Conducting monetary policy is a difficult business. It’s easy enough to set a policy goal such as price stability or low and stable inflation, but because monetary policy affects the economy with a lag, achieving those goals requires an ability to peer into the future. A change in the money supply or interest rates today won’t affect inflation for months, or even years, down the road. Consequently, policymakers use economic models and forecasts to help them make decisions.

Historically, economists and policymakers have used two major approaches to help predict future outcomes. The first approach relies on large-scale statistical models of the economy that capture historical relationships among hundreds of economic variables. The second approach is simpler: focus on small models containing just a few variables, such as money growth and interest rates, that seem to provide information about future output growth, employment, and inflation.

Unfortunately, such models can fail to give reliable predictions, especially when factors affecting the economy change in a major way. Thus, after the oil-price shocks of the 1970s, fore-
casting models had difficulty predicting the simultaneous occurrence of high unemployment and high inflation. In the 1980s, major changes were made to regulations governing financial institutions, and many models were affected by the breakdown of formerly trusted links between money, inflation, and output growth. When forecasting models become unreliable, policymakers and economists find it more difficult to predict how today’s actions will affect the future.

Nature abhors a vacuum in economics no less than in physics. As the old ways of divining the future came under fire, new proposals quickly emerged. Many of these proposals suggested monetary policy be guided by, or even target, certain variables that are sensitive to the market’s expectation of inflation, such as commodity prices and interest rate spreads. Perhaps these indicators could give advanced warning if inflation was about to accelerate, allowing policymakers to take steps to ward it off.

Focusing on, or targeting, a small number of expectations-sensitive indicators seems to make life a lot simpler for policymakers. Why bother with complicated economic models or unreliable measures of the money supply if you can act on a few indicators that send good signals about future inflation or output growth? After all, building economic models is difficult, as is trying to understand the causes of inflation and growth.

Unfortunately, there is little reason to believe a monetary policy based primarily on expectations-sensitive forecasting indicators would be successful in the long run. The link between these indicators and economic outcomes is sensitive to many factors in the economy, including the way in which monetary policy is conducted. If policymakers change their method of implementing monetary policy in a substantive way, previously reliable forecasting relationships could easily break down. In addition, market expectations can change for reasons unrelated to future inflation or output growth. A policy focused too closely on expectations-driven indicators could easily be led astray. Another concern is the possibility of bad interaction between market expectations, forecasting indicators, and monetary policy. The result might be self-fulfilling changes in expectations and greater economic instability, precisely the outcome these new approaches seek to avoid. Monetary policy is more likely to succeed if it is guided primarily by variables tied to the underlying causes of inflation and economic growth and not by variables tied to expectations of inflation and growth.

**MONETARY POLICY: A FORWARD-LOOKING VENTURE**

Many economists and policymakers are convinced monetary policy can best contribute to maximum sustainable economic growth by delivering a stable price level or low and stable inflation. (See *The Benefits of Low and Stable Inflation.*) Because monetary policy affects inflation with a lag, policymakers who want to keep inflation low and stable must rely, to some extent, on forecasts or indicators of future inflation. In the 1960s and early 1970s, one standard way of predicting inflation relied on the Phillips curve. The Phillips curve posits a link between the unemployment rate and inflation: when the unemployment rate is low, inflation is high, and when the unemployment rate is high, inflation is low. Before the 1970s, a stable Phillips curve seemed evident in the data. But the relationship broke down following the oil-price shocks and high inflation of the 1970s.¹

The breakdown of the Phillips curve returned attention to the relationship between money growth and inflation. A key concept for understanding the link between money and inflation is money demand, which links money to prices, interest rates, and output. Individuals and businesses need money to carry on their daily activi-

¹For more on the Phillips curve, see the article by Robert King and Mark Watson. The 1999 article by Thomas Sargent contains a discussion of the breakdown of the Phillips curve in the context of monetary policy.
The Benefits of Low and Stable Inflation

High and variable inflation hurts economic performance in several ways. Variable inflation raises the uncertainty faced by debtors and creditors. Loan agreements typically specify an interest rate based on the expected rate of inflation over the life of the loan. If inflation differs from what was expected, so will the real return debtors pay creditors differ from what was expected. Since people generally don’t like uncertainty, variable inflation can reduce the flow of credit in the economy. U.S. taxes on interest income and capital gains are not indexed to inflation, so high inflation reduces the after-tax return on saving. As a result, peoples’ decisions on how much to save and spend are distorted. And then there’s what economists call the “shoe-leather” cost of inflation: the higher the inflation rate, the more time and resources people spend minimizing their holdings of currency and demand deposits; that leaves less effort and fewer resources devoted to productive activity. Price stability avoids these distortions.

Monetary policy in many countries reflects an emerging consensus on these benefits. In the United States, central bankers have emphasized a commitment to price stability as a means to achieve maximum sustainable growth. The governments or central banks of some countries, including the United Kingdom, New Zealand, Sweden, and Canada, have adopted numerical targets for very low to zero inflation. Such goals do not mean policymakers are unconcerned about short-run economic developments. But policymakers recognize that monetary policy affects the economy with a lag, making it impossible to smooth out all short-run fluctuations. Moreover, they realize that sometimes a policy can have beneficial short-term results but unacceptable long-run consequences.

While some empirical studies, such as the one by Robert Barro, find moderate rates of inflation are inversely related to economic growth, others find no significant correlation between the two (see the paper by McCandless and Weber). However, very high rates of inflation are associated with lower economic growth.

This conclusion is based on the debate over rules vs. discretion and the time consistency of monetary policies. See chapter 12 in Bennett McCallum’s book for a good discussion of discretionary vs. rule-based policymaking. The article by Herb Taylor and the one by Chari, Kehoe, and Prescott discuss time consistency and economic policy.
in the link between money supply growth and inflation:

\[
\text{inflation} = \text{money supply growth} + \text{growth of velocity} - \text{output growth.}
\]

In principle, policymakers can use their knowledge of money velocity to control long-run inflation. If velocity is constant and money grows faster than output, the result is more money chasing the same amount of goods and services in the economy, so prices rise. Too rapid growth in the money supply eventually leads to higher prices and inflation.3

However, if money demand moves unpredictably, policymakers can’t predict velocity and so are uncertain about the link between growth in the money supply and future inflation. Unfortunately, this appears to be what happened beginning in the 1980s (see Unstable Money Demand and Velocity). Economists Benjamin Friedman and Kenneth Kuttner, focusing on the period since 1975, have shown that the money supply was a useful predictor of the future economy from 1975 until the mid-1980s. But from about the mid-1980s on, they found that the money supply (M1 and M2) had little or no predictive power for future inflation and output. A study by economists Arturo Estrella and Frederic Mishkin, which used the monetary base and M2 as measures of the money supply, came to a similar conclusion.4 Estrella and Mishkin looked at data since 1979 and found that the money supply didn’t predict future inflation or output well, and it’s not a very good indicator of the stance of monetary policy either. In short, something happened in the 1980s that caused a serious breakdown in the link between money and economic activity.

Why did the breakdown occur? Likely candidates include technological innovation, changing financial regulations, and the financial innovation that both responded to and drove regulatory change. For example, before 1980, banks faced regulatory limits on the interest rates they could pay on certain types of deposits counted in the money supply. In the late 1970s, short-term interest rates were very high, so consumers were reluctant to hold funds in deposit accounts that didn’t pay competitive rates. Banks responded by finding a way to, in effect, pay interest on checking deposits: they created negotiable order of withdrawal (NOW) accounts. Many other financial innovations have occurred over the last 20 years or so, including money market mutual funds, expanded use of credit cards, and widespread use of ATM machines. Financial innovations such as these can have a great impact on money demand and make that demand unstable.5

Before the 1980s, simple models that used the growth rate of the money supply to predict future inflation worked fairly well, once they accounted for the oil-price shocks of the 1970s. The instability of money demand in the 1980s meant the models started to have larger-than-usual forecasting errors. In short, the breakdown of the money-demand relationship meant money growth was no longer a reliable variable on which to anchor monetary policy.

IF NOT MONEY GROWTH OR THE PHILLIPS CURVE, WHAT?

Once it became clear that the simple Phillips curve and growth of the money supply had be-

---

3The story is more complicated than we have made it seem. When the Fed increases the money supply, it does so by purchasing bonds in an open-market operation, which increases bank reserves. Banks then increase lending, which results in more money circulating in the economy. In addition, changes in the money supply are associated with changes in real activity, such as employment and output, at least in the short run.

4The monetary base consists of the currency component of the money supply plus bank reserves.

5There is an extensive literature on the instability of money demand. See the review article by Stephen Goldfeld and Daniel Sichel.
Unstable Money Demand and Velocity

The increasing instability of money demand can be easily seen in plots of short-term interest rates and the velocity of money (Figures 1 and 2). Velocity is defined as nominal GDP divided by the money supply. It is a measure of how frequently money changes hands in the economy. When velocity is high, money is turning over rapidly because a relatively small quantity of money suffices to do a year’s transactions in goods and services in the economy. Alternatively, low velocity means a relatively large quantity of money is held to buy goods and services, so money must not be changing hands as frequently.

Velocity is positively related to nominal interest rates. When interest rates are high, people tend to reduce money holdings because money generally earns less interest than other assets. With less money held for economywide spending, velocity tends to be high. As long as velocity can be reliably related to interest rates (or other variables), all is well. But if the relationship changes, the Fed will have trouble figuring out how much money to supply to the economy.

For M1, the trends of velocity and interest rates were similar up until the early 1980s: interest rates were generally rising as was velocity. In addition, velocity was not very sensitive to swings in interest rates. The combination made it easy to predict M1 velocity. After that, velocity started to show much wilder swings than before, and even though interest rates came down quite a bit in the 1990s, velocity didn’t. The relationship between M2 velocity and interest rates also broke down, though the timing is different. The link between velocity and interest rates held up fairly well until the early 1990s: M2 velocity moved closely with interest rates. After that, M2 velocity increased dramatically at the same time interest rates declined.

The breakdown of the link between velocity and interest rates is a symptom of unstable money demand. Policymakers found the money supply was becoming unreliable in that it did not do well in predicting future economic outcomes.
come increasingly unreliable for setting monetary policy, a search for alternative indicators commenced. In the mid-1980s, inflation averaged about 4 percent, but people retained vivid memories of inflation that exceeded 10 percent in 1979 and 1980. Concerns about inflation were reflected in early suggestions that monetary policymakers try to stabilize commodity prices, an action that, it was argued, would lead to general price stability in the economy.  

The notion that monetary policy should stabilize commodity prices or commodity-price indexes met with a good deal of skepticism. Many commodity prices are very volatile and seem to move in response to factors other than movements in the general price level. If policymakers tried stabilizing commodity prices, they would have to respond to sources of volatility that have little to do with inflation. Though the proposal to stabilize commodity prices was flawed, some economists and policymakers argued that commodity prices could still play a useful role in formulating monetary policy. If commodity prices were good predictors of inflation, they might be useful indicators of inflationary pressures in the economy, even if they were not good targets for policy. After all, a desirable monetary policy is one that responds to inflationary pressures well before actual inflation begins to rise. Perhaps by focusing on other predictors of inflation, rather than variables such as money growth and unemployment, policymakers could better achieve goals of low inflation and long-run price stability.

Expectations, Indicators, and Monetary Policy: Caveats

Unfortunately, even if a proposed indicator has been a good predictor of inflation, it may not be a good guide for monetary-policy decisions. If the indicator is driven by market expectations of inflation, it may respond to changes in underlying causes of inflation. But it may also respond to other factors, ones unrelated to future inflation. In the latter case, policymakers face greater uncertainty about the signal an indicator gives: false signals are likely.

Suppose policymakers rely on variables, such as the price of gold, that are driven by market expectations about inflation. If expectations about future inflation change, it doesn’t matter whether those expectations are well founded—the price of gold today will change. If many people start to believe inflation will be higher in the future, that belief will be reflected in higher gold prices, even if the fundamentals of the economy are unchanged. Other potential indicators that are sensitive to market expectations are spreads between long-term and short-term interest rates and survey-based measures of inflation expectations. Of course, these variables respond to other economic factors besides changes in inflation expectations. Gold prices respond to discoveries of new gold deposits and extraction techniques and to wars. Interest rate spreads may change in response to financial uncertainties, such as those brought about by the recent Asian crisis. In addition, changes in the way expectations are formed can lead to large swings in expectations-driven indicators.

Pitfalls When Using Expectations-Driven Indicators for Policymaking. Michael Woodford’s article discusses several pitfalls for policymakers who focus too narrowly on expectations-driven indicators that have forecasted inflation in the past.

First, as people recognize that monetary policy has shifted its focus to a new set of forecasting indicators, those indicators can become unreliable. The basis for this idea is the Lucas
critique, named after Nobel Prize-winning economist Robert E. Lucas, who pointed out that the beliefs of households and firms about the conduct of monetary and fiscal policy are important determinants of economic behavior. (See *The Lucas Critique*.) In making economic decisions, households and firms take into account their beliefs about the course of economic policy and how it will affect them. When policy changes, new beliefs are formed, and economic decisions can then change.

What does this mean for monetary policy? Suppose policymakers announced that henceforth the price of gold will be targeted because gold prices are sensitive to expectations about inflation. Once policy is switched, the old link between gold prices and inflation would almost certainly change, since individuals would buy and sell gold with the knowledge that policymakers are responding to its price. Policymakers would then have to figure out how to interpret the new relationship between gold prices and inflation expectations. The same potential problem would be faced with any variables driven primarily by expectations about inflation or output growth. If policymakers begin to target these variables, it may well happen that their usefulness as forecasters of future in-

---

**The Lucas Critique**

In 1976, Robert Lucas published a now-famous article arguing that the large-scale macroeconomic models then in vogue could, in principle, provide no useful information about the actual consequences of alternative monetary and fiscal policies, even though the models might be very good at short-term forecasting. The argument became known as the Lucas critique and has been very influential in macroeconomics.

The core of Lucas’s argument is the recognition that expectations about policy are an important ingredient in the behavior of households and firms. Implicit in the estimated large-scale models were a set of beliefs by households and firms about the future course of monetary and fiscal policy. If policies changed substantially, new beliefs would be formed, and households and firms would change their behavior. Thus, it made little sense to use models estimated under one set of beliefs to evaluate the consequences of economic policies that would generate new beliefs and different behavior. Models estimated under the old beliefs would likely give incorrect answers.

One way to think about Lucas’s point is in terms of a game: if the rules of the game change, the players will adapt to the new environment by changing their behavior. We can develop an analogy in terms of a football game. Football fans will have noticed that, during the 1998 season, the Philadelphia Eagles almost always punted when confronted with a fourth down in their own territory. It didn’t matter whom the Eagles were playing or where; it was a safe bet they would punt on fourth down when in their own end of the field. On the basis of this history, we would do well to predict that the Eagles would continue to punt on fourth down, and we wouldn’t need any understanding of football to be right most of the time. But suppose we want to analyze a change in the rules. The new rules state that teams get six attempts to make a first down. For anyone who understands football, it is clear that our old model, which says the Eagles punt on fourth down, will not predict well in the new environment: the Eagles will change their behavior in response to the new policy.

---

*aSee the article “Econometric Policy Evaluation: A Critique.”
bThe football analogy was originally developed in a 1986 article by Thomas Sargent.*
flation and growth will change dramatically.

A classic example of this phenomenon is what happened to the Phillips curve in the early 1970s. As discussed earlier, the Phillips curve suggests an inverse link between the unemployment rate and inflation. Before the 1970s, this relationship was apparent in the data. But once policymakers accepted higher inflation to avoid higher unemployment in the 1970s, they found the relationship changed. Households and firms revised their expectations about future inflation as inflation remained high. Expecting higher inflation, workers adjusted their views about the real payoff to working additional hours and about their wage demands; the result was that higher unemployment no longer meant lower inflation. The old link between inflation and unemployment deteriorated when workers changed their beliefs about future inflation. Nowadays, most economists reject the notion of a stable long-run tradeoff between inflation and unemployment.

A second pitfall is that just because a variable doesn’t forecast inflation or output growth doesn’t mean it should be ignored. An example from outside the world of economics might be helpful. If you step on the gas pedal of a running car, the car will accelerate. But suppose you are in a dual-control car, and whenever you step on the gas, someone sitting next to you steps on the brake. Someone who only sees you pushing on the gas pedal might falsely conclude there isn’t much of a relationship between that action and the car’s acceleration.

So, too, with economic variables. Suppose accelerating wage growth indicates higher future inflation. If the central bank used tighter monetary policy to offset accelerating wage growth, it would appear as if there were little or no relationship between wage growth and inflation—even though wage growth was an important variable for policymaking. So, a variable may lack forecasting power because it is already being used in making decisions. In that case, it looks as if the variable is not helpful in making forecasts, even though it is tied to subsequent outcomes. In such a case, we would need a detailed statistical model of the underlying relationships in order to uncover the usefulness of the variable.

There is a third pitfall to beware of when using expectations-driven variables to guide monetary policy: the possibility for bad interactions between policy actions and forecasting variables. Expectations could become self-fulfilling, and the economy could become more volatile. Suppose the central bank uses the following policy rule: whenever three-month interest rates rise more than one percentage point above 10-year interest rates, the central bank will increase the growth rate of the money supply, and whenever 10-year interest rates rise more than one percentage point above three-month interest rates, the central bank will decrease the growth rate of the money supply. This policy might seem sensible, since, in the past, when three-month rates rose well above 10-year rates, the economy often ended up in a recession, and when 10-year rates rose much higher than three-month rates, it often signaled accelerating inflation.

Imagine that for some reason, perhaps an unexpected jump in oil prices, people think there will be a temporary rise in inflation, one that will taper off in the next year or so. The rise in expected inflation causes three-month interest rates to jump up right away; once inflation tapers off, rates are expected to fall back to normal levels. As a result of this expected pattern, three-month interest rates rise above 10-year interest rates. If policymakers respond too aggressively and pump a lot of money into the economy, the belief that there will be higher inflation could be ratified by substantially higher actual inflation. Thus, the mere expectation of inflation could become self-fulfilling, and the economy could destabilize under this policy rule. The key factor is that interest rate spreads are heavily influenced by expectations. Policymakers may find they are reacting to these expectations in a way that affirms changes in beliefs that aren’t tied to economic fundamentals.
BETTER ECONOMIC MODELS TO THE RESCUE? NOT YET

The many problems with using expectations-influenced forecasting variables to guide monetary policy suggest we shouldn’t be too quick to turn away from economic models when trying to understand the behavior of the economy and the implementation of monetary policy. Many economists argue that forecasting models broke down in the 1970s and 1980s because they were not built up from first principles of household and business behavior: they were not sufficiently explicit about the underlying factors governing household choice and business investment. Ideally, the way to assess how forecasting relationships change when policy changes is to have a model that is precise about what causes what in the economy. Such models can trace the effects of policy changes on the economy because they are explicit about the links between variables that cause inflation and growth; the decisions of households, firms, and policymakers; and final outcomes.8 Since the models are explicit about the fundamental determinants of economic behavior and the formation of expectations, alternative monetary policies can be fed into the models, and the effects on variables such as inflation, employment, and output growth can be examined. In addition, we could, in principle, trace how expectations-driven variables are affected by alternative policies.

This is a great idea in theory, but where do we get such models? Macroeconomists have been hard at work on models of the economy for quite some time, but there is, as yet, no consensus on models or modeling strategies.9 Some economists believe business-cycle fluctuations are driven primarily by demand-side factors, such as the money supply, investment, and consumption. Others attribute business-cycle fluctuations primarily to supply-side factors, such as technological progress. As a result, we have a menu of models from which to choose when investigating the consequences of alternative monetary and fiscal policies. This variety of models may not reveal truth, but it does help us analyze which differences in models are important for generating different predictions about policy responses and outcomes. Different models give us alternative ways to sort out and interpret economic data and allow us to frame questions and investigate implications more clearly.

WHAT CAN POLICYMAKERS DO?

Absent a reliable link between the money supply and inflation or a completely trustworthy model of the economy, how can a successful policy be implemented? Since the early 1990s, policymakers have found themselves in the position of not having a completely trustworthy indicator for use in setting monetary policy. Despite this, the performance of the economy has been quite good: inflation is low and the economy has been expanding since April 1991. Economists have investigated an approach that potentially avoids the pitfalls of narrowly focusing on expectations-driven indicators to guide monetary policy: design explicit rules that tell policymakers how to adjust variables they control directly (the federal funds rate or the monetary base) in response to observed deviations of target variables (inflation or nominal

---

8Robert E. Lucas’s 1977 article contains a nontechnical discussion of many of these points.

9Much work has been done over the last 20 years on building better foundational models for policy analysis, for example, the paper by Eric Leeper and Christopher Sims, and the book edited by Thomas Cooley. These sources offer examples of what economists call stochastic, dynamic, general equilibrium models. One key difference between these newer models and older, large-scale statistical models is that the newer models have restrictions across equations that account for how people respond to perceived changes in monetary and fiscal policies. In addition, the new models are based explicitly on the utility and profit-maximizing behavior of households and firms.
GDP) from values policymakers consider desirable. Such rules are less prone to the problems that arise when people change the way they form expectations. However, these rules can still be adversely affected by a change in the economy that alters the links between the variables policymakers control and the variables they ultimately care about.

Frederic Mishkin, in a 1997 article, argues that many central bankers have been following what he calls the “just do it” strategy, in which preemptive moves are made against inflation and recession. This strategy differs significantly from one that focuses narrowly on a small set of expectations-driven indicators. Under the “just do it” strategy, policymakers pay attention to many economic variables and use a variety of models in an effort to cut off inflationary pressure well before inflation actually increases. They also act to forestall a recession before it begins. There is no clearly articulated, explicit strategy for making monetary policy, but the strategy is coherent nonetheless. The policy is not very transparent, since markets don’t know for certain how policymakers are reading the economy at any given moment. But the policy has been pretty successful so far: U.S. inflation is below 2 percent, the unemployment rate is well below 5 percent, and real output has been growing for eight years.

But even under such a strategy, policymakers have to look at economic indicators to judge the stance of monetary policy and assess inflationary pressures. Thus, the warnings we raised about expectations-driven indicators are still important. That is not to say that such indicators are not useful. Policymakers must get a sense of how financial markets expect the future to unfold. However, focusing too narrowly on these variables can quickly lead to the problems enumerated above.

Mishkin points out some drawbacks to the “just do it” approach that could lead to trouble in the future. For one, the lack of transparency—the fact that markets can’t look at a specific set of indicators and infer how policymakers are reading the economy—could result in greater financial and economic uncertainty. As a result, the economy may not operate as efficiently as it otherwise might. Another drawback is that the success of the “just do it” approach depends on the individuals who make policy decisions. Individual policymakers can differ in their abilities and in their ranking of various policy objectives. Under a “just do it” approach, a change in the persons making policy can more easily lead to a change in economic outcomes.

CONCLUSION

The fact that monetary policy affects the economy with a lag means that central bankers have to make decisions today based on how they expect policy to affect output and inflation in the future. In assessing the likely consequences of policy actions, policymakers must pay attention to variables tied to the ultimate causes of inflation. In the past, the money supply filled that role, but it became increasingly unreliable as a tool for the conduct of policy. The alternative of focusing on a narrow set of expectations-driven indicators raises its own set of problems. Forecasting indicators sensitive to expectations can easily lead policy astray and become unreliable when they become the focus of policy.

The best way to understand the interrelationship of monetary policy and forecasting indicators is to develop models of the economy that are explicit about the fundamental determinants of households’ and firms’ choices about spending and investment. By focusing on variables tied to the fundamental causes of inflation and not solely on expectations-driven variables, monetary policymakers are more likely to achieve the goal of low and stable inflation. Building better models of the economy is an ongoing project for macroeconomists.
Unfortunately, policymakers do not yet have the luxury of models that closely match the U.S. economy and are always usable for policy analysis. Consequently, the current policy regime is eclectic: policymakers look at many variables and different models to gauge the appropriate stance of policy and the degree of inflationary pressure.

**REFERENCES**


REFERENCES (continued)


The announcement of Citicorp’s merger with Travelers Insurance, and much of the accompanying commentary by banking observers and consultants, included many references to cross-marketing opportunities, one-stop shopping, and other potential scope economies—the gains from having a single firm provide different goods or services. Some skeptics have noted that these scope economies, which seem so plausible, have been difficult to exploit profitably. The diversification strategies of financial firms from Sears to American Express have been notable more for unfulfilled expectations than for high profits.

While financial firms keep searching for the secret formula to make profits out of providing multiple financial services under one roof, non-financial firms seem headed in the opposite direction. For well over a decade, CEOs in the non-financial sector have increasingly shunned terms such as synergies; currently, their mantras are corporate focus and a renewed emphasis on core businesses and core competencies. For the last few years, economists have been working to document, understand, and evaluate both mar-

*Mitchell Berlin is a senior economist and research advisor in the Research Department of the Philadelphia Fed.
ket participants’ romance with the idea of corporate focus and the problems of diversified firms. So far, the evidence mainly supports business consultants’ arguments that focused firms are more profitable.

Why are financial firms diversifying while nonfinancial firms are becoming more focused? It’s possible that different industries require different strategies. The ruling wisdom in nonfinancial markets may not apply to financial markets, which have been fragmented by regulatory restrictions for much of the last century. Also, there are vagaries in the ways we define products and services. It is not clear whether Citicorp and Travelers even provide different products, if “products” are broadly defined; both firms deliver retail financial services.

Still, the evidence from nonfinancial markets may be instructive to bankers, investors, and also policymakers as they evaluate the potential benefits and costs of the growth of large financial conglomerates, the most likely immediate outcome of the product expansions envisioned in recent financial modernization bills. Plus, a better understanding of the changing patterns of diversification in nonfinancial industries can provide insight into the likely evolution of financial markets less fettered by regulatory boundaries.

TRENDS IN DIVERSIFICATION

When a business chooses to expand into a new product line, it can take a number of directions. Consider Mama Tried, Inc. (MT), a specialized manufacturer of buckled rubber galoshes for children, whose managers anticipate a period of slow growth in demand, as the last baby-boomers’ children become teenagers. MT could seek to exploit its reputation for producing quality raingear by buying a new factory to produce raincoats. Alternatively, the firm could branch into producing rubber hoses, hoping for quantity discounts from the suppliers of its most important input (synthetic rubber). Finally, MT might try to eliminate haggling with its suppliers altogether by purchasing its main supplier of synthetic rubber outright.

Most of us would consider any of these choices as related product extensions that do not change the firm’s degree of focus: (i) raincoats and galoshes are similar products; (ii) rubber hoses and rubber galoshes share a common input; and (iii) MT and its rubber supplier are vertically related, that is, MT uses an input produced by the rubber supplier. We would not view any of these moves as a change of focus, that is, a move into a new product market. (But see Classifying Product Markets in a World That Resists Classification.) Of course, the firm could branch into producing auto parts, a market that almost anyone would classify as unrelated to the production of galoshes. For example, if MT purchases another firm that owns an automotive metal stamping plant, the merger would be classified as a conglomerate merger and a change in focus.

Firms Diversified in the 1960s and 1970s...

Between the end of World War II and the beginning of the 1970s, firms increasingly sought to expand into unrelated product lines (Table 1). These expansions occurred as the pace of merger activity quickened.

What about the economy as a whole during this period? Although there are no systematic data on changes in the degree of diversification for all firms, David Ravenscraft and Frederick Scherer examined a sample of 471 manufacturing firms that participated in the Federal Trade Commission’s (FTC) Line of Business survey between 1975 and 1978. These firms owned approximately 70 percent of the plant and equipment of the U.S. manufacturing sector at that time, so they offer a fairly comprehensive picture of changing business practices in the manufacturing sector and, perhaps, in the economy as a whole.

The authors found that firms in all size classes operated in many more product markets in 1975 than they had 25 years earlier (Table 2). This conglomerate merger wave generated a host of
Classifying Markets
In a World That Resists Classification

The standard industrial classification (SIC) has been used for the last 60 years by economists to identify the products produced by a firm (or firm segment). The code identifies a product in much the same way that a zip code identifies a neighborhood. The first number assigns a product to a very broad category, for example, 0 places a product in the category agriculture, forestry, and fishing, while 1 identifies the category mining. Each subsequent number distinguishes the product at a progressively finer level. Although some establishments report detail down to the seven-digit level, the finest level of detail available for all establishments is the six-digit level, for example, SIC 282104, Plastic resins consumed in the form of granules, pellets, powders, liquids, etc., except sheets, rods, tubes, and shapes.

SIC codes are tremendously useful, but as with any system of categories, there will inevitably be arbitrary choices. Let’s take a look at Mama Tried. At the four-digit level, children’s galoshes are classified under Rubber and Plastics Footwear (SIC 3021). If MT starts producing children’s raincoats, it has entered a new two-digit industry, Apparel and Other Textile Products (SIC 23), which includes Waterproof Outerwear (SIC 2385). But if it begins to produce rubber hoses—included in Rubber and Miscellaneous Plastics Products (SIC 30)—MT hasn’t entered a new two-digit industry, since both galoshes and hoses share the same first two digits (30). Are children’s galoshes and raincoats more or less related than galoshes and hoses are? Now, if MT buys its supplier of synthetic rubber—the main input into galoshes—it enters a new industry at the one-digit level, Chemicals and Allied Products (SIC 28), since the first digit of the new industry (2) is different from that of galoshes (3). But shouldn’t vertically related products be treated as related?*

But even if products could be defined with complete precision, an even more fundamental problem remains. The number of products produced by a firm is only one way of determining a firm’s degree of focus, a point made forcefully by Steve Kaplan and coauthors Mark Mitchell and Karen Wruck. They present a case study of Cooper Industries Inc., a highly efficient manufacturer of commodity-type industrial products, that is, industrial products for which brand image, post-sale product support, or other personal relationships don’t matter much. The firm’s expertise lies in the type of production process at which it excels, rather than the particular products it manufactures. Thus, a firm might produce a wide range of products, spanning the SIC system of codes, yet reasonably be viewed as a highly focused firm. This problem is widely recognized by researchers, but it has not yet been successfully addressed in empirical studies.

In March 1999, the Census Bureau began replacing the SIC system with the new North American Industry Classification System (NAICS). While the earlier system defined product categories according to multiple criteria, including the type of customer served and the underlying production process, the new system adopts a single criterion, the nature of the production process. Reflecting the rapid introduction of new products in recent years, the NAICS has updated industry groupings and introduced more categories at the one-digit level. The Census Bureau intends to review and revise the NAICS every five years to keep pace with changes in the economy.

*Thus, researchers can easily disagree about whether a particular decision to merge or acquire assets involves a move into a related or an unrelated industry. Researchers have used various techniques to overcome this problem. Allowing individual firms to have multiple SIC classifications or using broad groupings (say, two-digit SIC codes) is a conservative approach, which avoids making excessively sharp distinctions between firms’ products. Some researchers have used input-output tables to ensure that vertically related products are not classified as unrelated. At least one researcher, David Scharfstein, took the interesting approach of limiting his sample size so he could examine each firm in the sample and make a personal judgment. Of course, each of these approaches presents difficulties.


Theories about the potential benefits of corporate diversification. (See The Arguments for Diversification.)

...But They Got More Focused in the 1980s and 1990s. Using business segment data reported by firms to the Securities and Exchange Commission and measuring the degree of diversification a number of different ways, Robert Comment and Gregg Jarrell provide evidence that the earlier trend toward diversification was reversed and that firms focused on fewer product lines throughout the 1980s (Table 3). In fact, by the end of the 1980s, over half of the firms in their sample had only one business segment.¹

The wrenching corporate restructurings of the 1980s also point to a move toward greater focus.

¹While the weight of the evidence points to an increase in focus since the 1980s, Cynthia Montgomery makes the telling point that diversification across multiple product lines is still very much the norm for the largest businesses in the U.S. In fact, some of our most successful and profitable firms—General Electric, for example—remain highly diversified. Montgomery finds that between 1985 and 1989 (and 1989 and 1992) product diversification for the 500 largest U.S. public companies stayed roughly constant. This finding suggests that the increase in the average level of focus found by Comment and Jarrell was not true of the very largest firms in the sample.

### TABLE 1

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1948-77, Large Manufacturing &amp; Mining Firms</td>
<td>3.2%</td>
<td>15.9%</td>
<td>33.2%</td>
<td>49.2%</td>
</tr>
</tbody>
</table>

### TABLE 2

Increase from 1950 to 1975 in Number of Lines of Business* for Firms in the FTC’s Survey

<table>
<thead>
<tr>
<th>Company rank in terms of 1975 sales</th>
<th>1950</th>
<th>1975</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top 100</td>
<td>4.48</td>
<td>12.38</td>
</tr>
<tr>
<td>101-200</td>
<td>3.38</td>
<td>9.27</td>
</tr>
<tr>
<td>201-300</td>
<td>2.11</td>
<td>6.56</td>
</tr>
<tr>
<td>301-400</td>
<td>1.73</td>
<td>4.85</td>
</tr>
<tr>
<td>401-471</td>
<td>1.62</td>
<td>3.42</td>
</tr>
<tr>
<td>All 471 Firms</td>
<td>2.73</td>
<td>7.54</td>
</tr>
</tbody>
</table>

*Lines of business are classified into 261 separate manufacturing categories at the four-digit level (compared to 451 four-digit SIC codes for manufacturing in 1972). Respondents were instructed not to report separate lines of business for vertically related products and were permitted to exclude lines of business with sales less than $10 million.

Source: Ravenscraft and Scherer, p. 28.
The Arguments for Diversification

The Benefits of Internal Capital Markets. A firm’s cash flows can be paid out to bondholders and stockholders, who can then choose how to reinvest the funds, perhaps investing them in another firm. Alternatively, the firm’s head office can retain earnings internally and decide how funds should be distributed within the firm—what economists call an internal capital market. Oliver Williamson has argued that the head office of a conglomerate may be better equipped than market participants to evaluate the relative profitability of investment projects. Milton Harris and Artur Raviv showed that a firm’s head office can better design efficient reward and evaluation schemes for capital budgeting purposes. Glenn Hubbard and Darius Palia have provided some evidence that during the diversification wave of the 1960s, investors recognized the benefits of internal capital markets.

The Resource View. Cynthia Montgomery and Birger Wernerfelt have argued that firms seek to extend the use of fixed resources, such as managerial skill, as long as such use is profitable. For example, a firm might diversify outside its core market into less profitable markets if it has exhausted profitable opportunities in its core market but the firm’s managers have spare time. Interestingly, this view predicts a negative relationship between diversification and measures of the profitability of the firm’s investment opportunities (because the firm has moved into a less profitable market). But contrary to theories of inefficient diversification (which are examined in greater detail in the body of this article), the firm’s diversifying investments don’t reduce firm profits.

Tax Benefits. A temporarily unprofitable division’s losses can be set off against the total profits of the firm to reduce the firm’s tax bill. Empirical evidence shows that such tax benefits help explain a firm’s gains from diversification, but the potential gains from being able to offset profits are relatively small.

in the United States. The 1980s were a volatile period in the market for control over corporate assets. Assets were sold off to acquirers and spun off as independent firms by corporate raiders, and the incumbent managers of firms did much the same, partly in fear of becoming a raider’s next target. To a significant extent, we can view the restructurings of the 1980s as an undoing of the excesses of the preceding movement toward diversification.

At least three studies have documented this process. Following the fortunes of the 471 manufacturing firms in the FTC’s Line of Business survey—which accounted for three-fourths of the mergers among all manufacturing firms between 1950 and 1976—Ravenscraft and Scherer estimated that one-third of assets purchased between 1960 and 1976 were ultimately divested. In numerous case studies and using econometric analyses, they found that conglomerate mergers were the most likely to be undone by selling unrelated assets.

Looking at a narrower sample of 271 large acquisitions completed between 1971 and 1982, Steven Kaplan and Michael Weisbach discovered a similar pattern. In their sample, 44 percent of the firms purchased were ultimately sold off to other firms, spun off as free-standing firms, or liquidated outright. These divestitures typically led to an increase in focus, both for the divesting firms and for firms in the aggregate.2 Sanjai Bhagat, Andrei Shleifer, and Robert

---

2The original target firm—the firm initially acquired—was four times more likely to be sold off eventually if its business was unrelated to that of its initial acquirer. Only 20 percent of the assets divested were purchased by an unrelated firm, while 43 percent of the sales were to firms in related businesses. Kaplan and Weisbach view two firms as unrelated if the firms’ four most important lines of business, as reported in Dun and Bradstreet’s Million Dollar Directory, don’t have at least one three-digit SIC code in common.
Vishny provided corroborating evidence of a “return to specialization” in the 1980s through corporate restructurings. Using a sample of 62 successful hostile takeovers between 1984 and 1986, these authors found that fully 72 percent of the assets changing hands ultimately ended up in the hands of buyers in related businesses.3

**WHY HAVE NONFINANCIAL FIRMS INCREASED FOCUS?**

Studies suggest that firms usually become less profitable when they become more diversified. Ravenscraft and Scherer followed their 471 manufacturing firms for a number of years, both preceding and following a merger. They found that during the conglomerate merger wave of the

---

3Bhagat, Shleifer, and Vishny include firms in which the hostile takeover led to temporary control by a leveraged buyout firm that subsequently sold off segments to purchasers in related industries. The authors are not explicit about their criteria for deciding whether two products are related.
1960s and early 1970s, the typical purchaser acquired a firm that had been performing better than others in its industry. Nonetheless, the merged firms’ performance usually deteriorated, and conglomerate mergers were the ones that fared the worst, according to accounting measures of performance.

Firms also seem to become more profitable when they increase their focus. Lane Daley, Vikas Mehrotra, and Ranjini Sivakumar examined the subsequent performance of a sample of 212 spinoffs between 1975 and 1991. In a spinoff, part of a firm is split off and becomes independent. Prior to Daley, Mehrotra, and Sivakumar’s work, many studies had found that a parent firm’s stock price typically rises when a spinoff is announced (in economist’s lingo, a positive announcement effect), evidence that market participants view a spinoff as good news about the firm’s future profitability. But why? The authors discovered that this positive announcement effect was actually confined to those firms spinning off an unrelated segment, but other spinoffs had no positive effect on stock prices at all, evidence that market participants expect higher profits with an increase in focus.

The authors then examined the subsequent performance of both the original firms and of the spun-off firms to see whether market participants’ optimism was rewarded. Corroborating the stock market evidence, the authors discovered improved operating performance at the parent firm, but only if the parent and the spinoff were in unrelated markets.

More Evidence From the Stock Market About the Benefits of Focus. In the 1980s and 1990s, stock market participants have consistently rewarded focus and penalized diversification by paying less for the stocks of diversified firms than for those of their specialized counterparts, the so-called diversification discount.

Philip Berger and Eli Ofek have written a number of influential articles developing the “chop shop” approach to measuring the diversification discount. In this approach, the authors look at each of the diversified firm’s segments separately and compare them to a representative specialized firm in the same industry. This com-

4But note that economists are always skeptical of the uncritical use of accounting measures of economic performance. See the article by Steven Kaplan and coauthors Mark Mitchell and Karen Wruck for some interesting examples of instances in which accounting measures misrepresent the performance of firms following a merger.

5Spinoffs are especially interesting because an analyst can follow the subsequent performance of a firm that is spun off, as well as the performance of its former owner. When a firm is sold, such an analysis often can’t be done because the firm becomes part of the acquirer.

6In Daley, Mehrotra, and Sivakumar’s study, the parent and the spinoff were defined as unrelated if they did not share an industry classification at the two-digit level. At this broad level of classification, most analysts would agree whether two firms are related.

7The chop shop approach was first used by Dean LeBaron and Lawrence Speidell. There are two main variants. In the first, Berger and Ofek calculate a representative stand-alone firm’s ratio of sales to market value and apply this ratio to a segment’s sales to calculate the segment’s hypothetical market value. (The representative firm in an industry is defined as the specialized firm with the median value of the ratio.) They also repeat the procedure using the ratios of assets to market value and earnings to market value to obtain two additional estimates. The second approach compares the diversified firm’s actual market value with a hypothetical market value for the sum of its segments. This hypothetical value is calculated using q—the ratio of a firm’s market value to the book value of its assets, which is often used to measure the market’s valuation of a firm’s investment prospects. The diversified firm’s hypothetical market value is a weighted average of the q-ratios of representative specialized firms, one for each segment of the diversified firm. Berger and Ofek’s 1995 article has a good discussion of the problems with each approach.

8Berger and Ofek define the industry at the four-digit level if there are at least five stand-alone firms in that four-digit-level industry; the authors move to the three-digit level if there are less than five stand-alone firms at the four-digit level, and so on.
parison is used to hypothetically chop up the firm and determine how much each segment would fetch in the market—its hypothetical stand-alone value. Comparing the stand-alone values of the diversified firm’s segments and the firm’s actual market value provides a measure of the diversification discount.9

In their 1995 study, Berger and Ofek found not only that a diversification discount exists but that this discount is large: it averages about 15 percent of the value of the firm as a whole. If such a large discount exists in practice (and not just hypothetically), why don’t the stockholders of such a firm force its managers to sell or spin off assets or demand other changes in the diversified firm’s policy to raise its value?

In a 1997 study, Berger and Ofek, as well as other researchers, answer that this is exactly what happens. Firms with larger diversification discounts are substantially more likely to restructure to increase the firm’s focus. Echoing the results for spinoffs, investigators have found that stock market participants respond favorably to asset sales and divestitures that increase the selling firm’s degree of focus.10

SOURCES OF THE DIVERSIFICATION DISCOUNT

Recent empirical evidence alerts us to the difficulties of running a diversified firm profitably. Why is it so difficult to make a profit?

Jack of All Trades, Master of None? One simple and intuitive explanation has found particular favor in the business press. Staying focused on closely related products is the corporate version of following the well-known adage “stick to what you know.” In this view, a manager’s expertise is specific to a particular product; operating in multiple markets may tax the manager’s attention and abilities. In a closely related view, some economists say that designing the right incentives for managers gets progressively harder as new products are added and the operating environment becomes more complicated.11

The empirical evidence for this view comes mostly from case studies of conglomerate mergers, in which the managers of the acquiring firm simply didn’t understand the subtleties of the new markets they were entering.12 Kaplan and coauthors quote a former manager for Premark, a diversified producer of plastic laminates and other home products (including Tupperware). Referring to his firm’s troubled acquisition of a previously successful producer of decorative kitchen tiles, he said, “We did not know the decorative products business; we knew laminates.”

Kaplan and coauthors also illustrated the complexities of maintaining appropriate incentives and controls for different types of markets. In a case study of Cooper Industries’ purchase of Cameron Inc., they detailed Cooper’s centralized control over even the tiniest expenses—a key to Cooper’s prowess in the production of standardized products. (Cooper’s success in introducing this system in the firms it acquired was so widely admired that industry observers would not necessarily support any such change.

9The sum of the stand-alone values can be lower than that of the actual firm. In that case, the market would be saying the firm is worth more than the sum of its parts, and we would be speaking of a diversification premium.

10Kose John and Eli Ofek present the evidence for asset sales, and Berger and Ofek present the evidence for divestitures in their 1996 study. Of course, even if market participants are correct in discounting the value of a diversified firm, the costs of making a transition to a new (more focused) organizational form may outweigh any potential gains, and shareholders would not necessarily support any such change.

11Chaim Fershtman and Ehud Kalai propose an attractive model of the difficulties for managers of firms operating in complex environments based on managers’ limited cognitive abilities. Julio Rotemberg and Garth Saloner analyze some of the difficulties in giving managers appropriate incentives when a firm operates in multiple markets.

12Ravenscraft and Scherer provide numerous case studies that support this view.
had coined the term cooperization.) However, detailed expense control from the top was viewed as wasteful meddling by Cameron’s managers, who saw personal relationships with customers—not saving nickels and dimes—as the key to success.

The Dark Side of Internal Capital Markets: Corporate Socialism? The idea that the managers of a diversified firm might have an advantage over poorly informed capital market participants in deciding how to allocate funds is one of the most convincing arguments in favor of diversification. But the empirical evidence has increasingly supported an alternative, less optimistic view of internal capital markets. Recent evidence suggests that internal capital markets are prone to propping up poorly performing segments by shifting corporate money from more profitable to less profitable uses.

First, diversified firms that have a segment with negative cash flow suffer a greater loss in value than specialized firms with negative cash flow, evidence that diversified firms stick with losers longer than would the marketplace. Second, segments of diversified firms typically invest more than comparable stand-alone firms in the same industry. But does higher investment indicate a problem or a benefit of internal capital markets? A third finding says that it’s a problem. Segments in industries with strong investment prospects invest less than their stand-alone counterparts, while segments in industries with poor investment prospects invest more than their stand-alone counterparts. Finally, Rajan, Servaes, and Zingales found that the shifting of funds from winners to losers is greatest in those firms in which different segments have widely divergent investment prospects. This result is consistent with the authors’ own view of the internal capital market, in which the head office “buys” the cooperation of segments with weaker prospects by shifting funds in their direction.

13 An internal capital market is an economist’s term for a firm’s practice of allocating cash flows to new investments within the firm, rather than returning cash to stockholders and bondholders as dividends or interest payments and, thus, allowing the funds to be reallocated through financial markets.

14 Much of the empirical evidence for inefficiency is based on the assumption that market participants’ evaluation of a firm’s prospects is usually right. Vojislav Maksimovic and Gordon Phillips have also criticized the previous literature, arguing that an explicit model of optimal investment behavior by multiproduct firms is needed to determine whether a particular pattern of observed behavior is inefficient. They provide evidence that differences in the investment behavior between conglomerate and specialized producers can be explained as the outcome of efficient investment decisions in a market that includes both conglomerate and specialized firms.

15 This has been documented by Berger and Ofek (1995).

16 This has been documented by Berger and Ofek (1995); Hyun-Han Shin and Rene Stulz; David Scharfstein; and Rajan, Servaes, and Zingales.

17 This tendency for firms to shift funds from winners to losers has been termed corporate socialism by David Scharfstein. It has been documented by Owen Lamont, by Scharfstein, and by Rajan, Servaes, and Zingales. Shin and Stulz also show that a segment’s own investment level doesn’t depend on its own prospects compared with those of the other segments of the firm. This finding is further evidence of inefficient investments.

18 In this interesting model, the firm’s divisions bargain for funds, and divisions can sometimes increase their bargaining power by making investments that do not benefit the corporation as a whole. This behavior is most likely for divisions with poor investment prospects. By increasing weaker divisions’ share of corporate funds, the head office buys their cooperation and induces them to make profit-maximizing investments. For example, a weak division’s managers can be induced to forgo an investment that increases the division’s own visibility outside the corporation—but doesn’t increase the corporation’s profits—only as part of an ongoing bargain in which they get a disproportionate share of corporate funds for investment. In the model, inefficient transfers arise only when a firm’s segments have very dissimilar prospects; diversification per se is not the source of the problem.
Some Evidence From Banking Markets. Although the literature has mostly examined non-financial firms, banking scholars have uncovered evidence of internal capital markets in bank holding companies. Consistent with the evidence from nonfinancial firms, findings from a study by Joel Houston, Christopher James, and David Marcus showed that loan growth in a bank subsidiary is affected by its holding company’s cash flow, rather than its own cash flow. In addition, loan volume grows faster at a bank when the cash flows of nonbank subsidiaries of the bank’s holding company are high, thus demonstrating that the holding company’s internal capital market leads to bank lending decisions different from those that would be made by a stand-alone bank. However, unlike the literature on nonfinancial organizations, there is, so far, no empirical evidence of inefficiency.19

IF FOCUS IS SO GREAT, WHY DID FIRMS DIVERSIFY IN THE FIRST PLACE?

The extensive evidence for a diversification discount has given rise to lively debate. How did it arise in the first place, and why does it persist? Have stockholders always penalized diversified firms? If the answer to this last question is yes, then managers, not stockholders, were the intended beneficiaries of diversification. For example, some financial economists argue that stockholders suffer from managers’ desire to build great empires rather than to increase profits.20

The Behavior of the Diversification Discount Over Time Is a Puzzle. In three separate papers, Henri Servaes, John Matsusaka, and Peter Klein provided evidence that stockholders supported diversifying mergers during the early 1970s, the latter part of the conglomerate merger wave. That is, it appears that the diversification discount did not exist during the first half of the 1970s. This finding supports the view that the diversification wave of the late 1960s and early 1970s may have been a mistake, but not the view that it was a victory for managerial self-interest over stockholder value.

But Servaes also found that stock markets imposed a significant penalty on diversifying mergers throughout the 1960s, the period in which conglomerate mergers first picked up speed and then peaked (according to most measures). Thus, evidence from the stock market doesn’t fully support the view that it is only with 20-20 hindsight that stockholders found out that most conglomerate mergers were against their best interests. This historical pattern in which stockholders penalized diversification throughout the 1960s, changed their minds in the first half of the 1970s, then shifted back again in the 1980s and 1990s remains a puzzle that no existing theory has explained convincingly.

Maybe the Economy Has Changed. The increasing diversification that occurred from 1950 to 1975, which was followed by a trend toward greater focus in the 1980s and 1990s, has led some economists to hunt for changes in the economy that might have made diversification attractive at first, but less so later on.

19Peter Klein and Marc Saidenberg found that banks within bank holding companies lend more than otherwise similar stand-alone banks and are equally profitable. Hence, beneficial geographical diversification effects within bank holding companies are not outweighed by the types of organizational difficulties that have been found in diversified nonfinancial firms. However, note that this result does not imply that product diversification in financial firms will prove to be more efficient than for nonfinancial firms. It is reasonable to view lending in Delaware and lending in Illinois as substantially the same financial product, at least once a bank has gained sufficient experience.

20Empire building is only one reason managers might have a desire to diversify even against stockholders’ interests. Some economists, for example, Yakov Amihud and Baruch Lev, have argued that managers seek to reduce their own risks at investors’ expense by diversifying into different markets.
For example, antitrust policy changed under the Reagan administration and could have generated a trend toward greater focus. Until the 1980s, horizontal mergers—mergers between firms in the same product markets—were viewed with hostility by the antitrust authorities. According to one account, the hostility of the Justice Department kept a lid on combinations between large firms in the same market, until the Reagan years ushered in a new antitrust policy that was less suspicious of concentrated markets. This change permitted a return to specialization.

Unfortunately, the evidence doesn’t offer much support for this view. John Matsusaka has presented compelling evidence from the early 1970s that large firms—firms subject to surveillance by the Justice Department—and small firms—firms unlikely to attract the Justice Department’s attention—were equally likely to merge with a firm in an unrelated market. Since antitrust considerations don’t seem to explain conglomerate mergers, why view the move to greater focus as a belated attempt to assemble firms that were previously forbidden? In fact, there is evidence that the mergers, acquisitions, and asset sales that occurred between 1981 and 1989 didn’t actually increase concentration on average.

Another possibility is that capital markets became more efficient at allocating funds and assets and at disciplining poor managers. In this view, internal capital markets may have been necessary in the 1950s and 1960s, but the stock and bond markets can now do the job better.

It is easy to point to fundamental changes that have increased the efficiency of capital markets as a means of mobilizing funds and keeping a check on managers. For example, the deregulation of investment banking fees and shelf registration have made it cheaper and easier for firms to go to capital markets to get funds. Using junk bonds to finance mergers was an innovation of the 1980s. The growth of large institutions with big investments in individual firms has forced managers to pay more attention to investors.

This vision of competing corporate governance mechanisms is a tempting explanation, but it is also one that relies on loose, albeit plausible, connections among a lot of different events rather than on formal empirical tests. Before accepting this explanation, most economists would insist on more concrete evidence.

Managers and Stockholders Don’t Always Agree. Although the evidence doesn’t provide strong support for the view that the conglomerate merger wave was driven by managers’ interests alone, there is growing evidence that managers and stockholders often don’t see eye to eye on the benefits of diversification.

Perhaps the most convincing evidence for conflict between managers and stockholders has been assembled by Berger and Ofek (1997), who compare a sample of firms that refocused between 1983 and 1994 to another group of otherwise similar diversified firms. Firms that refo-

21This hypothesis was initially posed by Ravenscraft and Scherer.

22Julia Porter Liebeskind, Tim Opler, and Donald Hatfield’s study found a modest increase in concentration, but it was the result of exit by small firms and the internal growth of existing firms. They define industries at the four-digit SIC level.

23See Amar Bhide’s article.

24Shelf registration allows firms to issue securities multiple times, without having to undergo the entire battery of reporting and registration requirements imposed by the SEC for a new issue.

25Note, the diversification discount doesn’t, by itself, prove that diversified firms are being held together over stockholders’ opposition. For example, managers and stockholders might agree that an increase in focus would raise profits, but not enough to outweigh the costs of reallocating assets through sale, spinoff, or merger.
cused were much more likely to have replaced their management during the preceding year—evidence that stockholders intervened not only to increase firm focus but also to replace the management identified with the firm’s prior strategy. Firms that refocused were also much more likely to have introduced a stock option compensation scheme for managers. Often, such schemes are intended to motivate managers to make decisions in stockholders’ interests. Thus, refocusing is often undertaken as part of a strategy to make stockholders’ interests paramount. In general, Berger and Ofek’s evidence paints a picture in which firms undertake a strategy of increasing firm focus mainly when stockholders are successful in exerting more control.

Other studies have shown that if a firm’s ownership structure promotes stockholders’ interests first and foremost, the firm is also more likely to be focused.26 Firms with higher managerial stock holdings are typically more focused, perhaps because higher stock holdings give managers financial incentives more like those of stockholders. In addition, firms with large outside investors—that is, investors not closely allied with top management—are more likely to be focused. Large outside investors can exercise more influence over a firm’s policy than can small investors, and they can exercise this influence on behalf of stockholders if they are independent of top management.

CONCLUSION

Nonfinancial firms have become more focused since the 1980s. This shift reversed the previous postwar trend of increasing diversification, which culminated in the conglomerate merger wave of the 1960s and early 1970s. Financial market participants now reward focus and penalize diversification by paying lower stock prices for diversified firms, the so-called diversification discount. This pattern of rewards and penalties appears to be more than a whimsical choice by stockholders, since studies suggest that diversification often hurts a firm’s performance. While the costs of diversification remain somewhat mysterious, recent empirical evidence points to inefficient investment decisions by diversified firms. In particular, there is evidence that diversified firms tend to prop up poorly performing divisions by transferring resources from more profitable divisions.

In light of the evidence from nonfinancial markets, investors may be skeptical about claims of large benefits from diversification by financial firms, especially if firms diversify very far afield from their core businesses. They should be particularly skeptical of the claim that diversified financial firms are best capable of responding quickly to shifting market conditions, that is, that resources can be more easily shifted from declining to promising markets within a diversified firm than through the marketplace. In fact, preliminary evidence points to problems in shifting resources toward their most profitable use as the most likely culprit behind the diversification discount.

Recent evidence from nonfinancial markets may also provide policymakers with hints about how markets will evolve over time if financial firms can offer a wider range of products and services. The evidence does not support the view that giant, diversified financial and commercial conglomerates are a necessary outcome of product deregulation. Financial firms free to choose their mix of products are unlikely to obey regulatory boundaries drawn in the first half of the 20th century, but product deregulation may still ultimately lead to markets composed of specialized financial firms. Evidence from nonfinancial markets says there are limits to profitable product diversification, and market participants have increasingly pressed firms to recognize these limits.

26The evidence in this paragraph is from the articles by Berger and Ofek (1996), and David Denis, Diane Denis, and Atulya Sarin (1997).
REFERENCES


REFERENCES (continued)


REFERENCES (continued)


