Inside the Black Box: The Credit Channel of Monetary Policy Transmission

Ben S. Bernanke and Mark Gertler

Most economists would agree that, at least in the short run, monetary policy can significantly influence the course of the real economy. Indeed, a spate of recent empirical research has confirmed the early finding of Friedman and Schwartz (1963) that monetary policy actions are followed by movements in real output that may last for two years or more (Romer and Romer, 1989; Bernanke and Blinder, 1992; Christiano, Eichenbaum and Evans, 1994a,b). There is far less agreement, however, about exactly how monetary policy exerts its influence: the same research that has established that changes in monetary policy are eventually followed by changes in output is largely silent about what happens in the interim. To a great extent, empirical analysis of the effects of monetary policy has treated the monetary transmission mechanism itself as a “black box.”

Of course, conventional views of how monetary policy works are readily available. According to many textbooks, monetary policymakers use their leverage over short-term interest rates to influence the cost of capital and, consequently, spending on durable goods, such as fixed investment, housing, inventories and consumer durables. In turn, changes in aggregate demand affect the level of production. But the textbook story is incomplete in several important ways.

One problem is that, in general, empirical studies of supposedly “interest-sensitive” components of aggregate spending have in fact had great difficulty in identifying a quantitatively important effect of the neoclassical cost-of-capital variable.¹ Indeed, the most common finding is that nonneoclassical factors—for example,

¹ In a frictionless market, the usual Jorgensonian formula for the neoclassical cost of capital, ignoring expected price changes, is \( (r + d) p_b \), where \( r \) is the required real return to lenders (which could include a premium for systematic risk), \( d \) is the depreciation rate and \( p_b \) is the price of a new capital good.
"accelerator" variables such as lagged output, sales or cash flow—have the greatest impact on spending. (For recent surveys on the determinants of spending components, see Blinder and Maccini, 1991, on inventories; Chirinko, 1993, on business fixed investment; and Boldin, 1994, on housing.) Further, in the relatively few cases in which evidence for a cost-of-capital effect is found, the nonneoclassical determinants of spending typically remain significant (for example, Boldin, 1994; Cummins, Hassett and Hubbard, 1994). Empirical studies that have tested the neoclassical model in its equivalent "Tobin's q" formulation have generally been no more successful.

Beyond the problem of weak cost-of-capital effects in estimated spending equations, there is a presumption that monetary policy should have its strongest influence on short-term interest rates. For example, the federal funds rate, the most closely controlled interest rate, is an overnight rate. Conversely, monetary policy should have a relatively weaker impact on long-term rates, especially real long-term rates. It is puzzling, therefore, that monetary policy apparently has large effects on purchases of long-lived assets, such as housing or production equipment, which—to the extent that they are sensitive to interest rates at all—should be responsive primarily to real long-term rates.

These gaps in the conventional story have led a number of economists to explore whether imperfect information and other "frictions" in credit markets might help explain the potency of monetary policy. Collectively, the mechanisms that have been described in this literature are known loosely as the credit channel of monetary transmission. In this article, we summarize our current view of the credit channel and its role inside the "black box" of monetary policy transmission.

We don't think of the credit channel as a distinct, free-standing alternative to the traditional monetary transmission mechanism, but rather as a set of factors that amplify and propagate conventional interest rate effects. For this reason, the term "credit channel" is something of a misnomer; the credit channel is an enhancement mechanism, not a truly independent or parallel channel. Moreover, this nomenclature has led some authors to focus—inappropriately, in our view—on the behavior of credit aggregates, a point that we discuss later. However, it is probably too late to change the terminology now.

According to the credit channel theory, the direct effects of monetary policy on interest rates are amplified by endogenous changes in the external finance premium, which is the difference in cost between funds raised externally (by issuing equity or debt) and funds generated internally (by retaining earnings). The size of the external finance premium reflects imperfections in the credit markets that drive a wedge between the expected return received by lenders and the costs faced by potential borrowers. According to the "credit view," a change in monetary policy that raises or lowers open-market interest rates tends to change the external finance premium in the same direction. Because of this additional effect of policy on the external finance

---

2 Although we focus on the external finance premium in this article, it should be clear that credit market imperfections may also affect the safe real interest rate in general equilibrium (for example, Stein, 1995b). In emphasizing the external finance premium, we implicitly assume that, over the time span under consideration, the Fed can set the short-term safe rate at the level that it desires.
premium, the impact of monetary policy on the cost of borrowing broadly defined—and, consequently, on real spending and real activity—is magnified.

Why should actions taken by the central bank have any effect on the external finance premium in credit markets? In this article we describe two possible linkages. The first of these, the balance sheet channel, stresses the potential impact of changes in monetary policy on borrowers' balance sheets and income statements, including variables such as borrowers' net worth, cash flow and liquid assets. The second linkage, the bank lending channel, focuses more narrowly on the possible effect of monetary policy actions on the supply of loans by depository institutions.

In our view, the existence of a balance sheet channel seems fairly well established. The bank lending channel is more controversial. Institutional changes during the past 15 years or so have rendered the bank lending channel, at least as traditionally conceived, somewhat less plausible. On the other hand, certain other developments may have increased the importance of bank lending in monetary transmission. In this paper we do not attempt to draw strong conclusions about the relative importance of the balance sheet and bank lending channels. Instead, we try to make the case for the broader view that allowing for a credit channel of some type is important for understanding the response of the economy to changes in monetary policy.

How the Economy Responds to Monetary Policy Shocks: Facts and Puzzles

To set the stage for our discussion of how monetary policy works, we first fill in some of the details about what happens in the economy after a change in monetary policy (a tightening, say) occurs. We do so by extending recent empirical work on the effects of monetary policy to consider its impact on some key components of GDP. We emphasize four basic facts about the response of the economy to monetary policy shocks.

Fact 1: Although an unanticipated tightening in monetary policy typically has only transitory effects on interest rates, a monetary tightening is followed by sustained declines in real GDP and the price level.

Fact 2: Final demand absorbs the initial impact of a monetary tightening, falling relatively quickly after a change in policy. Production follows final demand downward, but only with a lag, implying that inventory stocks rise in the short run. Ultimately, however, inventories decline, and inventory disinvestment accounts for a large portion of the decline in GDP.

Fact 3: The earliest and sharpest declines in final demand occur in residential investment, with spending on consumer goods (including both durables and non-durables) close behind.

\footnote{Both linkages have been extensively discussed in the literature. Surveys of related material include (among others) Bernanke (1993a), Kashyap and Stein (1994), Hubbard (1994) and Bernanke, Gertler and Gilchrist (forthcoming).}
Fact 4: Fixed business investment eventually declines in response to a monetary tightening, but its fall lags behind those of housing and consumer durables and, indeed, behind much of the decline in production and interest rates.

Responses to Policy Shocks: Evidence from Vector Autoregressions

These four facts are illustrated by Figures 1–3, which show the dynamic responses of various economic aggregates to an unanticipated tightening of monetary policy. Figures 1–3 are generated by the technique of "vector autoregression." A vector autoregression, or VAR, is a system of ordinary least-squares regressions, in which each of a set of variables is regressed on lagged values of both itself and the other variables in the set. VARs have proved to be a convenient method of summarizing the dynamic relationships among variables, since, once estimated, they can be used to simulate the response over time of any variable in the set to either an "own" disturbance (that is, a disturbance to the equation for which the variable is the dependent variable) or a disturbance to any other variable in the system. The VARS that we employ here include various combinations of macroeconomic variables and, additionally, the federal funds interest rate. Following Bernanke and Blinder (1992), Christiano, Eichenbaum and Evans (1994a,b) and others, we employ the federal funds rate as an indicator of the stance of monetary policy; this means that we identify the disturbances to the funds-rate equation in the VAR as shocks to monetary policy, and we interpret the responses of other variables in the system to a funds-rate shock as the structural responses of those variables to an unanticipated change in monetary policy. Because we are interested in observing the fine timing of responses to monetary shocks, we use monthly data. The sample period on which Figures 1–3 are based is January 1965 through December 1993 (subsample results are similar).

Figure 1 is based on a VAR system that includes the log of real GDP, the log of the GDP deflator, the log of an index of commodity prices and the federal funds rate (in percentage points), in that order. Real GDP and the GDP deflator are included as broad measures of economic activity and prices, and the commodity price index is intended to control for oil price shocks and other supply-side factors influencing output and

---

4 The use of VARs in macroeconomics was pioneered by Sims (1980); for a comprehensive recent discussion, see Watson (1994).

5 Bernanke and Blinder (1992) argue that the Fed has often used the funds rate, which is the interest rate prevailing in the market for bank reserves, as its primary policy indicator (particularly before 1979). Bernanke and Mihov (1995) estimate a model of the Fed's operating procedures and find that funds-rate targeting describes Fed behavior particularly well prior to 1979 and from 1988 to the present. In any case, the results discussed here are not dependent on using the funds rate as the monetary policy indicator; similar results are obtained when using reserves-based indicators (see, for example, Strongin, 1992) or indicators developed through historical analysis (for example, Romer and Romer, 1989).

6 We constructed monthly data for real GDP and the GDP deflator by interpolation methods, using a variety of monthly series to provide the within-quarter information. Bernanke and Mihov (1995) offer details. Results using noninterpolated monthly output and price data—for example, the industrial production index and the CPI (excluding shelter)—yield very similar results. Twelve lags of each variable and a constant term are included in each equation of the VAR.
inflation. The figure shows the estimated dynamic responses of log real GDP, the log of the GDP deflator and the funds rate to a positive, one-standard-deviation shock to the funds rate (which we interpret as an unanticipated tightening of monetary policy). As output and prices are measured in logs, the responses can be interpreted as proportions (that is, .001 = 0.1 percent) of baseline levels.

According to the estimated response patterns shown in Figure 1, GDP begins to decline about four months after a tightening of monetary policy, bottoming out about two years after the shock. The price level remains inert for about a year, then begins to decline, well after the drop in GDP begins. (The commodity price index, not shown, drops more quickly.) Finally, after rising sharply initially, the funds rate begins to fall precipitously after three to four months. Nine to 12 months after the policy innovation, the deviation of the funds rate from its baseline path is only a quarter or less of the initial shock; at two years out, the funds rate is essentially back to trend. These patterns are summarized above as Fact 1. These results in monthly data are quite consistent with the patterns observed by others in quarterly data and for various shorter sample periods. Although we do not show standard error bands in the figures, these responses and all that we report subsequently are statistically significant at conventional levels.

Sims (1992) and Christiano, Eichenbaum and Evans (1994b) discuss in detail the rationale for including the index of commodity prices; see also Sims and Zha (1995). Inclusion of this variable, along with measures of output and the general price level, has become conventional in the recent VAR-based literature on monetary policy.
Figure 2
Responses of Final Demand and Inventories to a Monetary Policy Shock

To allow us to look at the economy’s response to a monetary shock in closer detail, Figure 2 replaces log real GDP from the first VAR with two variables that sum to GDP, final demand and inventory investment. We measure both final demand and inventories relative to trend GDP, proxied by a six-year moving average of GDP; this normalization makes the magnitudes of the changes in the two variables comparable (both can be interpreted as fractions of trend GDP) and avoids taking the log of a series (inventory investment) that is sometimes negative. Note from Figure 2 that final demand drops quickly following an unanticipated tightening of monetary policy. In contrast, inventories build up for a period of several months before beginning to decrease, implying that the fall in final demand leads the decline in aggregate production (real GDP). The fall in inventories, when it occurs, appears to account for a substantial portion of the initial drop in output, which is consistent with Blinder and Maccini’s (1991) evidence on the importance of inventory disinvestment in recessions. These results are summarized as Fact 2.

Next, we explore what happens when we include various components of GDP in the VAR used in Figure 1. These series are added one at a time to the base VAR, although adding them in combinations gives very similar results. To make the magnitudes of changes comparable, the GDP components are also left in levels and normalized by trend GDP.\(^8\) Figure 3 shows the responses to a monetary contraction of

\(^8\) We continue to order the funds rate last, which has the effect of assuming that the Fed uses contemporaneous economic information, but that innovations in monetary policy do not feed back to the rest of the economy until the next month (Bernanke and Blinder, 1992; Bernanke and Mihov, 1995).
some important components of private domestic spending. As noted in Fact 3 above, residential investment drops sharply following a monetary tightening and accounts for a large part of the initial decline in final demand. Next in importance are consumer durables and nondurables, which also contribute significantly to the fall in final demand. (Nondurables react by much less in percentage terms than durables do, but they make a similar total contribution to the downturn owing to their larger share in overall economic activity.) Figure 3 shows that business fixed investment also declines following a monetary tightening, but with a greater lag than other types of spending (Fact 4). An interesting result (not shown in the figure) is that equipment investment accounts for nearly all of the decline in fixed investment; structures investment by businesses appears to respond very little to a monetary policy shock.

**Does the Conventional Story Fit the Facts?**

In a number of ways, the behavior of the economy shown in Figures 1–3, and as summarized by Facts 1–4, is consistent with the conventional analysis of monetary policy transmission. According to the standard story, the Fed has leverage over the short-term real rate because prices are sticky. In turn, the change in real rates affects aggregate demand. The slow response of the GDP deflator in the wake of the monetary contraction (Figure 1) and the quick response of final demand (Figure 2) are consistent with this scenario. Apparently, so too is the fact that durables spending—traditionally thought to be the most interest-sensitive part of aggregate demand—displays large responses to monetary policy shocks.

But there are some important puzzles. First among these is the magnitude of
the policy effect. We and many other researchers have found that the real economy is powerfully affected by monetary policy innovations that induce relatively small movements in open-market interest rates. But as we discussed earlier, empirical studies have not typically found commensurately strong cost-of-capital effects on the various components of private spending.

Second is the issue of timing. As shown by Figure 1, the interest-rate spike associated with an unanticipated monetary tightening is largely transitory; the federal funds rate is virtually back to trend at eight to nine months after the shock. Yet some important components of spending do not begin to react until after most of the interest-rate effect is past. For example, the bulk of the response of business fixed investment occurs during the period between six and 24 months after the shock (Figure 3).\(^9\) Inventories actually rise during the first three to four months after a monetary tightening; they begin to decline only in the period during which interest rates are falling sharply back to trend (Figure 2). The poor correspondence in timing between changes in interest rates and movements in some components of spending observed in Figures 1–3 no doubt helps to explain why robust effects of interest rates on spending have been hard to pin down empirically.

A third issue is the composition of the spending effects. Because monetary policy has its most direct effects on short-term rates, it would seem that it should have its most significant impact on spending on assets with shorter lives—some types of inventories and consumer durables, for example. Yet the most rapid (and in percentage terms, by far the strongest) effect of monetary policy is on residential investment (Figure 3). This finding is puzzling because residential investments are typically very long-lived and thus (according to the conventional view) should be most sensitive to long-term real interest rates, not the short-term rates most directly influenced by the Fed. At the same time, the other major type of long-lived investment, business structures investment, does not seem to be much affected by monetary policy actions. It is not immediately obvious why residential and business structures investment behavior should differ in this way.

The Credit Channel of Monetary Transmission

We have argued that it is difficult to explain the magnitude, timing and composition of the economy's response to monetary policy shocks solely in terms of conventional interest-rate (neoclassical cost-of-capital) effects. The mechanisms collectively known as the credit channel help to fill in the gaps in the traditional story.

Underlying our conception of the credit channel is the following basic premise: whenever frictions—such as imperfect information or costly enforcement of

---

\(^9\) Adjustment costs are an alternative potential explanation of the lagged response of investment to interest rates. However, most attempts to fit neoclassical investment equations with adjustment costs have estimated those costs to be implausibly large. If adjustment costs are important, there remains the question of why managers should respond at all to interest-rate fluctuations that they should expect to be temporary.
contracts—interfere with the smooth functioning of financial markets, we expect to observe a wedge between the cost of funds raised externally (for example, through the issuance of imperfectly collateralized debt) and the opportunity cost of internal funds. This wedge, which we call the external finance premium, reflects the deadweight costs associated with the principal-agent problem that typically exists between lenders and borrowers. Among the factors reflected in the external finance premium are the lender’s expected costs of evaluation, monitoring and collection; the “lemons” premium that results from the fact that the borrower inevitably has better information about its prospects than does the lender; and the costs of distortions in the borrower’s behavior that stem from moral hazard or from restrictions in the contract intended to contain moral hazard (for example, restrictive covenants or collateral requirements).

According to advocates of the credit channel, monetary policy affects not only the general level of interest rates, but also the size of the external finance premium. This complementary movement in the external finance premium may help explain the strength, timing and composition of monetary policy effects better than is possible by reference to interest rates alone. Two mechanisms have been suggested to explain the link between monetary policy actions and the external finance premium: the balance sheet channel (sometimes called the net worth channel) and the bank lending channel. In describing how these channels work, in the rest of this section, we focus on the behavior of firms. Subsequently, we will discuss how the analysis may be extended to consumer behavior, like spending on housing and consumer durables.

The Balance Sheet Channel

The balance sheet channel is based on the theoretical prediction that the external finance premium facing a borrower should depend on borrower’s financial position. In particular, the greater is the borrower’s net worth—defined operationally as the sum of her liquid assets and marketable collateral—the lower the external finance premium should be. Intuitively, a stronger financial position (greater net worth) enables a borrower to reduce her potential conflict of interest with the lender, either by self-financing a greater share of her investment project or purchase or by offering more collateral to guarantee the liabilities she does issue. This basic insight underlies many real-world financial arrangements, such as the requirement that borrowers meet certain financial ratios, that they post collateral and/or that they make down payments.

Since borrowers’ financial positions affect the external finance premium, and thus the overall terms of credit that they face, fluctuations in the quality of borrowers’ balance sheets similarly should affect their investment and spending decisions. An extensive theoretical literature has exploited this idea to argue that endogenous procyclical movements in borrower balance sheets can amplify and propagate business cycles, a phenomenon that has been referred to as the “financial accelerator.” This approach has been supported by a wide range of empirical work linking balance sheet and cash flow variables to firms’ decisions concerning fixed investment, inventories and other factor demands, and to household purchases of durables and housing. Bernanke, Gertler and Gilchrist (forthcoming) offer a discussion and references.
The balance sheet channel of monetary policy arises because shifts in Fed policy affect not only market interest rates per se but also the financial positions of borrowers, both directly and indirectly.

For example, a tight monetary policy directly weakens borrowers' balance sheets in at least two ways. First, to the extent that borrowers have outstanding short-term or floating-rate debt, rising interest rates directly increase interest expenses, reducing net cash flows and weakening the borrower's financial position. Because many firms rely heavily on short-term debt to finance inventories and other working capital, this direct effect of monetary policy on business net cash flows is quite important quantitatively, as we show below.\(^\text{10}\) Second, rising interest rates are also typically associated with declining asset prices, which among other things shrink the value of the borrower's collateral. For example, many observers would agree that the crash of Japanese land and equity values in the latter 1980s was the result (at least in part) of monetary tightening; and that this collapse in asset values reduced the creditworthiness of many Japanese corporations and banks, contributing to the ensuing recession. Borio, Kennedy and Prowse (1994) argue that a similar pattern of asset-price boom and bust leading to real fluctuations occurred during the 1980s in a number of major industrialized countries.

Tight monetary policy may also reduce net cash flows and collateral values indirectly. Consider a manufacturing firm that produces for downstream customers. If a monetary tightening reduces spending by those customers (either for cost-of-capital or balance sheet reasons), the firm's revenues will decline while its various fixed or quasi-fixed costs (including interest and wage payments) do not adjust in the short run. The resulting increase in the "financing gap" (the difference between the firm's uses and sources of funds) erodes the firm's net worth and creditworthiness over time. As we discuss later, this mechanism may explain why the impact of the credit channel on spending, particularly on inventory and investment spending, may persist well beyond the period of the initial monetary tightening.

Is there evidence that links monetary policy to the financial positions of borrowers? For firms, direct high-frequency measures of balance sheet condition are hard to obtain (a major problem is that reporting is done in book rather than market values). A common and useful summary measure of a firm's financial condition is the "coverage ratio," the ratio of interest payments by nonfinancial corporations to the sum of interest payments and profits.\(^\text{11}\) The coverage ratio has the advantage of being easy to measure, and there is good reason to believe that it is highly correlated with the other indicators of firms' financial health. Further, a number of studies have found that for firms with imperfect access to credit, movements in the coverage ratio have a significant effect on factor demands.

\(^\text{10}\) This rise in borrowers' interest expenses is, of course, "only" a redistribution from borrowers to lenders. As discussed in the literature on debt deflation (Bernanke and Gertler, 1989), such redistributions are not macroeconomically neutral if lenders are relatively unconstrained in credit markets or if they do not have access to the same investment and spending opportunities as do borrowers.

\(^\text{11}\) Actually, the coverage ratio is more usually defined as the inverse of the way we define it. However, we believe that our definition makes for a more intuitive discussion and graphical presentation.
Figure 4
The Funds Rate and the Coverage Ratio

Figure 4 plots the coverage ratio along with the federal funds rate. The close relationship of the two variables is evident from the figure. If we interpret the direction of causality as running from the funds rate to the coverage ratio, which is most plausible, it seems that increases in the funds rate translate almost immediately into increases in the coverage ratio and hence, ultimately, into weaker balance sheet positions. The coverage ratio rises because of both the direct and indirect effects of monetary policy on firms' financial positions. The rise in the funds rate, which affects other short-term rates, directly increases interest expenses (the numerator of the coverage ratio); at the same time, the impact of tightening monetary policy on final demand, and thus on firm revenues, reduces profits (which are included in the denominator). Another important factor is that short-term borrowing rises following a monetary tightening (in large part to finance the inventory buildup; see Figure 2), which further increases interest expenses.\(^\text{12}\)

Figure 5, which follows closely the analysis of Galati (1994), gives a more complete breakdown of the impact of monetary policy on the components of corporate cash flows, using the same type of VAR technique employed in the previous section. This VAR uses quarterly data for the period 1965:I to 1994:II and includes four lags of each variable. Each of the components of cash flow is normalized by the six-year moving average of gross income, so that magnitudes of changes in the figure are

comparable. The figure shows the dynamic responses of the interest payments, profits, gross income and employee compensation of nonfinancial corporate business to a positive innovation of one standard deviation in the funds rate (a tightening of monetary policy).\textsuperscript{13}

The picture illustrates the two reasons that corporate cash flows and profits decline following a tightening of monetary policy. First, increases in interest payments directly reduce profits; Figure 5 suggests that something over 40 percent of the short-run decline in corporate profits is the result of higher interest payments. Second, Figure 5 shows that, following a monetary shock, corporate income tends to fall more quickly than costs (such as employee compensation), which tend to be squeezed during a period of monetary tightening. Interestingly, the cash squeeze appears to peak about six to nine months after the monetary tightening, about the time that output, inventories and investment begin to decline (see Figures 1–3). Interest expenses remain high for several more quarters even though short-term interest rates themselves are dropping, because of the rise in short-term debt outstanding. These results are consistent with empirical findings that cash flow and other "accelerator" variables are important for inventory accumulation and investment spending.

The effects of the corporate cash squeeze on economic behavior depend largely on firms' ability to smooth the drop in cash flows by borrowing. Firms that have relatively poor access to credit markets may have to respond to declining cash flows by cutting production and employment, while firms with good access to credit will face less financial pressure. The differential impact of a cash squeeze on different types of firms has been studied by Gertler and Gilchrist (1993, 1994), who considered how a tightening of monetary policy affected the inventories and short-term debt of large and small manufacturing firms.

Gertler and Gilchrist (1993, 1994) found striking differences in behavior between large and small firms. Larger firms, who are more likely to have recourse to commercial paper markets and other sources of short-term credit, typically respond to an unanticipated decline in cash flows by increasing their short-term borrowing. The inventories of large firms grow following a tightening of monetary policy, suggesting that these firms are at least temporarily able to maintain their levels of production and employment in the face of higher interest costs and declining revenues.\textsuperscript{14} In contrast, small firms—who in most cases have more limited access to short-term credit markets—respond to the cash flow squeeze principally by decumulating inventories, that is, by cutting work-hours and production. Apparently, small firms are not able to increase short-term borrowing.\textsuperscript{15} Oliner and Rudebusch

\textsuperscript{13} In the results shown, we ordered interest payments and profits after the funds rate, in the belief that monetary policy would affect those variables within the quarter. The results are qualitatively the same when the funds rate is ordered last, except that the movement in interest payments is somewhat smaller.

\textsuperscript{14} This is not to deny the possibility that at certain times even large firms may have imperfect access to credit. For example, recent work by Carpenter, Fazzari and Petersen (1994) suggests that, while cash flow effects on inventory investment during the 1981–82 recession were stronger for small firms than for large firms, these effects were also present for large firms.

\textsuperscript{15} The relative decline in loans to small firms is not offset by large firms supplying increased trade credit
Figure 5
Responses of Corporate Cash Flows to a Monetary Policy Shock

(1994) have shown that analogous differences occur between size classes in the response of fixed investment to a monetary policy shock.

A spirited debate has developed that offers several competing explanations for Gertler and Gilchrist’s (1993, 1994) results and related results by others. We cannot do justice to that debate here, but two brief points are worth mentioning. First, it is difficult to account for the Gertler-Gilchrist findings by appealing to differences in technologies or other nonfinancial aspects of small firms. The observed differences in inventory and borrowing behavior between the size classes of firms remain after controlling for the growth of sales; that is, the differential behavior is not explained simply by differences in demand. In addition, these differences between large and small firms arise mainly in recessions and in tight-money periods, which is consistent with a prediction of the theory that the response of the external finance premium to financial conditions should be greatest during recessions, when liquidity constraints are likely to bind across a broader cross-section of firms; during booms, small firms appear to smooth production in much the same way that large firms do. Finally, using firm-level data, Bernanke, Gertler and Gilchrist (forthcoming) show that the results survive when industry dummies are included and when the sample is split by financial criteria (such as dependence on banks or access to the commercial paper market) rather than size.

to small firms (Gertler and Gilchrist, 1993). In any case, the terms of trade credit are typically so unfavorable as to make it a poor alternative to bank loans and other forms of short-term credit.
A second point concerns the macroeconomic significance of the Gertler and Gilchrist findings. It is true that "small" firms (defined by Gertler and Gilchrist to be firms with assets under $250 million in 1991 dollars) are not dominant in manufacturing, accounting for about 30 percent of aggregate sales. Nevertheless, it appears that small firms account for a significantly disproportionate share of inventory fluctuations in manufacturing (Gertler and Gilchrist, 1994; Kashyap, Lamont and Stein, 1994). In other sectors of the economy—such as retailing, wholesaling, construction and services—small firms make up a much larger part of the total, increasing the likelihood that credit effects are important. Thus it is not the case that small-firm behavior is of trivial significance to the economy as a whole.

The Bank Lending Channel

Beyond its impact on borrowers' balance sheets, monetary policy may also affect the external finance premium by shifting the supply of intermediated credit, particularly loans by commercial banks. This is the bank lending channel. Banks, which remain the dominant source of intermediated credit in most countries, specialize in overcoming informational problems and other frictions in credit markets. If the supply of bank loans is disrupted for some reason, bank-dependent borrowers (small and medium-sized businesses, for example) may not be literally shut off from credit, but they are virtually certain to incur costs associated with finding a new lender, establishing a credit relationship and so on. Therefore, a reduction in the supply of bank credit, relative to other forms of credit, is likely to increase the external finance premium and to reduce real activity.

That banks play an important role in overcoming informational problems in credit markets, and that as a result many borrowers are substantially bank-dependent, seems well established. The more controversial question about the bank lending channel is whether monetary policy can significantly affect the supply (or relative pricing) of bank loans. Bernanke and Blinder's (1988) model of the bank lending channel suggested that open market sales by the Fed, which drain reserves and hence deposits from the banking system, would limit the supply of bank loans by reducing banks' access to loanable funds. This effect, transmitted through the level and composition of bank assets, is over and above the traditional (IS-LM) money supply and interest-rate effects, which are reflected in declining bank liabilities.

A key assumption of the Bernanke and Blinder (1988) model is that banks cannot easily replace lost (retail) deposits with other sources of funds, such as certificates of deposit (CDs) or new equity issues. Arguably, this assumption was a good one for the United States prior to 1980, for several reasons. First, the so-called

---

16 For example, two recent studies of micro data on retail sector firms suggest that borrowing constraints have a large impact on the overall behavior of inventories, investment and employment in that sector (Zakrajsek, 1994; Kwon, 1995).

17 Additionally, the Bernanke and Blinder (1988) model must assume that loans and securities are imperfect substitutes in bank portfolios, so that banks are not willing to absorb deposit losses completely by reducing securities holdings. This assumption is empirically plausible.
"Regulation Q," imposed by the Federal Reserve, placed a ceiling on the interest rates banks could pay; when open-market interest rates went above the ceiling, banks had no means of competing for funds and indeed suffered sharp declines in deposits (the "disintermediation" phenomenon). Second, reserve requirements were much more onerous at that time than they are today; as late as 1980 the Federal Reserve imposed marginal reserve requirements on large CDs, hampering banks' ability to raise loanable funds. Third, markets for bank liabilities were much less liquid and less developed then than now. Thus, until about 15 years ago, the institutional facts fit the assumptions of the Bernanke-Blinder model quite well.

However, since about 1980, as emphasized by Romer and Romer (1990), banks' ability to raise funds on the margin—in particular, through issuance of large CDs and other "managed liabilities"—has become less restricted. Regulation Q has been removed. Markets for bank liabilities have greatly deepened, although new equity issues remain (as they have always been) a costly source of funds for banks, particularly in periods of economic stress. Reserve requirements have been eliminated for most bank liabilities, with the exception of pure transactions accounts. Clearly, the Bernanke and Blinder (1988) model is a poorer description of reality than it used to be, at least in the United States.

Nevertheless, the existence of a bank lending channel does not require banks to be totally incapable of replacing lost deposits. As emphasized by Kashyap and Stein (1994) among others, it is sufficient that banks do not face a perfectly elastic demand for their open-market liabilities, so that an open-market sale by the Fed—which shrinks banks' core deposit base and forces them to rely more on managed liabilities—also increases banks' (relative) cost of funds. An increase in the cost of funds to banks should shift the supply of loans inward, squeezing out bank-dependent borrowers and raising the external finance premium.

Despite financial reform and innovation, it remains likely that the demand for banks' managed liabilities is not perfectly elastic: For example, large CDs are incompletely protected by deposit insurance, which raises the usual issues of evaluation and monitoring by lenders. Indeed, small and poorly capitalized banks typically cannot issue large CDs at all. Further, the market for CDs is still not especially deep or liquid relative to some other public markets—as shown by the fact that many CDs are nonnegotiable or difficult to trade on secondary markets—so that new investors can be called forth only by paying higher rates of interest.

Figure 6 presents some data relevant to the bank lending channel. The variable entitled "terms of lending" shows the net percentage of small businesses who reported to the National Federation of Independent Businesses that credit was harder to get relative to the previous quarter. Also shown in Figure 6 are three interest-rate spreads: the difference between the bank CD rate and the T-bill rate, the spread between the federal funds rate and the T-bill rate and the prime rate–T-bill rate spread, all in quarterly averages. The vertical lines indicate business cycle peaks.

The behavior of interest-rate spread and terms of lending shown in Figure 6 offers some support for predictions of the bank lending channel. During tight-money periods, when open-market interest rates rise, the prime rate rises by even more (that is, the prime rate–safe rate spread opens up), and credit terms
become more onerous.\textsuperscript{18} CD interest rates increase by significantly more than the T-bill rate during a monetary tightening (so that the CD spread and funds-rate spread are highly correlated), consistent with our claim that the demand for banks' managed liabilities is not perfectly elastic.\textsuperscript{19} Note that a particularly sharp spike in the spread between CD rates and T-bill rates occurred in 1981, a period in which monetary policy was restrictive but reserve requirements on CDs were no longer in force.

But while the behavior of interest-rate spreads and terms of lending are consistent with the bank lending channel as conceived by Bernanke and Blinder, it must be noted that they are also potentially consistent with the operation of a balance sheet channel. Since a tightening of monetary policy leads to a worsening of both borrowers' and banks' balance sheets, balance sheet effects alone could conceivably explain why borrowing becomes more expensive and difficult, and why banks must promise higher interest to sell their CDs. It is extremely difficult to carry out an empirical test that would conclusively separate the bank lending channel from the balance sheet channel. For this reason, we are more confident in the existence of a credit channel in general than we are in our ability to distinguish sharply between the two mechanisms of the credit channel.

In summary, because of financial deregulation and innovation, the importance of the traditional bank lending channel has most likely diminished over time. While we believe this channel is still empirically relevant, obtaining sharp measurements of its potency is a challenging task. In any case, however, this framework may still be of value for interpreting historical data, for assessing the impact of institutional changes on the transmission of monetary policy and for comparing monetary transmission mechanisms across countries. Additionally, as we have already suggested, it may well be the case that the balance sheet channel—described earlier in reference to non-financial borrowers—has become increasingly relevant to banks as well (Kashyap and Stein, forthcoming). For example, if rising interest rates lower the value of securities, impairing banks' capital, then the ability of banks to attract funds and hence their capacity to make loans may be reduced. During the recent peso crisis, a number of commentators cited the exposure of the capital of Mexican banks to rising interest rates as a key concern; similarly, the Swedish devaluation of a few years ago was attributed in part to fears of the effects that rising interest rates would have on Sweden's banking system. In the United States, increased bank holdings of volatile securities and derivative instruments may have increased the sensitivity of bank lending to interest rates via the balance sheet channel. The health of the banking system remains intertwined with the performance of the economy as a whole.

\textsuperscript{18} The increase in the prime spread understates the increase in cost to the typical bank borrower, as fewer borrowers qualify for the prime rate during tight-money periods.

\textsuperscript{19} Changes in default risk cannot plausibly account for these differences, as realized rates of default tend to be much smaller in magnitude than the spreads. Indeed, the aggregate data underestimate the increase in interest rates faced by a typical borrower, as lenders discriminate heavily toward stronger borrowers during recessions and tight-money periods; see Bernanke, Gertler and Gilchrist (forthcoming).
Evidence from the Forecasting Power and Cyclical Timing of Credit Aggregates: A Critique

The empirical literature on the credit channel is large, and we cannot do justice to it here. While this literature (at both the macro and micro levels) generally supports the existence of a credit channel, more work needs to be done to establish the credit channel's quantitative significance in monetary transmission and to determine the relative importance (at particular points in time) of the balance sheet and bank lending channels.

However, one popular approach to "testing" the credit channel seems to us generally invalid and so merits a brief comment. This approach has focused on evaluating the forecasting power of credit aggregates (ranging from total private credit to bank commercial-industrial loans) for real activity, often relative to the forecasting power of money aggregates. Such work can be useful for some purposes, but as a test for the credit channel it suffers from the incorrect premise that this approach treats credit aggregates as an independent causal factor affecting the economy. Except in rare circumstances,\(^{20}\) credit is not a primitive driving force; rather, credit conditions—best measured, by the way, by the external finance premium and not the aggregate quantity of credit—are an endogenous factor that help shape the dynamic response of the economy to shifts in monetary policy. Thus the theory has no particular implications about the relative forecasting power of

---

\(^{20}\) These could be circumstances such as a spontaneous financial crisis whose effects spill over onto the real economy. This scenario may be relevant to the Great Depression, for example, but is not relevant under "normal" conditions.
credit aggregates. Indeed, as discussed further by Bernanke (1993b), even the extreme finding that credit aggregates have no marginal predictive power for real variables, once measures of the stance of monetary policy are accounted for, is not necessarily inconsistent with an important role for the credit channel in propagating monetary disturbances.

Along similar lines, examining the dynamic responses of various credit aggregates to shifts in monetary policy (or comparing these responses to those of monetary aggregates) is also largely uninformative about the existence or importance of the credit channel. Credit aggregates, like monetary aggregates, are determined jointly by supply and demand. In particular, while money demand is procyclical, credit demand appears to contain a significant countercyclical component, which arises from the desire of households and firms to smooth the impact of cyclical variations in income on spending or production. For example, as we have seen, firms often attempt to finance an inventory buildup following a monetary tightening, which means that they may attempt to increase their borrowing early in the cycle, even though the external finance premium is rising. Similarly, consumers may want to borrow to make up for shortfalls in income. As a result, at the same time that credit conditions (as measured by the external finance premium) are worsening, aggregate credit may continue to rise. Thus the countercyclical demand for credit helps to explain the common finding that credit aggregates are slower to respond to a change in Fed policy than are monetary aggregates. These “perverse” short-run movements in credit aggregates are not inconsistent with the theory of the credit channel, which does not require that firms and households literally reduce their borrowing following a monetary tightening; the theory predicts only that they will borrow and hence spend less than they would if credit markets were perfect.

**Housing and Other Consumer Expenditures**

So far we have focused on the behavior of firms, as has most of the literature on the credit channel. However, the credit market frictions that affect firms should also be relevant to the borrowing and spending decisions made by households, particularly spending on costly durable items such as automobiles and houses. Thus, the workings of the credit channel may help explain the high sensitivity of residential investment and consumer durables to monetary policy shocks that was shown in Figure 3.

For example, it is likely that both the bank lending channel and the balance sheet channel have at various times affected residential investment (which, remember, accounts for the largest fraction of the decline in economic activity immediately following a monetary tightening). The bank lending channel was probably most relevant in the period before 1980, when ceilings on deposit rates were in effect. We have already mentioned the phenomenon of disintermediation, where high market interest rates in the 1960s and 1970s sometimes led deposits to flee banks,

---

21 Notwithstanding the fact that, in a previous life, Bernanke has performed similar exercises. Mea culpa.
thus disrupting the flow of mortgage credit. A number of formal econometric studies have shown that including proxies for the impact of disintermediation improves the fit of equations explaining residential investment during the Regulation Q period (for example, Ryding, 1990). Both informal and formal studies suggest that the impact of monetary policy on housing (and, perhaps, on the overall economy) has weakened since the phasing out of interest rate ceilings and the introduction of innovations such as a liquid secondary market for mortgages.

There are also natural connections between the balance sheet channel and residential investment. Housing demand is linked directly to the condition of consumer balance sheets by features like down-payment requirements, up-front transactions costs like closing costs and "points," and minimum income-to-interest-payment ratios (see Stein, 1995b, for an analysis of down-payment constraints on the dynamics of the housing market). In an interesting empirical study of housing demand, Boldin (1994) constructed a measure of the pressure that purchasing a home puts on the typical household's balance sheet. His "mortgage burden" variable, which is (approximately) the ratio of mortgage payments to income for the median new home buyer, can be considered analogous to the coverage ratio for firms. Boldin found that the mortgage burden significantly influences housing demand, after controlling for both conventional neoclassical factors and variables that capture disintermediation effects (as in Ryding's study).

In fact, Boldin's mortgage burden variable has a close positive correlation with the federal funds rate, which we have used as a proxy for changes in monetary policy (Figure 7). The correlation arises both because nominal interest rates rise and household incomes fall following a tightening of monetary policy. The funds rate is also positively correlated with indicators of up-front housing costs, such as the average number of "points" charged on mortgages. The effects of monetary policy on such variables as the mortgage burden and mortgage terms help explain its strong impact on housing demand, despite the presumably weak link between monetary policy and long-term real interest rates. Most likely, these balance-sheet effects have not been eliminated in the recent era of deregulation and innovation, although it would be interesting to explore the implications of the introduction of adjustable-rate mortgages, for example, in this context.

Although we have emphasized housing, the balance sheet channel likely applies to other expenditures on durable goods by households. (Important early research on the topic was done by Mishkin, 1978; an interesting recent study is by Chah, Ramey and Starr, 1995.) An important goal of future research should be to give the role of consumers' balance sheets in monetary policy transmission the same attention that has been paid to the balance sheets of corporations.

Conclusion: The Role of the Credit Channel in Monetary Transmission

As we have pointed out, the conventional analysis of monetary transmission has a number of shortcomings. However, introducing the credit channel potentially helps resolve these puzzles.
First, while monetary policy clearly has a strong impact on durable goods spending, most research has had difficulty identifying significant cost-of-capital effects of the pure neoclassical type. However, conventional cost-of-capital measures rarely make allowance for a variable external finance premium. To be sure, obtaining direct measures of the external finance premium is difficult. Yet numerous recent studies have shown that including proxies or indicators for this premium in durables demand equations greatly improves their performance (as one example, see Kashyap, Lamont and Stein, 1994, on the demand for inventory stocks).

A second shortcoming of the textbook analysis is that, after a tightening of monetary policy, much of the decline in both inventories and nonresidential investment occurs with a lag—in particular, after interest rates have begun to fall back to baseline levels. Accounting for variation in the external finance premium may be helpful in explaining both the timing of the response of inventories and investment and the fact that these variables seem better explained by sales or cash flows than by interest rates. For example, following a tightening of monetary policy, firms' balance sheets deteriorate as interest rates rise, cash flows decline and short-term borrowings increase to fund inventory accumulation, payrolls and so on. The sustained weakening of balance sheets implies that the premium for external finance (and therefore the cost of capital, broadly defined) may be continuing to rise as the recession sets in, even though market interest rates are falling. This buildup in "financial pressure" helps explain the delayed but sharp drop in inventories and investment.

A third problem for the standard view is the large and rapid response of housing investment to changes in monetary policy, despite the fact that monetary policy
presumably has at best limited control over real long-term interest rates. Yet again, the solution may be to allow for movements in the external finance premium that result from the change in policy. As we discussed in this article, both the balance sheet and bank lending channels have played important roles in the housing market. Quantitatively successful analyses of monetary policy transmission should (and in many cases, already do) include these mechanisms.

- The authors wish to thank Timothy Taylor, Rick Mishkin, Alan Auerbach, Michael Boldin, Jordi Gali, Anil Kashyap and Jeremy Stein for helpful comments. The National Science Foundation and the C.V. Starr Center provided financial support.

References


