise

relative to the level of consumption in the rest of the world. Notice, once again, theory predicts that following a technological improvement, a country's consumption should be higher when its real exchange rate is lower. But are these predictions consistent with the data?

To verify whether improvements in technology affect economies as theory predicts, we conducted an analysis based on an empirical model, a simple vector autoregression (VAR). A VAR is a system of linear equations that link different variables together over time. For instance, a VAR with two variables — let's say the real exchange rate and consumption — would also have two equations. One equation would try to explain the movements in the real exchange rate; the other would try to explain the movements in consumption. To do so, both equations would use previous values of the real exchange rate and consumption.

Our VAR included five variables: labor productivity, real GDP, real consumption, net exports, and the real exchange rate.¹² We used a rise in U.S. labor productivity vis-à-vis the


rest of the OECD countries as a proxy for technological improvement in the U.S.¹³ Using our model, we estimated the effect that a sudden increase in the rate of U.S. technological progress would have on the U.S. and foreign economies. We did that by determining the impact that the change in labor productivity would have on the other variables in our statistical model.

We can chart the responses of the variables in our model to a one-time, unanticipated increase in the growth rate of labor productivity (Figure 3). The dotted line represents the estimated response of the variable to the sudden change in labor productivity; the grey area around the dotted line tells us how much confidence we can place in this estimate. In particular, when the entire area is above zero or below zero, we can say with a 90 percent level of confidence that the estimated response of, say, the real exchange rate to the unanticipated jump in productivity is significantly different from zero — that is, the unanticipated jump has an impact on the variable.

For instance, following the jump in labor productivity, the growth rate of output in the U.S. increases relative to the rest of the OECD countries. The rise in productivity is also accompanied by a rise in relative real GDP and consumption growth. These effects are the standard ones predicted by theory. However, contrary

to what theory predicts, the U.S. real exchange rate *appreciates* following an improvement in productivity (that is, the real exchange rate falls), which implies, once again, that consumption is higher when its price is higher. The appreciation of the real exchange rate hinders risk-sharing. As the real exchange rate appreciates, foreign countries can consume fewer imported products, a situation that makes it more difficult for the foreign country to sustain its level of consumption. This is reflected in the fact that net exports of U.S. goods fall following an increase in labor productivity.¹⁴

SUMMARY

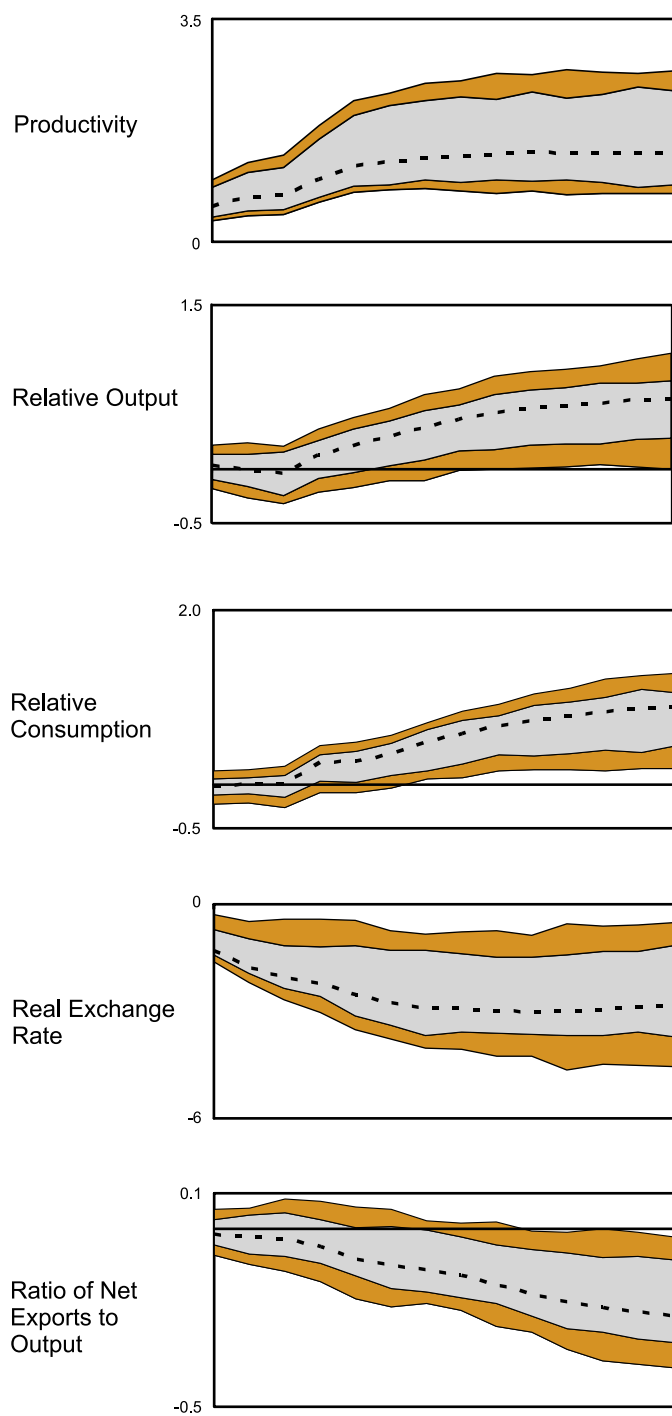
Notwithstanding the emergence of globalization over the last couple of decades, economies remain, to some extent, strikingly insular. Indeed, theory predicts that as the world becomes more integrated, consumption should be higher in countries where the relative price of consumption, the real exchange rate, is lower. In fact, we observe the exact opposite in the data: Consumption is higher in countries where the relative price of consumption is higher! One reason for this puzzling fact is that real exchange rates often move in a way that hinders the risk-sharing process in response to technological changes, accentuating the benefits to winners and the losses to losers. 

¹¹ Note that a productivity increase can theoretically raise the real exchange rate if the productivity improvement is concentrated in the traded-goods sector and countries produce very similar traded goods. However, models in which countries specialize in the production of a particular array of traded goods generally predict a depreciation of the real exchange rate following a technological improvement.

¹² All of our variables are in growth rates. For labor productivity, real GDP, and real consumption, we take the difference between the growth rate of these variables in the U.S. and in the rest of the OECD countries. Our measure of labor productivity is that of the manufacturing sector.

¹³ We also looked at the sensitivity of our results when we substituted total factor productivity for labor productivity: Our results are robust to this change. See my working paper with Giancarlo Corsetti and Luca Dedola for more details.

¹⁴ In our working paper, Corsetti, Dedola, and I detail the theoretical reasons underlying an appreciation of the real exchange rate and the terms of trade following an increase in the productivity of the traded-goods sector.

FIGURE 3**Impulse Responses to a Technology Shock in the U.S.**

The charts describe the responses from a five-variable VAR, using quarterly data. The variables are labor productivity, the real exchange rate (constructed using CPI indices), relative consumption (i.e., domestic minus foreign consumption), relative output, and net exports. All series are in percent.

REFERENCES

- Corsetti, Giancarlo, Luca Dedola, and Sylvain Leduc. "International Risk-Sharing and the Transmission of Productivity Shocks," Federal Reserve Bank of Philadelphia Working Paper 03-19, 2003.
- Crucini, Mario J., Chris I. Telmer, and Mario Zachariadis. "Understanding European Real Exchange Rates," Vanderbilt University Working Paper 01-W20 (2001).
- Engel, Charles. "Accounting for U.S. Real Exchange Rate Changes," *Journal of Political Economy* 107 (1999), pp. 507-538.
- Engel, Charles, and John H. Rogers. "How Wide Is the Border?" *American Economic Review* 86 (1996), pp. 1112-1125.
- Froot, Kenneth A., Michael Kim, and Kenneth Rogoff. "The Law of One Price Over 700 Years," National Bureau of Economic Research Working Paper 5132 (1995).
- Nakamura, Leonard. "The Retail Revolution and Food-Price Mismeasurement," Federal Reserve Bank of Philadelphia *Business Review*, May/June 1998.
- Sill, Keith. "The Gains from International Risk-Sharing," Federal Reserve Bank of Philadelphia *Business Review*, Third Quarter 2001, pp. 23-32.
- Rogers, John H., and Michael Jenkins. "Haircuts or Hysteresis? Sources of Movements in Real Exchange Rates," *Journal of International Economics* 38 (1995), pp. 339-60.
- Rogoff, Kenneth. "The Purchasing Power Parity Puzzle," *Journal of Economic Literature* 34 (1996), pp. 647-68.
- Warnock, Francis E. "Home Bias and High Turnover Reconsidered," *Journal of International Money and Finance* 21, pp. 795-805.

Legal Uncertainty and Contractual Innovation

BY YARON LEITNER

I

nnovative contracts are important for economic growth, but when firms face uncertainty as to whether contracts will be enforced, they may choose not to innovate.

Legal uncertainty can arise if a judge interprets the terms of a contract in a way that is antithetical to the intentions of the parties to the contract. Or sometimes a judge may understand the contract but overrule it for other reasons. How does legal uncertainty affect firms' decisions to innovate? In this article, Yaron Leitner explores issues related to legal uncertainty, particularly the amount of discretion judges have and the types of evidence they consider.

Innovation — which is important for growth and prosperity — is inherently uncertain. When a firm launches a new product, it faces uncertainty regarding the product's success. Similarly, when two firms (or individuals) enter a contract containing novel terms, they face uncertainty as to whether the contract will be enforced in court. In other words, they face legal uncertainty. New contracts are important for economic growth as

they enable the coordination of novel activities and relationships; however, when firms face legal uncertainty, they may choose not to innovate.¹

Legal uncertainty can stem from the fact that the judge interprets the contract differently from the parties' intentions when they entered the contract. It can also stem from "active judges" who understand the contract but overrule it for some other reason, such as concerns for third parties who might be affected by the underlying arrangement.

How does legal uncertainty af-

fect the new contracts we enter? How can courts affect legal uncertainty and firms' decisions about whether to innovate? I will explore these questions and related issues in this article. In particular, I will focus on the amount and type of evidence judges consider and the amount of discretion judges have.

AN EXAMPLE OF LEGAL UNCERTAINTY

Let's begin by illustrating legal uncertainty that results from an interpretation of a word. Even a simple word such as mandatory can sometimes be ambiguous. Take the case of Eternity Global Master Funds Ltd. ("Eternity") against Morgan Guaranty Trust Company of New York and JP Morgan Chase Bank ("Morgan") in 2002.² Eternity lent money to Argentina (it purchased Argentina's bonds) and protected itself against the risk that Argentina would fail to meet its debt payments by purchasing credit swaps contracts from Morgan.³ The contracts between Eternity and Morgan incorpo-

² The following description is based on the court's rulings. See *Eternity Global Master Fund Limited v. Morgan Guaranty Trust Company of N.Y. and JP Morgan Chase Bank*, United States District Court for the Southern District of N.Y., October 29, 2002, and June 5, 2003.

³ Credit swaps are a common way for lenders to protect themselves against the risk that a borrower will default. These swaps usually work as follows: The buyer promises to pay fixed periodic payments. In return, if a third party defaults, the seller pays the buyer the loss due to the default. Thus, you can think of the seller as providing the buyer with long-term insurance against default in return for an annual insurance premium. In our case, Eternity was the buyer, Morgan was the seller, and Argentina was the third party.



Yaron Leitner is an economist in the Research Department of the Philadelphia Fed.

¹ Negotiable debt instruments and the limited liability corporation are examples of contractual innovations that have been important for economic growth, yet subject to significant legal uncertainty.

rated terms from the 1999 ISDA *Credit Derivatives Definitions* published by the International Swaps and Derivatives Association (ISDA). In particular, the contracts said that Morgan would pay Eternity should a “credit event” occur, and the definition of a credit event included a few scenarios capturing the idea that Argentina will fail (or has failed) to meet its originally agreed-upon debt obligations.⁴

A dispute between Eternity and Morgan came up when Argentina, facing financial problems, announced a “voluntary debt exchange,” in which it offered its lenders the opportunity to exchange their debt for new loans with less favorable terms. Eternity argued this was a credit event, whereas Morgan maintained it was not. The judge, of course, had to decide.

The problem was that the definition of a credit event in the contract did not explicitly raise the possibility of a voluntary exchange, but it did raise the possibility of a *mandatory* exchange, which, according to the contract, qualified as a credit event. Morgan argued that since Eternity had the option of not exchanging its debt, the exchange was voluntary rather than mandatory; therefore, a credit event had not occurred. In contrast, Eternity argued that “mandatory” should be read to encompass situations that are “economically coercive,” and therefore, Argentina’s exchange offer qualified as a credit event. Eternity might have meant, for example, that even though, in principle, it had the option of not exchanging its debt, in practice, it had to do so because otherwise Argentina would not have paid anything on its original debt.

⁴Incorporating standard terms, such as those published by the ISDA, is an example of boilerplate or off-the-shelf text that reduces writing costs as well as legal uncertainty.

The judge, interestingly enough, presented two different views. At a first trial, he did not take a stand on the word mandatory and, instead, used a different reasoning to rule that a credit event had occurred.⁵ However, at a later trial, he reversed himself, saying that “upon further study, the court believes its analysis was incorrect.” This time, he ruled that a credit event had not occurred, basing his decision on the dictionary meaning of mandatory.

LEGAL UNCERTAINTY AFFECTS INNOVATION

Innovation May Not Take Place.

When a firm is not sure whether a new technology will succeed, it may sometimes choose to stick with an old one, even though the new technology might be more efficient. Similarly, when the contracting parties are not sure how courts will interpret a new contractual term, they may choose not to incorporate it into their contract and, instead, use terms that are more familiar. In other words, they may choose not to innovate.

To illustrate the point above, go back roughly 200 years, and consider the following example: As the owner of a small business, you try to raise money to finance a project that looks very promising. The bank is willing to lend you some money but requires that you post the building and machines as collateral. This means, of course, that you cannot sell those assets without permission from the bank. It also means that if you default, the bank will take immediate possession of the assets; so it knows it will get its money

⁵He ruled that the exchange qualified as a credit event because there had been an agreed-upon deferral of payments.



back. But there is one problem: The amount the bank is willing to lend you is only half of what you need. What will you do?

One option is simply to forget the project. Another option is to create a new type of mortgage contract that will allow you to raise more money without exposing the bank to too much risk of not getting its money back. One way to do it is for you to increase the amount of collateral you can post, say, by putting up your entire business as collateral; in particular, you will pledge not only the assets you own today but also the assets you may own in the future, such as inventories or accounts receivable. Since this creates more collateral, the bank will be willing to lend you more, so that you will have all the money you need to take on the new investment opportunity.

Sound like a good idea? In principle, it does. But unfortunately, the bank is not willing to lend you the extra money, saying it does not want to take the risk that the courts will not enforce this innovative contract.

In a working paper, Julian Franks and Oren Sussman discuss two cases in which companies entered contracts similar to the one above. The ultimate outcomes were very different, however.

The first case occurred in England in 1870. A steamship company called the Panama, New Zealand, and Australian Royal Mail Co. borrowed

money using its “undertaking and all sums of money arising therefrom” as collateral.⁶ When the case came before the courts, the judge interpreted “undertaking” as covering all of the assets owned by the company at the time of default. According to Palmer’s textbook on company law, the judge essentially recognized that a mortgage can be placed not only on an object currently owned by the company but also on a class of assets that may be acquired in the future.

The second case occurred in the U.S.⁷ It involved a loan made in 1839 from Winslow to a cutlery manufacturer. The borrowing company used the “machinery, tools and implements...which we may anytime purchase” as collateral for the loan. When the company went bankrupt, Winslow took possession of some of the machinery, tools, and stock in trade that were mortgaged to him under the original contract. Mitchell filed suit on behalf of all of the other creditors, claiming that Winslow was not entitled to the property and that the mortgage instrument was not valid because it was on goods that were not yet in the possession of the manufacturer. A state judge dismissed Mitchell’s claim, arguing that the mortgage was properly registered and disclosed. However, superior courts later accepted the claim that this type of collateral, “the floating lien,” was not a valid instrument, arguing that a mortgage could be secured only on current (existing) property. If new property were acquired, a new mortgage had to be taken out.⁸ It took nearly 100 years before the restrictions against this type of security were abolished in the U.S.

In both the U.S. and England,

these initial rulings had lasting effects because they became precedents for subsequent courts.

New Contracts May Set Inefficient Standards. When previous rulings set precedents for future rulings, subsequent firms face less legal uncertainty than the innovators.⁹ Thus, the contract written by the innovators need not be the best one for those who use it afterward. Nonetheless, these followers may stick with the tried-and-true contract because judges will enforce this contract consistently.

To illustrate this, return to the

lender will take control only in half of the cases in which you default. In the other half, you will keep control, even though you have failed to pay. (This other half might correspond to cases in which you breached the contract but eventually came into compliance.) In some cases, such a contract may be optimal both for the borrower and for the lender. The lender is happy because the threat of losing the business gives the borrower the incentive to put a lot of effort into running the business, which, in turn, increases the probability that the borrower will be able

When previous rulings set precedents for future rulings, subsequent firms face less legal uncertainty than the innovators.

previous example in which you wanted to borrow against your entire business. In practice, debt contracts often include covenants that give the borrower a fixed period of time to get back into compliance. In many cases, the borrower has one or two months to remedy an initial breach of contract and avoid default. This gives the borrower more time to come up with the funds and thus reduces the chances that the creditor will seize the borrower’s assets if the borrower breaches the terms of the covenant.

More generally, in theory, we can think of contracts that specify the *probability* that the lender will be able to take control of the borrower’s assets if the borrower defaults. In particular, in our example, suppose that rather than saying that if you default, the lender automatically takes control of your business, you want to say that the

to pay back the loan. The borrower is happy because he gets some protection against bad luck — situations in which he was unable to make a payment, even though he put a lot of effort into the business.¹⁰

Now suppose the contracting parties think there is a 50 percent chance the court will not enforce their contract. Assume that, in that case, the lender will not be able to take control of the business. Then it may be optimal for the parties to enter a contract that does not reflect their true intentions. The reason is that if they enter a contract that reflects their true intentions (saying that the lender takes control only in half of the cases in which the borrower defaults), and the court enforces it only half of the time, the lender will effectively gain control only in a quarter of the cases. If, on the other hand, the parties enter a contract that says if the borrower

⁶ According to the Merriam-Webster online dictionary, “undertaking” means the business of an entrepreneur.

⁷ *Mitchell v. Winslow*, 1843.

⁸ *Jones v. Lewis Richardson*, 1845.

⁹ In practice, rulings made by high courts usually bind lower courts, but a single ruling of a lower court need not become a precedent for other courts.

¹⁰ Simply transferring control to the lender will not generally be efficient. The assets are often more valuable in the borrower’s hands; however, the lender may care only about his own share.

defaults, the lender always takes control, and the court enforces it only half of the time, the lender will effectively gain control in half of the cases. This is exactly the outcome the contracting parties intended when they entered the contract, even though they specified something else in the contract.

The problem, according to Franks and Sussman, is that if previous rulings become precedents for future rulings, once the court enforces the first contract, firms in the future may prefer to enter exactly the same contract, rather than incur the cost of revising it. This is because by doing so, they can avoid legal uncertainty — they know the judge will enforce the contract. Consequently, entering a contract that says “always transfer control” may become the standard, even though the outcome involved is optimal only for the innovating firm and not for other firms.

THE EVIDENCE COURTS CONSIDER CAN AFFECT INNOVATION

We have seen how legal uncertainty can negatively affect the innovation process. Legal uncertainty, in turn, may depend on the way courts act when they face a new contract. Different judicial practices can either facilitate innovation or stand in its way.

One feature of a judicial process that might affect legal uncertainty is the amount and type of evidence courts can use to interpret an ambiguous contract. A British judge, for example, often won't take account of evidence of informal promises different from the explicit contractual terms. However, the Uniform Commercial Code, which governs commercial transactions in the U.S., directs a U.S. judge to consider such evidence when explicit contractual terms are vague. The Uniform Commercial Code also captures the idea that an agreement is

to be read in light of the parties' previous transactions (“course of dealing”).

This raises many questions: Should courts consider evidence of prior negotiations between the parties to interpret an ambiguous contract? If so, should courts be allowed to consider prior negotiations to decide whether the language is actually ambiguous?

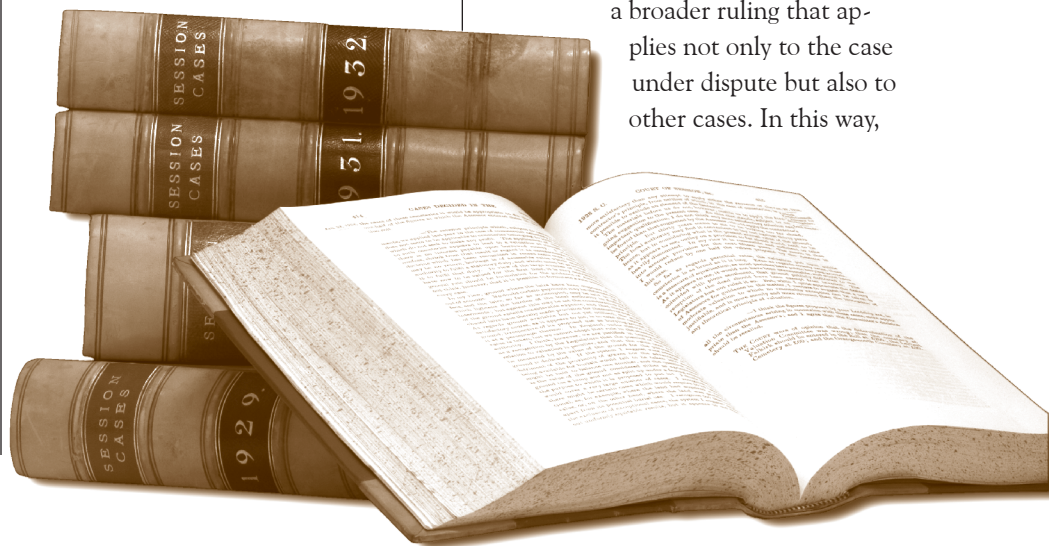
More Evidence Can Help the Judge Interpret the Agreement... As part of our current research agenda, Mitchell Berlin and I have investigated these issues as well as related ones. We start by assuming that when companies introduce new contractual terms, they face legal uncertainty; they can never be sure how courts will interpret their contract. This, as we have already seen, can keep firms from innovating. We also assume that when the judge considers more evidence, such as prior negotiations or course of dealing, he is more likely to “rule correctly.” In other words, he is more likely to guess correctly the intentions the contracting parties had when they entered the contract.¹¹ This can motivate firms to innovate new contracts because the legal uncertainty they face is reduced.

¹¹ Thus, we differentiate between the *written contract* and the *implicit agreement* that reflects the parties' intentions. The assumption is that the judge is more likely to rule correctly when he looks at evidence that tells something about the specific agreement.

What we have in mind are contracts that specify future payments. You can think of the insurance contract between Eternity and Morgan or the mortgage contract from the previous section. We assume there is no disagreement between the two parties when they enter a contract. In other words, they agree on what should happen in each possible scenario. However, at a later stage, when one party has to pay, he may prefer to go to court, hoping the judge will not enforce the contract because of misinterpretation.

...But May Make It Harder to Build Precedents. While looking at more evidence may help the judge interpret the contract correctly, it may not be good for everyone. As in the previous section, what's optimal for the first firms that innovate may not be optimal for subsequent firms. In the previous example, precedents not only reduced uncertainty but they also induced subsequent firms to use inefficient contracts. In our case, the problem is that precedents may not be established at all. If the court uses evidence that is too case specific, subsequent firms or individuals using the same contractual term may not learn how the judge will interpret the novel term in *their* case. This is because the evidence used in the first case may not apply in other cases. If, instead, the court does not use case-specific evidence to interpret the contract, it needs to set a precedent, that is,

a broader ruling that applies not only to the case under dispute but also to other cases. In this way,



legal uncertainty is reduced for subsequent firms.¹²

An interesting implication of the tradeoff above relates to the speed with which the innovation is adapted: When judges look at more case-related evidence, the innovation process may start earlier, but it may take more time for the innovation to be widely diffused. The intuition behind this result is that the higher the legal uncertainty firms face, the less likely they are to incorporate new terms into their contract. This is because they always have the alternative of sticking with familiar terms and old standards. When the judge is more likely to rule correctly because he looks at more evidence, it may be easier to find a company willing to be the first to innovate. That's why the innovation process may start earlier. However, after the first innovation is brought to court, it may not become easier to find another firm that will use the new terms. Thus, the innovation spreads slowly to other firms. If, on the other hand, the judge did not use evidence to interpret a contract, it could be more difficult to find a firm willing to take the first step and use an unfamiliar term. However, once a case is brought to court and the judge makes a broad ruling, more firms are likely to use the new term because they are faced with less legal uncertainty.

Irrelevant Evidence May Make Innovation More Costly. The assumption that more evidence helps the judge interpret the contract may depend on the process by which evidence is collected. In the civil law countries of Europe (e.g., France), the judge is in charge of collecting evidence; so he can make sure that only evidence relevant to the case is collected. In contrast, in the U.S., lawyers are in

charge of collecting evidence. They need to collect all evidence before the trial begins; therefore they try to collect as much evidence as possible. In his article, John Langbein suggests that this process can lead to inefficiencies because lawyers may choose to collect evidence that is not relevant to the case, and that can lead the court to make wrong decisions. To prevent these mistakes, the contracting parties may try to write very detailed contracts. But when new contracts are very different from old ones, doing so may make innovation more costly.¹³

The example above shows how legal uncertainty can lead to very detailed contracts. However, in some cases, legal uncertainty can actually lead to contracts that are not as detailed as they could have been. (See *Legal Uncertainty Can Also Make Contracts More Incomplete*.)

JUDICIAL DISCRETION AFFECTS INNOVATION

Another factor that may affect legal uncertainty, and thus the innovation process, is the amount of discretion judges have when they face a contract that is not ambiguous. In England, judges have been *formalist*, adopting an attitude of deference toward the contractual agreements of private parties.¹⁴ For example, when the London Pressed Hinge Company Limited failed in 1905, the judge concentrated control in the hands of debt holders — even though he thought

it was unfair to do so — because this was what the contract said. The judge was concerned about other creditors that might be harmed, particularly suppliers or trade creditors, who were too weak to contract on their own, and whose junior position in the case of default was not a result of a deliberate contracting decision, but rather a result of their failure to contract at all. Nonetheless, the judge ruled in favor of debt holders because he thought they obtained their rights in a lawful and valid contract.¹⁵

In contrast, in the U.S., judges have been more active, in the sense that they intervened in the innovation process, sometimes in blunt violation of contracted agreements.¹⁶ We have already seen one example in which the courts in the U.S. voided a contract, arguing that a mortgage could be secured only on current property. Another example relates to the failure of the Wabash Railway in 1884. Here, courts in the U.S., wanting to preserve the railroad as a going concern, violated the debt contract by allowing Wabash to appoint two of its own directors as those who would take control of the firm's assets.

Franks and Sussman suggest that the different rulings in the U.S. and England were caused by the differences in views about the appropriate role of judges, rather than the differences of opinion. In both cases, the judges thought it was unfair to concentrate control in the hands of a single person (for example, by pledging the whole

¹² Thus, a judicial precedent is a public good.

¹³ According to many observers, contracts in the United States are much more detailed than contracts originating in the civil law countries of Europe. Langbein's article discusses a number of theories as to why this might be so.

¹⁴ The English corporation was granted the right to contract freely by a series of Acts of Parliament between 1848 and 1856 (the Limited Liability Act), consolidated in the Companies Act of 1862.

¹⁵ See Franks and Sussman.

¹⁶ The U.S. Constitution has allocated the power to innovate new insolvency procedures away from the parties and into the hands of Congress and the federal government. (According to Article 1, Section 8, of the 1789 Constitution, "Congress shall have the power...to establish... uniform laws on the subject of bankruptcies throughout the United States.")

business as collateral). However, they intervened in the U.S., but not in England. According to Franks and Sussman, this difference in approach helps to explain why English bankruptcy law is more creditor oriented (its principal focus is to make sure debts are paid), while American law is more debtor oriented (its principal focus is on rescuing firms in distress).¹⁷

An important issue, then, is how much discretion judges should have. Unfortunately, there is no clear answer. However, economists have begun to explore some of the tradeoffs.

¹⁷ To learn more about the different bankruptcy procedures, read the paper by Julian Franks, Kjell Nyborg, and Walter Torous.

Active Judges Can Protect Contracting Parties from Unforeseen Contingencies. In a recent working paper, Luca Anderlini, Leonardo Felli, and Andrew Postlewaite consider a model with active judges. They show that in some cases, active judges, who are allowed to void contracts, can actually *reduce* the legal uncertainty the contracting parties face, thereby reducing the risk of innovating. In particular, by voiding contracts, courts can protect the contracting parties from “unforeseen contingencies.” The idea is that the contracting parties cannot think of everything; so enforcing the contract “as it is” may subject them to very high cost in situations that could not be foreseen when the contract was entered. One example

they mention is the case of *Spalding & Sons, Incorporated v. The United States*. Spalding had a contract to harvest timber on U.S. government land, and the Bureau of Land Management cancelled the contract after a fire on adjacent property required unforeseen remedial action. When the case was brought before the court, the court upheld the Bureau of Land Management’s right to cancel.

The problem, of course, is that before voiding the contract, the court must decide whether an unforeseen contingency has occurred. This may not always be that simple. Often, judges cannot rely on the contracting parties to say truthfully whether a contingency was foreseen or unforeseen because once the issue has

Legal Uncertainty Can Also Make Contracts More Incomplete

H

ow legal uncertainty makes contracts more incomplete is illustrated in a working paper by Shurojit Chatterji and Dragan Filipovich. In their example, two individuals enter a contract that specifies which action each individual should take. The judge then enforces the contract. The problem is that the

judge may choose actions different from those initially intended by the contracting parties, and this can impose a high cost on one of the two individuals. To hedge against this possibility, the individuals enter a contract that does not specify as much as it could. This gives the individual who can be negatively affected by an erroneous court ruling more flexibility to protect himself.

The logic behind this result builds on the idea that some intrinsic incompleteness — in this case arising from the judge’s difficulty in figuring out the intentions of the contracting parties — can lead to further incompleteness. Douglas Bernheim and Michael Whinston show that when the contracting parties cannot specify some things in a contract, they may intentionally leave other things open, even though they could be specified at no extra cost. In their model, the judge can distinguish among some actions, but not among others. For example, he may be able to tell whether a university gave a faculty member a particular office or whether the faculty member obtained a wage increase. But he may not be able to tell whether the faculty member has put a lot of effort into providing services that benefit the university (e.g., helping in the recruiting process). Thus,

the contract between the university and the faculty member can specify the obligations of the university, but it cannot specify all the obligations of the faculty member. The judge will simply not be able to learn whether the faculty member acted according to the contract, and so he will not be able to enforce it. Thus, the contract between the university and the faculty member is intrinsically incomplete.


Bernheim and Whinston show that this intrinsic incompleteness can lead to further incompleteness. In particular, the contracting parties may *choose* not to specify some of the university’s obligations, even though they could be easily specified in the contract and enforced by the judge. Choosing not to specify allows the university to punish the faculty member (say, by reducing his future pay raises) if the latter shirks his obligations. At the same time, it protects the faculty member from being maltreated by the university.

The logic is as follows: If the contract specified all of the university’s obligations, the faculty member could go to court if the university reneged on its contractual obligations; however, the university could not go to court if the faculty member shirked because the court would not be able to tell whether he had, in fact, done so. In contrast, if the contract left some of the university’s obligations unspecified, the university could punish the faculty member if he shirked. If, instead, the university reneged, the faculty member could punish the university by exerting less effort. Thus, overall, choosing to enter such an incomplete contract could be beneficial to both parties.

come to court, the parties' interests are opposed. When judges mistakenly identify an event as unforeseen, judicial discretion has a cost. Contractual remedies that the parties had knowingly agreed to when the contract was signed are undermined. Whenever agents are concerned that a contract will not be enforced, they are less likely to innovate.¹⁸

CONCLUSION

We have seen that when parties face legal uncertainty, they may choose not to innovate new contractual terms and instead stick with old standards. We have also seen that the way the court rules may affect the uncertainty the contracting parties face, which, in turn, may affect the innovation process. For example, when courts look at case-specific (and relevant) evidence, legal uncertainty is reduced for the first firms that innovate. However, precedents are not established, so uncertainty is not reduced for subsequent firms.

We have also seen that allowing judges to overrule or void contracts may have ambiguous effects. On the one hand, doing so can protect the parties from unforeseen contingencies, and it can protect the interests of third parties. On the other hand, it opens the door to potential judicial mistakes that may undermine incentives and increase the legal uncertainty the parties face. 

¹⁸ In another working paper, Anderlini, Felli, and Postlewaite suggest that voiding contracts can sometimes be good for the contracting parties because it protects them from the risk that one of them will have an information advantage. For example, I might be more willing to buy a car from you if I knew the court would void the contract if I found out that you "forgot to mention" the car was involved in an accident.

REFERENCES

- Anderlini, Luca, Leonardo Felli, and Andrew Postlewaite. "Courts of Law and Unforeseen Contingencies," Working Paper (May 2003).
- Anderlini, Luca, Leonardo Felli, and Andrew Postlewaite. "Should Courts Always Enforce What Contracting Parties Write?" Working Paper (November 2003).
- Berlin, Mitchell, and Yaron Leitner. "Legal Systems and Contractual Innovation," manuscript (December 2003).
- Bernheim, Douglas B., and Michael D. Whinston. "Incomplete Contracts and Strategic Ambiguity," *American Economic Review*, 88, 1998, pp. 902-32.
- Chatterji, Shurojit, and Dragan Filipovich. "Ambiguous Contracting: Natural Language and Judicial Interpretation," Working Paper, ITAM (February 2002).
- Franks, Julian, and Oren Sussman. "Financial Innovations and Corporate Insolvency," Working Paper, London Business School (August 1999).
- Franks, Julian, Kjell Nyborg, and Walter Torous. "A Comparison of US, UK, and German Insolvency Codes," *Financial Management*, 25, 1996, pp. 86-101.
- Langbein, John H. "Comparative Civil Procedure and the Style of Complex Contracts," *American Journal of Comparative Law*, 35, 1987, pp. 381-94.
- Palmer, Francis B. *Palmer's Company Law: A Practical Handbook*. London: Stevens and Sons, 1905.