Distinguishing Theories of the Monetary Transmission Mechanism

Stephen G. Cecchetti

Traditional studies of monetary policy's impact on the real economy have focused on its aggregate effects. Beginning with Friedman and Schwartz (1963), modern empirical research in monetary economics emphasizes the ability of policy to stabilize the macroeconomy. But casual observation suggests that business cycles have distributional implications as well. One way of casting the debate over the relative importance of different channels of monetary policy transmission is to ask if these distributional effects are sufficiently important to warrant close scrutiny.

The point can be understood clearly by analogy with business cycle research more generally. If recessions were characterized by a proportionate reduction of income across the entire employed population—for example, everyone worked 39 rather than 40 hours per week for a few quarters—then economists would pay substantially less attention to cycles. It is the allocation of the burden or benefit of fluctuations, with some individuals facing much larger costs than others, that is of concern. There are two ways for an economist to address this problem. The first is to attempt to stabilize the aggregate economy, the traditional focus of policy-oriented macroeconomics. The second is to ask why the market does not provide some form of insurance.

The recent debate over the nature of the monetary transmission mechanism can be thought of in similar terms. According to the original textbook IS-LM view of money, changes in policy are important only insofar as they affect aggregate outcomes. Only the fluctuation in total investment is important since policies only affect the required rate of return on new investment projects, and so it is only the least profitable projects (economy-wide) that are no longer funded. But since the most profitable projects continue to be undertaken, there are no direct efficiency losses associated with the distributional aspects of the policy-induced interest rate increase.

In contrast, the "lending" view focuses on the distributional consequences of monetary policy actions. By emphasizing a combination of capital market imperfections and portfolio balance effects based on imperfect asset substitutability, this alternative theory suggests the possibility that the policy's incidence may differ substantially across agents in the economy. Furthermore, the policy's impact has to do with characteristics of the individuals that are unrelated to the inherent creditworthiness of the investment projects. An entrepreneur may be deemed unworthy of credit simply because of a currently low net worth, regardless of the social return to the project being proposed. It is important to understand whether the investment declines created by monetary policy shifts have these repercussions.[1]

In this essay, I examine how one might determine whether the cross-sectional effects of monetary policy are quantitatively important. My goal is to provide a critical evaluation of the major contributions to the literature thus far. The discussion proceeds in three steps. I start in the first section with a description of a general framework that encompasses all views of the transmission mechanism as special cases, thereby highlighting the distinctions. In the second section, I begin a review of the empirical evidence with an assessment of how researchers typically measure monetary policy shifts. The following two sections examine the methods

[1] The financial accelerator, in which the impact on investment of small interest changes is magnified by balance sheet effects, is also an important part of many discussions of the lending view.

used for differentiating between the theories. Studies fall into two broad categories depending on whether they use aggregate or disaggregate data. The third section discusses the aggregate data, while the fourth section describes the use of disaggregate data. A conclusion follows.

MONETARY POLICY: THEORY

A General Framework

One way of posing the fundamental question associated with understanding the monetary transmission mechanism is to ask how seemingly trivial changes in the supply of an outside asset can create large shifts in the gross quantity of assets that are in zero net supply. How is it that small movements in the monetary base (or nonborrowed reserves) translate into large changes in demand deposits, loans, bonds and other securities, thereby affecting aggregate investment and output?

The various answers to this puzzle can be understood within the framework originally proposed by Brainard and Tobin (1963). Their paradigm emphasizes the effects of monetary policy on investor portfolios, and is easy to present using the insights from Fama's (1980) seminal paper on the relationship between financial intermediation and central banks.

Fama's view of financial intermediaries is the limit of the current type of financial innovation, because it involves the virtual elimination of banks as depository institutions. The setup focuses on an investor's portfolio problem in which an individual must choose which assets to hold given the level of real wealth. Labeling the portfolio weight on asset i as w_i, and total wealth as W, then the holding of asset, i—the asset demand—is just $X_i = w_i W$.

In general, the investor is dividing wealth among real assets—real estate, equity and bonds—and outside money. Each asset has stochastic return, $\tilde{z}_i$, with expectation $\bar{z}_i$; and the vector of asset returns, $\tilde{z}$, has a covariance structure Γ. Given a utility function, as well as a process for consumption, it is

possible to compute the utility maximizing portfolio weights. These will depend on the mean and variance of the returns, $\bar{z}_i$ and Γ, the moments of the consumption process, call these μ_c, and a vector of taste parameters that I will label ϕ and assume to be constants. The utility maximizing asset demands can be expressed as $X_i^* = w_i^*(\bar{z}, \Gamma, \mu_c, \phi)W$.[2]

This representation makes clear that asset demands can change for two reasons. Changes in either the returns process ($\bar{z}, \Gamma$) or macroeconomic quantities (μ_c, W) will affect the X_i^*'s.[3]

At the most abstract level, financial intermediaries exist to carry out two functions. First, they execute instructions to change portfolio weights. That is, following a change in one or all of the stochastic processes driving consumption, wealth or returns, the intermediary will adjust investors' portfolios so that they continue to maximize utility. In addition, if one investor wishes to transfer some wealth to another for some reason, the intermediary will effect the transaction.

What is monetary policy in this stylized setup? For policy to even exist, some government authority, such as a central bank, must be the monopoly supplier of a nominally denominated asset that is imperfectly substitutable with all other assets. I will call this asset "outside money." In the current environment, it is the monetary base. There is a substantial literature on how the demand for outside money arises endogenously in the context of the type of environment I have just described.[4] But in addition, as Fama emphasizes, there may be legal requirements that force agents to use this particular asset for certain transactions. Reserve requirements and the use of reserves for certain types of bank clearings are examples.

Within this stylized setup, a *policy action* is a change in the nominal supply of outside money. For such a change to have any effects at all, (1) the central bank controls the supply of an asset that is both in demand and for which there is no perfect substitute, and (2) prices must fail to adjust fully and instantaneously. Otherwise, a change in the nominal quantity of outside money cannot have any impact on the real interest rate, and will have no real effects. But, assuming that

[2] See Ingersoll (1987) for a complete description of this problem.

[3] Following the traditional financial economics approach, I have avoided discussing demand and supply explicitly. Instead, the asset demands are derived from the idea of arbitrage relationships among all of the assets.

[4] These include the limited participation models based on Lucas (1990). See the survey by Feurst (1993), as well as the summary in Christiano and Eichenbaum (1992).

the policymaker can change the real return on the asset that is monopolistically supplied, investors' portfolio weights must adjust in response to a policy change.

The view of financial intermediaries that is implicit in this description serves to highlight the Brainard and Tobin (1963) insight that monetary policy can be understood by focusing solely on the endogenous response of investor portfolios. Understanding the transmission mechanism requires a characterization of how asset holdings change in response to policy actions.

Second, even though there need be no banks as we know them, there will surely be intermediaries that perform the service of making small business loans. The agency costs and monitoring problems associated with this type of debt will still exist, and specialists in evaluation will emerge. While they will have such loans as assets, they most likely will not have bank deposits as liabilities. Such entities will be brokers, and the loans will be bundled and securitized.

With this as background, it is now possible to sketch the two major views of the monetary transmission mechanism. There are a number of excellent surveys of these theories, including Bernanke (1993a), Gertler and Gilchrist (1993), Kashyap and Stein (1994a) and Hubbard (1995). As a result, I will be relatively brief in my descriptions.

THE MONEY VIEW

The first theory, commonly labeled the *money* view, is based on the notion that reductions in the quantity of outside money raise real rates of return.[5] This, in turn, reduces investment because fewer profitable projects are available at higher required rates of return—this is a movement along a fixed marginal efficiency of investment schedule. The less substitutable outside money is for other assets, the larger the interest rate changes.

There is no real need to discuss banks in this context. In fact, there is no reason to distinguish any of the "other" assets in investors' portfolios. In terms of the simple portfolio model, the money view implies

that the shift in the w_i^*'s for all of the assets excluding outside money are equal.

An important implication of this traditional model of the transmission mechanism involves the incidence of the investment decline. Since there are no externalities or market imperfections, it is only the least socially productive projects that go unfunded. The capital stock is marginally lower. But, given that a decline is going to occur, the allocation of the decline across sectors is socially efficient.

This theory actually points to a measure of money that is rarely studied. Most empirical investigations of monetary policy transmission focus on M2, but the logic of the portfolio view suggests that the monetary base is more appropriate. It is also worth pointing out that investigators have found it extremely difficult to measure economically significant responses of either fixed or inventory investment to changes in interest rates that are plausibly the result of policy shifts. In fact, most of the evidence that is interpreted as supporting the money view is actually evidence that fails to support the lending view.

THE LENDING VIEW: BALANCE SHEET EFFECTS

The second theory of monetary transmission is the *lending* view.[6] It has two parts, one that does not require introduction of assets such as bank loans, and one that does. The first is sometimes referred to as the *broad* lending channel, or financial accelerator, and emphasizes the impact of policy changes on the balance sheets of borrowers. It bears substantial similarity to the mechanism operating in the money view, because it involves the impact of changes in the real interest rate on investment.

According to this view, there are credit market imperfections that make the calculation of the marginal efficiency of investment schedule more complex. Due to information asymmetries and moral hazard problems, as well as bankruptcy laws, the state of a firm's balance sheet has implications for its ability to obtain external finance. Policy-induced increases in interest rates (which are both real and nominal) can cause a deterioration in

[5] Terminology has the potential to create confusion here. I have chosen the traditional term for this textbook IS-LM or "narrow" money view. I do not mean to imply that this is the "monetarist" view of the transmission mechanism.

[6] I follow Kashyap and Stein's (1994a) terminology rather than the more common credit view to emphasize the importance of loans in the transmission mechanism. Bernanke and Gertler (1989, 1990) provide the original theoretical underpinnings for this view.

[7] Bernanke, Gertler and Gilchrist (1994) refer to this as a financial accelerator since it causes small changes in interest rates to have potentially large effects on investment and output.

[8] It may be particularly difficult to distinguish these effects from those that arise from varying cyclicality of different firms' sales and profitability.

[9] See James (1987) for a discussion of the uniqueness of bank loans.

[10] With nominal rigidity, a decrease in outside money reduces the price level slowly, and so the real return to holding money increases. This channel of transmission requires that investors shift away from loans in response.

[11] Kashyap and Stein (1994b) point out that large banks can issue CDs in a way that insulates their balance sheets from contraction in deposits, but small banks cannot. So long as small banks are an important source of funds for some bank-dependent firms, there will still be a *bank* lending channel. In other words, for bank lending to be an important part of the transmission mechanism, credit market imperfections must be important for banks.

the firm's net worth, by both reducing expected future sales and increasing the real value of nominally denominated debt. With lower net worth, the firm is less creditworthy because it has an increased incentive to misrepresent the riskiness of potential projects. As a result, potential lenders will increase the risk premium they require when making a loan. The asymmetry of information makes internal finance of new investment projects cheaper than external finance.

The balance sheet effects imply that the shape of the marginal efficiency of investment curve is itself a function of the debt-equity ratio in the economy and can be affected by monetary policy.[7] In terms of a simple textbook analysis, policy moves both the IS and the LM curves. For a given change in the rate of return on outside money (which may be the riskless rate), a lender is less willing to finance a given investment the more debt a potential borrower has. This points to two clear distinctions between the money and the lending views—the latter stresses both the distributional impact of monetary policy and explains how seemingly small changes in interest rates can have a large impact on investment (the financial accelerator).

Returning to the portfolio choice model, the presence of credit market imperfections means that policy affects the covariance structure of asset returns. As a result, the w_i^*'s will shift differentially in response to monetary tightening as the perceived riskiness of debt issued by firms with currently high debt-equity ratios will increase relative to that of others.[8]

THE LENDING VIEW: LOANS FROM INTERMEDIARIES

The second mechanism articulated by proponents of the lending channel can be described by dividing the "other" assets in investors' portfolios into at least three categories: outside money, "loans" and all the others. Next, assume that there are firms for which loans are the only source of external funds—some firms cannot issue securities.[9] Depending on the solution to the portfolio allocation problem, a policy action may directly change both the interest rate and

the *quantity* of loans. It is not necessary to have a specific institutional framework in mind to understand this. Instead, it occurs whenever loans and outside money are complements in investor portfolios; that is, whenever the portfolio weight on loans is a negative function of the return on outside money for given means and covariances of other asset returns.[10]

The argument has two clear parts. First, there are borrowers who cannot finance new projects except through loans, and second, policy changes have a direct effect on loan supply. Consequently, the most important impact of a policy innovation is cross-sectional, as it affects the quantity of loans to loan-dependent borrowers.

Most of the literature on the lending view focuses on the implications of this mechanism in a world in which banks are the only source of loans and whose liabilities are largely reservable deposits. In this case, a reduction in the quantity of reserves forces a reduction in the level of deposits, which must be matched by a fall in loans. The resulting change in the interest rate on outside money will depend on access to close bank deposit substitutes. But the contraction in bank balance sheets reduces the level of loans. Lower levels of bank loans will only have an impact on the real economy insofar as there are firms without an alternative source of investment funds.

As a theoretical matter, it is not necessary to focus narrowly on contemporary banks in trying to understand the different possible ways in which policy actions have real effects. As I have emphasized, bank responses to changes in the quantity of reserves are just one mechanism that can lead to a complementarity between outside money and loans. As pointed out by Romer and Romer (1990), to the extent that there exist ready substitutes in bank portfolios for reservable deposits such as CDs, this specific channel could be weak to nonexistent.[11] But it remains a real possibility that the optimal response of investors to a policy contraction would be to reduce the quantity of loans in their portfolios.

The portfolio choice model also helps to make clear that the manner in which policy actions translate into loan changes need not

be a result of loan rationing, although it may.[12] As Stiglitz and Weiss (1981) originally pointed out, a form of rationing may arise in equilibrium as a consequence of adverse selection. But the presence of a lending channel does not require that there be borrowers willing to take on debt at the current price who are not given loans. It arises when there are firms which do not have equivalent alternative sources of investment funds and loans are imperfect substitutes in investors' portfolios.

Obviously, the central bank can take explicit actions directed at controlling the quantity of loans. Again, lowering the level of loans will have a differential impact that depends on access to financing substitutes. But the mechanism by which explicit credit controls influence the real economy is a different question.[13]

DISTINGUISHING THE TWO VIEWS: GENERAL CONSIDERATIONS

Distinguishing between these two views is difficult because contractionary monetary policy actions have two consequences, regardless of the relative importance of the money and lending mechanisms. It both lowers current real wealth and changes the portfolio weights.[14]

Assuming that there are real effects, contractionary actions will reduce future output and lower current real wealth, reducing the demand for all assets. In the context of standard discussions of the transmission mechanism, this is the reduction in investment demand that arises from a cyclical downturn.[15]

The second effect of policy is to change the mean and covariance of expected asset returns. This changes the w_{it}'s. In the simplest case in which there are two assets, outside money and everything else, the increase in the return on outside money will reduce the demand for everything else. This is a reduction in real investment.

The lending view implies that the change in portfolio weights is more complex and in an important way. There may be some combination of balance sheet and loan supply effects.

This immediately suggests that looking at aggregates for evidence of the right degree of imperfect substitutability or timing of changes may be very difficult. What seems promising is to focus on the other distinction between the two views—the lending view's assumption that some firms are dependent on loans for financing.

In addition to differences stemming from the relative importance of shifts in loan demand and loan supply, the lending view also predicts cross-sectional differences arising from balance sheet considerations. These are also likely to be testable. In particular, it may be possible to observe whether, given the quality of potential investment projects, firms with higher net worth are more likely to obtain external funding. Again, the major implications are cross-sectional.

EMPIRICAL EVIDENCE: PRELIMINARIES

Before discussing any empirical examination of the monetary transmission mechanism, two questions must be addressed. First, do nominal shocks in fact have real effects? Unless monetary policy influences the real economy, it seems pointless to study the way in which policy changes work. Second, how can we measure monetary policy? In order to calculate the impact of monetary policy, we need a quantitative measure that can reliably be associated with policy changes.

Here I take up each of these issues. In the following section, I will weigh the evidence on the real effects of money. This is followed by a discussion of ways in which recent studies have attempted to identify monetary shocks.

THE REAL EFFECTS OF NOMINAL SHOCKS

Modern investigation of the impact of money on real economic activity began with Friedman and Schwartz (1963). In many ways, this is still the most powerful evidence in support of the claim that monetary policy plays an important role in aggregate fluctuations. Through an examination that spanned

[12] Since there must be firms that are loan-dependent, there is still some form of rationing in the security market.

[13] See Romer and Romer (1993) for a concise discussion of recent episodes in which the Federal Reserve has attempted to change the composition of bank balance sheets through means other than standard policy actions.

[14] The change in portfolio weights can arise either from any combination of a change in the return on the outside asset, a change in the covariance structure of returns, or a shift in the consumption process.

[15] In general equilibrium, there is an offsetting effect that arises from the increase in the interest rate. All other things equal, this would increase saving and therefore investment. But we can be fairly confident that so long as monetary policy tightening can cause a recession, the impact of the income and wealth declines will be large enough that investment will fall.

numerous monetary regimes, they argue that apparently exogenous monetary policy actions preceded output movements.

Recent researchers use more sophisticated statistical tools to study the correlations between money and income. This "money-income causality" literature is largely inconclusive, because it fails to establish convincingly either that money "caused" output or the reverse. In the end, the tests simply establish whether measures of money *forecast* output, not whether there is causation. Given that outside money—the monetary base—is less than 10 percent of the size of M2, it is not surprising that economists find the simultaneity problems inherent in the question too daunting and give up.

Two pieces of evidence seem reasonably persuasive in making the case that money matters. First, the Federal Reserve seems to be able to change the federal funds rate virtually without warning. (I am not arguing that this is necessarily a good idea, just that it is possible.) In the very short run, these nominal interest rate changes cannot be associated with changes in inflationary expectations, and so they must represent real interest rate movements. Such real interest rate changes almost surely have an impact on real resource allocations.[16]

The second piece of evidence comes from the examination of the neutrality of money in Cecchetti (1986, 1987). In those papers, I establish that output growth is significantly correlated with money growth at lags of up to 10 years! There are several possible interpretations of these findings, but they strongly suggest that monetary shocks have something to do with aggregate real fluctuations.

MEASURING INNOVATIONS TO MONETARY POLICY

It stands to reason that before one can study the monetary transmission mechanism, it is necessary to identify monetary shocks. A number of authors have argued convincingly that policy disturbances cannot be gauged by examining movements in the monetary aggregates. The reason is that the variance in the innovations to broad measures

of money are a combination of endogenous responses to real shocks (King and Plosser, 1984) and shifts in money demand (Bernanke and Blinder, 1992).

There have been two reactions to the fact that monetary aggregates provide little insight into policy actions. Both begin by looking at the functioning of the Federal Reserve and examining how policy is actually formulated. The first, due to Bernanke and Blinder (1992), note that the federal funds rate is the actual policy instrument that is used on a day-to-day basis. This suggests that innovations to the federal funds rate are likely to reflect, at least in part, policy disturbances. The main justification for their conclusion comes from examining the institutions of how monetary policy is carried out.

Romer and Romer (1989) suggest a second method. By reading the minutes of the Federal Open Market Committee (FOMC) meetings, they have constructed a series of dates on which they believe policy became contractionary.[17]

Innovations to the Federal Funds Rate

To understand the shortcomings of these two approaches, I will describe how each is used. In the first, researchers begin by specifying a vector autoregression. For the purposes of the example, I will use the formulation in Bernanke and Blinder's (1992) Section IV. They employ a six-variable specification with the total civilian unemployment rate, the log of the CPI, the federal funds rate, and the log of three bank balance sheet measures, all in real terms: deposits, securities and loans. The assumption is that the federal funds rate is a "policy" variable, and so it is unaffected by all other contemporaneous innovations.[18]

Following Bernanke and Blinder, I estimate the VAR with six lags using seasonally adjusted monthly data.[19] Figures 1 and 2 plot some interesting results from this VAR. The first figure shows the estimated residuals from the federal funds rate equation. The solid vertical lines are National Bureau of Economic Research (NBER) reference cycle peaks and troughs, while the dashed vertical lines are the Romer and Romer dates, intended to indicate the onset of contractionary

[16] The equivalent open economy observation is that in small open economies, exchange rates move in response to changes in policy.

[17] Boschen and Mills (1992) describe a related technique.

[18] See Hamilton (1994) for a complete description of the methodology.

[19] The exact measures and sample follow those of Kashyap and Stein (1994a), who kindly supplied the data.

monetary policy episodes.

This series looks extremely noisy and it is hard to see how it could represent policy changes. The 1979-82 period is the only one with large positive or negative values. Although it is surely the case that there are unanticipated policy changes both when the Federal Reserve acts and when it does not, one would expect small normal shocks with occasional spikes. If decisions are really this random, there is something fundamentally wrong with the policymaking apparatus. Furthermore, since the federal funds rate itself is the equilibrium price in the reserves market, given technicalities of the way that monetary policy is actually carried out, the market-determined level of the funds rate is not a policy instrument.[20]

The second figure shows the response of the log of the CPI to a positive one percentage point innovation in the federal funds rate. To understand how this is computed, begin by writing the vector autoregression as

$$A(L)y_t = \varepsilon_t,$$

where $A(L)$ is a matrix of polynomials in the lag operator L ($L^i y_t \equiv y_{t-i}$), y_t is the vector of variables used in the estimation, and ε_t is mean zero independent (but potentially heteroskedastic) error. The first step is to estimate the reduced form version of equation 1 by assuming that no contemporaneous variables appear on the right-hand side of any equations ($A(0) = I$). This results in an estimate $\hat{A}(L)$ along with an estimated covariance matrix for the coefficient estimates—call this $\hat{\Omega}$. The impulse response functions are obtained by inverting the estimated lag polynomial $\hat{B}(L) = \hat{A}^{-1}(L)$.[21]

But the point estimate of the impulse response function is not really enough to allow us to reach solid conclusions. It is also important to construct confidence intervals for the estimates. There are two ways to do this. The first involves the technique that has been called Monte Carlo Integration. This is a Bayesian procedure that involves presuming that the distribution of the vector of errors in equation 1—the ε_t's—is i.i.d. normal.[22] To avoid making such stringent assumptions, I choose to estimate confidence bands using an alternative technique grounded

Figure 1

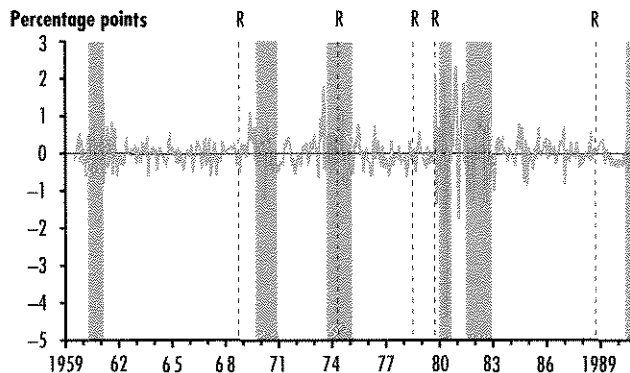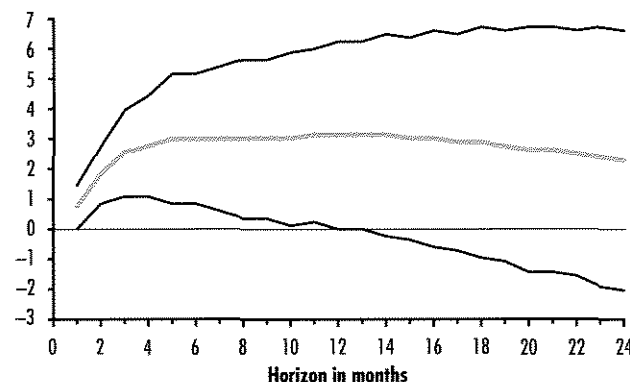## Estimated Innovations to the Federal Funds Rates

Figure 2

Response of the Log CPI to a Change in the Federal Funds Rate

in classical statistics.

The delta method is the simple procedure that comes from noting that if the estimates of the coefficients in lag polynomials are asymptotically normally distributed, then any well-behaved function of these parameters will also be asymptotically normally distributed. Stacking all of the parameters in $A(L)$ and calling the result θ, then

$$\sqrt{T}(\hat{\theta} - \theta) \xrightarrow{A} N(0, \Omega).$$

It follows that any function of these parameters $f(\theta)$—for example, the impulse response function—will be asymptotically normally distributed,

[20] This entire discussion ignores the possibility that *anticipated* monetary policy matters— something that researchers should consider bringing into the discussion.

[21] A simple way to calculate the vector moving-average form of equation 1 is to construct the companion form of the VAR as described in Sargent (1987). This is also discussed in Hamilton (1994).

[22] See Doan (1990).

$$\sqrt{T}\left(f[\hat{\theta}]-f[\theta]\right) \stackrel{A}{\sim} N\left(O,\Omega_f\right),$$

where

$$\Omega_f = \frac{\partial f[\theta]}{\partial \theta'}\Omega\frac{\partial f[\theta]}{\partial \theta},$$

which can be estimated numerically.

The result plotted in Figure 2 was first pointed out by Sims and is known as the "price puzzle." Paradoxically, the VAR estimates imply that monetary policy contractions lead to price *increases*! As is clear from the estimated standard-error bands, this price rise is significantly positive for approximately the first year. After two years, however, it is not possible to reject the hypothesis that a funds rate increase has no effect on the price level.[23]

The standard conclusion is that the VAR is misspecified in some way. One strong possibility is that the funds rate is not exogenous in the way that is required for this identification to be valid, and so these innovations do not accurately reflect policy movements.[24]

My conclusions may be too harsh for the following reason. As Ben Bernanke pointed out in the conference, the estimated innovations are the sum of true policy innovation, policy responses to omitted variables, and more general specification errors in the VAR. As a result, one would expect them to be noisy. Furthermore, as pointed out by Adrian Pagan, since one is primarily interested in the impulse response functions—the impact of unanticipated policy on output, prices and the like—then it may be immaterial that the estimated policy innovations are noisy even if the true innovations are not.

The Romer and Romer Dates

The Romer and Romer dates have been both widely used and extensively criticized.[25] They suffer from both technical and substantive problems. First, they are discrete. Presumably, policy changes have both an intensity and a timing. Ignoring the size of policy changes must have an impact on results. Second, Romer and Romer choose to focus their inquiry only on policy contrac-

tions, because they feel that expansions were more ambiguous. Since most models predict symmetric responses to positive and negative monetary innovations, this strategy throws out information.

But the main issue is the exogeneity of the policy shifts. It is difficult to believe that the actions of the FOMC, as reported in the minutes of the meetings, are truly exogenous events. There have been two responses to this. First, Hoover and Perez (1991) provide a lengthy discussion of why Romer and Romer's methods are not compelling in identifying output fluctuations induced by exogenous monetary shocks.

Taking a slightly different approach, Shapiro (1994) examines whether the FOMC is responding to changes in economic conditions, and so there is some reaction function implicit in policy. He estimates a probit model for the Romer and Romer dates using measures of inflation and unemployment, both as deviations from a carefully constructed target level, as determinants. Figure 3 reproduces his estimates of the probability of a date, with the vertical lines representing the dates themselves. The unanticipated policy action is 1 minus the estimated probability. As is clear from the figure, several of the dates were largely anticipated, and there were some periods when policy shifts were thought to be likely, and then did not occur. Overall, Shapiro's results suggest that the standard interpretation of the dummy variables as exogenous is incorrect to varying degrees over time.

There seems to be no way to measure monetary policy actions that does not raise serious objections. Given this, it might seem difficult to see how to proceed with the study of different theories of the transmission mechanism. But the literature proceeds in two directions. The first uses these measures directly in an attempt to gauge the influence of policy changes directly. The conclusions of these studies must be viewed with some degree of skepticism. The alternative approach is to note that investment declines account for the major share of output reductions during recessions. If one is able to show that the distribution of the contraction in investment is correlated with variables

[23] A test for whether all of the reactions are jointly zero rejects at the 5 percent level for the first seven months, and then fails to reject. The test for whether all the effects are zero simultaneously for all 24 months fails to reject with a p-value of 0.77.

[24] There is a further reason to view this measure of policy with some skepticism. Because of the large number of parameters estimated, the regressions are usually overfitted. As a result, they normally have very poor out-of-sample forecasting properties.

[25] See Friedman (1990) for a discussion of the strategy of using policymakers' statements to gauge their actions.

related to a firm's balance sheet and its access to bank loans, then this strongly suggests the existence of a lending channel.

USING AGGREGATE DATA

Numerous studies have used aggregate data in an attempt to distinguish the channels of monetary transmission. This literature can be divided into three categories: The first looks at the relative forecasting ability of different quantity aggregates; the second studies differences in the timing of the response of aggregate quantities to presumed policy shocks; and the third examines the behavior of interest rates.

Before examining the work on quantities, I will discuss the use of interest rate data.[26] As is clear from the discussion in the first section, the lending view does allow for movements in market interest rates. Furthermore, these movements are in the same direction as those predicted by the money view, and their magnitude depends solely on the degree of substitutability between outside money and various other assets. Where the two views differ is in their predictions for movements in the interest rate on loans. But since there is currently no secondary market for these securities, it is impossible to determine the interest rate on these loans.[27] This implies that market interest rates are of virtually no use in this exercise. There is no sense in which the behavior of interest rates could serve to distinguish between the money or lending views.

I now turn to the work on quantities. In the following section, I examine tests involving the relative forecasting ability of measures of money and credit. This is followed by a discussion of papers that emphasize aggregate timing relationships.

Relative Forecasting Ability

A number of papers have examined the ability of different financial aggregates to forecast output (or unemployment) fluctuations. Ramey (1993) is a recent example. The main methodology here is to ask whether measures of credit are informative about future output movements, once money has

Figure 3

Estimated Probability of a Romer and Romer Date
(Shapiro probit model using inflation and unemploymnt)
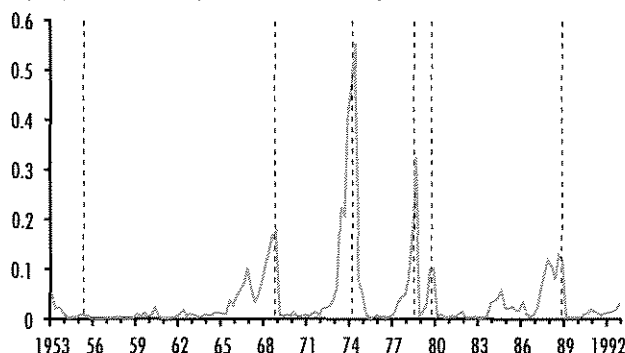
been taken into account. The problem with this is that credit is usually just a broader measure of money. To put it slightly differently, the balance sheet identity of the banking system implies that bank assets equal bank liabilities. As Bernanke (1993b) points out, monetary aggregates are a measure of bank liabilities, while credit aggregates are measures of bank assets. Since these are calculated slightly differently, they will not be identical. But it is these technical measurement differences that are likely to account for the differences in forecasting ability, not anything about the transmission mechanism.

More generally, the main finding is that credit lags output. Unfortunately, this tells us nothing about the transmission mechanism. The aggregate data do show that aggregate credit is *countercyclical*, but it is easy to find explanations for this that are consistent with the lending view. For example, Kiyotaki and Moore (1993) present a model in which individuals must continue to service credit even after income falls, and so credit falls after income even though it is the fundamental source of fluctuations. In the end, it is difficult to see how aggregate timing relationships can tell us anything at all about the way in which monetary policy affects real activity.[28]

Aggregate Timing Relationships

The second use of aggregate data has been to examine the response of various

[26] Miron, Romer and Weil (1994) study pre-World War II interest rates in an attempt to address these questions.

[27] In the presence of rationing, there is the added complication that one would need observations on the shadow price for a loan to a borrower who is deemed not to be creditworthy given the economic environment. Obviously, there is no easy way to infer such a price.

[28] This point is also made by Bernanke, Gertler and Gilchrist (1994).

Figure 4

Bernanke and Blinder Plot of Responses to Federal Funds Rate Shock

Horizon in months
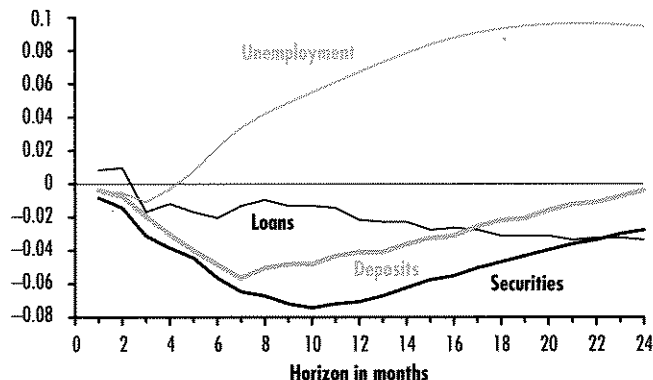
financial quantities to policy innovations. Returning to Bernanke and Blinder (1992), they study whether bank loans and securities respond differently to federal funds rate innovations.[29] The standard methodology is to calculate the impulse responses for the two variables and note that they look different. Figure 4 reports the common finding, calculated using the six-variable Bernanke and Blinder VAR estimated over the 1959-90 sample. In response to a positive 1 percentage point innovation in the federal funds rate, the unemployment rate rises by nearly 0.1 percentage point after one-and-a-half years, while bank securities fall 0.07 percent and loans decline 0.02 percent. Securities fall both by a larger amount and more quickly than loans.

But point estimates of these impulse responses do not tell the entire story. In Figure 5, I plot the point estimate and two standard error bands for the difference between the impulse response for loans and securities. This allows an explicit test of whether these two assets are imperfectly substitutable in response to the shock. The differences are individually greater than zero in only a few months, and a joint test of the first 24 months of the impulse response, which is asymptotically distributed as a Chi-squared, has a p-value of 0.70.

My conclusion is that reduced-form vector autoregressions are nearly incapable of providing convincing evidence of a differential impact of federal funds rate innovations on various parts of bank balance sheets. These results are based on the estimation of a large number of parameters with a relatively small amount of data—this VAR has 237 parameters and 354 data points—and so the estimates are fairly imprecise.[30]

But even if one were to find that the impulse responses differed significantly, this would only bear on the substitutability of the assets, and not directly on the validity of the lending view. Both the prices and quantities of perfect substitutes must have the same stochastic process, and so finding that this particular partial correlation is different would be evidence of imperfect substitutability. As Bernanke and Blinder (1992) make clear in discussing their findings, this is a necessary but not a sufficient condition for the lending view to hold. It is not possible, using reduced-form estimates based on aggregate data alone, to identify whether bank balance sheet contractions are caused by shifts in loan supply or loan demand. What is needed is a variable that is known to shift one curve but not the other.

Kashyap, Stein and Wilcox (1993) also provide evidence based on aggregate timing. They compare the response of bank loans to that of commercial paper issuance following policy innovations. They find that monetary policy contractions seem to decrease the mix of loans relative to commercial paper. Borrowers that can move away from direct bank finance following a tightening appear to do so. Both Friedman and Kuttner (1993), and Oliner and Rudebusch (1993) take issue with these findings and show that changes in the mix are due to increases in the amount of commercial paper issuance during a recession, but that the quantity of bank loans does not change. In addition, Oliner and Rudebusch show that once firm size is taken into account, and trade credit is included in the debt of the small firms, the mix of financing is left unaffected by policy changes.

It is worth making an additional point about the commercial paper market. First, Post (1992) documents that all commercial paper rated by a rating agency must have a backup source of liquidity, which is generally a bank line of credit or a standby letter of

credit. This means that commercial paper is an indirect liability of banks, albeit one that is not on their balance sheet. Furthermore, Calomiris, Himmelberg and Wachtel (1994) suggest that increases in commercial paper issuance are accompanied by an increase in trade credit. This means that a policy contraction may simply cause a re-shuffling of credit by forcing banks to move liabilities off of their balance sheet such that large firms issue commercial paper in order to provide trade credit to small firms that would have otherwise come from banks.

USING CROSS-SECTIONAL DATA

There is a large empirical literature using cross-sectional data that is relevant to understanding the channels of monetary policy. These studies fall into groups that separately address the two parts of the lending view. The first set of papers tries to gauge the importance of capital market imperfections on investment, and so is related to the balance sheet effects described in the first section. The second set, which is fairly small, examines time-series variation in cross-sectional data in an attempt to characterize the distributional effects of monetary policy directly. I will briefly describe each of these strategies.

Measuring Capital Market Imperfections

The literature on capital market imperfections is an outgrowth of the vast work done on the determinants of investment. The general finding in this literature is that internal finance is less costly than external finance for firms that have poor access to primary capital markets.

The empirical studies fall into two categories. The first examines reduced-form correlations, while the second looks directly at the relationship between the cost and expected return to a marginal investment project—they estimate structural Euler equations.
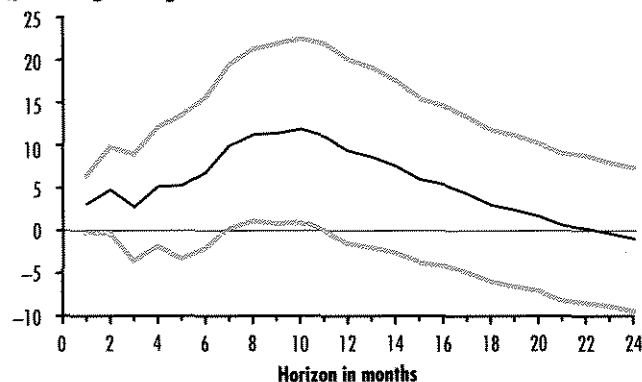
Reduced-Form Correlations

Fazzari, Hubbard and Petersen (1988)

Figure 5

Difference Between Loan and Security Response to Federal Funds Rate Shock
(percentage change at annual rate with two standard-deviation bands)

Horizon in months

pioneered the technique of dividing firm-level data into groups using measures thought to correspond to the project monitoring costs created by information asymmetries, and then seeing if the correlation between investment and cash-flow measures varies across the groups. The finding in a wide range of studies is that investment is more sensitive to cash-flow variables for firms who have ready access to outside sources of funds.[31]

The main issue in interpreting these results is whether the characteristics of the firm used to split the sample are exogenous to financing decisions. Measures of firm size, dividend policies, bond ratings and the like may be related to the quality of investment projects a firm has available, and so lender discrimination may not be a consequence of asymmetric information.

There are several examples in which researchers identify potentially constrained firms based on institutional characteristics, and so the endogeneity problems are mitigated. I will mention two. Hoshi, Kashyap and Scharfstein (1991) find that investment by Japanese firms that were members of a *keiretsu*, or industrial group, was not influenced by liquidity effects. Using data on individual hospitals, Calem and Rizzo (1994) find that investment depends more heavily on cash-flow variables for small, single-unit hospitals than for large, network-affiliated ones.

In the most convincing study of this type, Calomiris and Hubbard (1993) study

[31] Bernanke, Gertler and Gilchrist (1994) survey the large number of studies that use this approach.

the undistributed corporate profits tax in 1936 and 1937 to estimate the differences in financing costs directly from firms' responses to the institution of a graduated surtax intended to force an increase in the dividend payout rate. Their results, holding investment opportunities fixed, are that investment spending is affected by the level of internal funds only for those firms with low levels of dividend payments and high marginal tax rates. Furthermore, these tended to be smaller and faster growing firms.

Structural Estimation

The neoclassical theory of investment allows one to derive the complex equilibrium relationship among the capital stock, rates of return, future marginal value products and project costs that form the first-order conditions for a firm's problem. With the appropriate data, it is then possible to see whether these Euler equations hold. Hubbard and Kashyap (1992) is an interesting use of this technique. Following the work of Zeldes (1989) on consumption, they examine whether the ability of agricultural firms to meet this first-order condition depends on the extent of their collaterizable net worth. They find that during periods when farmers have high net worth, and so have better access to external financing, their investment behavior is more likely to look as if it is unconstrained.

These investment studies have been very successful in establishing the existence of capital market imperfections as well as their likely source in information asymmetries arising from monitoring problems. While the work has little to say about monetary policy directly, it does provide an excellent characterization of the distributional effects of changes in the health of firms' balance sheets regardless of the source.

Time-Series Evidence

The strategy in the second set of studies is to use the cross-sectional dimension to identify the transmission mechanism. The goal is to determine whether the reduction in loans during monetary contractions is a con-

sequence of shifts in loan demand or loan supply. My conclusion is that these studies fail to establish the desired result in a convincing way. Instead, they provide further evidence of capital market imperfections.

Three major studies use data on manufacturing firms. In the first, Gertler and Hubbard (1988) find that the impact of cash flow on investment is higher during recessions for firms that retain a high percentage of their earnings. The second, by Kashyap, Lamont and Stein (1992), shows that during the 1981-82 recession, the inventories of firms without ready access to external finance fell by more when their initial level of internal cash was lower. On the other hand, the inventory investment behavior of firms with ready access to primary capital markets showed no evidence of liquidity constraints. In the third, Gertler and Gilchrist (1994) use the *Quarterly Financial Report for Manufacturing Corporations (QFR)* to divide firms into asset-size categories and find that small firms account for a disproportionate share of the decline in manufacturing following a monetary shock.[32]

Both Kashyap and Stein (1994b) and Peek and Rosengren (forthcoming) focus on the behavior of lenders rather than borrowers. By examining the cyclical behavior of banks, Kashyap and Stein hope to find evidence for the importance of loan supply shifts. The strongest result in their paper is that, following a monetary contraction, the total quantity of loans held by small banks falls while that of large banks does not. By contrast, Peek and Rosengren study New England banks during the 1990-91 recession and find that poorly capitalized banks shrink by more than equivalent institutions with higher net worth. My interpretation is that both of these show that the capital market imperfections commonly found to apply to manufacturing firms apply to banks as well.

There are two difficulties inherent in any attempt to establish that the important transmission mechanism for monetary policy shocks is through bank loan supply shifts. First, as described at length in the second section, there is the problem of empirically identifying monetary policy. Beyond this, there is the subtlety of distinguishing loan

[32] Oliner and Rudebusch (1994) and Bernanke, Gertler and Gilchrist (1994) obtain similar results using the *QFR* data as well.

supply shifts from the balance sheet effects. Is the observed reduction in loans a consequence of their complementarity with outside money caused by the structure of the banking system, or is it the result of changes in the shape of the marginal efficiency of investment schedule brought on by the balance sheet effects? Kashyap, Lamont and Stein (1992) suggest one possible way of distinguishing these possibilities. If one can find a recessionary period that was *not* preceded by a monetary contraction, and show that interest rates rose but that bank dependence was irrelevant to individual firms' experiences, this would mean that banks are responsible for the distributional effects induced by monetary shocks. Unfortunately, such evidence is not readily available.

CONCLUSION

After a survey of the work that attempts to distinguish theories of the monetary transmission mechanism, where do we stand? My conclusion is that the myriad studies have succeeded in establishing the empirical importance of credit market imperfections. This means that monetary policy shifts have an important distributional aspect that cannot be addressed within the traditional money view. It is the smaller and faster growing firms that bear a disproportionate share of the burden imposed by a recession. Since these are likely to be firms with highly profitable investment opportunities, this has important implications for social welfare. Not only are recessions associated with aggregate output and investment declines, but the declines are inefficient.

Beyond this, there is the issue of distinguishing the two parts of the lending view. Do we care if we can distinguish changes in investment opportunities resulting from financial accelerator effects from bank loan supply shifts per se? Does the conclusion have implications for the actual conduct of monetary policy? I believe the answer is yes. If the complementarity of bank loans and outside money arises largely as a result of the financial regulatory environment, then, with financial innovation and liberalization, these effects are likely to become less important

over time. With the introduction of interstate banking and the development of more sophisticated instruments aimed at trading pools of loans, it is only the balance sheet effects that will remain. As a result, it is important to know which is the more important channel of monetary policy transmission.

REFERENCES

Bernanke, Ben S. "How Important is the Credit Channel in the Transmission of Monetary Policy? A Comment," *Carnegie-Rochester Conference Series on Public Policy 39* (December 1993) pp. 47-52.

_____. "Credit in the Macroeconomy," *Federal Reserve Bank of New York Quarterly Review* (spring 1993) pp. 50-70.

_____ and Mark Gertler. "Financial Fragility and Economic Performance," *Quarterly Journal of Economics* (February 1990), pp. 87-114.

_____ and Mark Gertler. "Agency Cost, Net Worth and Business Fluctuations," *The American Economic Review* (March 1989), pp. 14-31.

_____ and Alan S. Blinder. "The Federal Funds Rate and the Channels of Monetary Transmission," *The American Economic Review* (September 1992), pp. 901-21.

_____ Mark Gertler and Simon Gilchrist. "The Financial Accelerator and the Flight to Quality," *National Bureau of Economic Research Working Paper No. 4789* (July 1994).

Brainard, William C., and James Tobin. "Financial Intermediaries and the Effectiveness of Monetary Controls," *The American Economic Review* (May 1963), pp. 383-400.

Boschen, John F., and Leonard O. Mills. "The Effects of Countercyclical Monetary Policy on Money and Interest Rates: An Evaluation of Evidence from FOMC Documents," working paper (1992), College of William and Mary.

Calem, Paul, and John A. Rizzo. "Financing Constraints and Investment: New Evidence from Hospital Industry Data," working paper (1994), Yale University, School of Medicine, Department of Epidemiology and Public Health.

Calomiris, Charles W., Charles P. Himmelberg and Paul Wachtel. "Commercial Paper and Corporate Finance: A Microeconomic Perspective," working paper (April 1994), Stern School of Business, New York University.

_____ and R. Glenn Hubbard. "Tax Policy, Internal Finance, and Investment," Internal Finance and Investment: Evidence from the Undistributed Corporate Profits Tax of 1936-37, *National Bureau of Economic Research Working Paper No. 4288* (March 1993).

Cecchetti, Stephen G. "Testing Short Run Neutrality: International Evidence," *Review of Economics and Statistics* (February 1987) pp. 135-40.

_____. "Testing Short Run Neutrality," *Journal of Monetary Economics* (May 1986) pp. 409-23.

Christiano, Lawrence, and Martin S. Eichenbaum. "Liquidity Effects and the Monetary Transmission Mechanism," *The American Economic Review* (May 1992) pp. 346-53.

Doan, Thomas A. User's Manual, RATS Version 3.10. VAR Econometrics, 1990.

Fama, Eugene F. "Banking in the Theory of Finance," *Journal of Monetary Economics* (January 1980) pp. 39-57.

Fazzari, Steven M., R. Glenn Hubbard and Bruce C. Petersen. "Financing Constraints and Corporate Investment," *Brookings Papers on Economic Activity* (1988:1) pp. 141-95.

Feurst, Timothy. "Liquidity Effect Models and Their Implications for Monetary Policy," working paper (April 1993), Bowling Green State University, Department of Economics.

Friedman, Benjamin M. "Discussion of 'New Evidence on the Monetary Transmission Mechanism," *Brookings Papers on Economic Activity* (1990:1), pp. 204-9.

_____and Kenneth N. Kuttner. "Economic Activity and the Short-Term Credit Markets: An Analysis of Prices and Quantities," *Brookings Papers on Economic Activity* (1993:2), pp. 193-266.

Friedman, Milton, and Anna Schwartz. *A Monetary History of the United States*, 1867-1960. Princeton University Press, 1963.

Gertler, Mark, and Simon Gilchrist. "Monetary Policy, Business Cycles, and the Behavior of Small Manufacturing Firms," *Quarterly Journal of Economics* (May 1994) pp. 309-40.

_____ and _____. "The Role of Credit Market Imperfections in the Monetary Transmission Mechanism: Arguments and Evidence," *Scandinavian Journal of Economics* (1993) pp. 34-64.

_____ and R. Glenn Hubbard. "Financial Factors in Business Fluctuations," National Bureau of Economic Research *Working Paper No. 2758* (November 1988).

Hamilton, James D. *Time Series Analysis*. Princeton University Press, 1994.

Hoover, Kevin D., and Stephen J. Perez. "Post Hoc Ergo Propter Hoc Once More: An Evaluation of 'Does Monetary Policy Matter?' in the Spirit of James Tobin," *Journal of Monetary Economics* (February 1992), pp. 3-24.

Hoshi, Takeo, Anil K. Kashyap and David Scharfstein. "Corporate Structure, Liquidity and Investment: Evidence from Japanese Panel Data," *Quarterly Journal of Economics* (February 1991), pp. 33-60.

Hubbard, R. Glenn. "Is There a 'Credit Channel' for Monetary Policy?" this Review (May/June), pp. 63-82.

_____ and Anil K. Kashyap. "Internal Net Worth and the Investment Process: An Application to U.S. Agriculture," *Journal of Political Economy* (June 1992) pp. 506-34.

Ingersoll, Jonathan E. *Theory of Financial Decision Making*. Rowman and Littlefield, 1987.

James, Christopher. "Some Evidence on the Uniqueness of Bank Loans," *Journal of Financial Economics* (1987), pp. 217-36.

Kashyap, Anil K., and Jeremy C. Stein. "Monetary Policy and Bank Lending," in N. Gregory Mankiw, ed., *Monetary Policy*. University of Chicago Press for the National Bureau of Economic Research, 1994, pp. 221-62.

_____ and _____. "The Impact of Monetary Policy on Bank Balance Sheets," working paper (March 1994), University of Chicago, Graduate School of Business.

_____, Jeremy C. Stein and David W. Wilcox. "Monetary Policy and Credit Conditions: Evidence from the Composition of External Finance," *The American Economic Review* (March 1993), pp. 78-98.

_____, Owen A. Lamont and Jeremy C. Stein. "Credit Conditions and the Cyclical Behavior of Inventories: A Case Study of the 1981-1982 Recession," National Bureau of Economic Research *Working Paper No. 4211* (November 1992).

King, Robert G., and Charles I. Plosser. "Money, Credit and Prices in A Real Business Cycle," *The American Economic Review* (June 1984), pp. 363-80.

King, Stephen R. "Monetary Transmission: Through Bank Loans or Bank Liabilities?" *Journal of Money, Credit and Banking* (August 1986), pp. 290-303.

Kiyotaki, Nobuhiro, and John Moore. "Credit Cycles," working paper (March 1993), London School of Economics .

Lucas, Robert E., Jr. "Liquidity and Interest Rates," *Journal of Economic Theory* (1990), pp. 237-64.

Miron, Jeffrey, Christina D. Romer and David Weil. "Historical Perspectives on the Monetary Transmission Mechanism," in N. Gregory Mankiw, ed., *Monetary Policy*. University of Chicago Press for the National Bureau of Economic Research, 1994, pp. 263-306.

Morgan, Donald P. "The Lending View of Monetary Policy and Bank Loan Commitments," working paper (December 1992), Federal Reserve Bank of Kansas City.

Oliner, Stephen, and Glenn D. Rudebusch. "Is There a Bank Credit Channel for Monetary Policy?" Board of Governors of the Federal Reserve System, Finance and Economics Division *Discussion Paper No. 93-8* (February 1993).

Peek, Joe, and Eric Rosengren. "The Capital Crunch: Neither a Borrower Nor a Lender Be," *Journal of Money, Credit and Banking* (forthcoming).

Post, Mitchell A. "The Evolution of the U.S. Commercial Paper Market Since 1980," *Federal Reserve Bulletin* (December 1992), pp. 879-91.

Ramey, Valerie. "How Important is the Credit Channel in the Transmission of Monetary Policy?" *Carnegie-Rochester Conference Series on Public Policy* (December 1993), pp. 1-45.

Romer, Christina D., and David H. Romer. "Credit Channels or Credit Actions? An Interpretation of the Postwar Transmission Mechanism," in *Changing Capital Markets: Implications for Monetary Policy*. Federal Reserve Bank of Kansas City, 1993.

_____ and _____. "New Evidence on the Monetary Transmission Mechanism," *Brookings Papers on Economic Activity* (1990:1), pp. 149-213.

_____ and _____. "Does Monetary Policy Matter? A New Test in the Spirit of Friedman and Schwartz," National Bureau of Economic Research *Macroeconomics Annual*. MIT Press, 1989, pp. 121-70.

Sargent, Thomas J. *Macroeconomic Theory,* Second Edition. Academic Press, 1987.

Sellon, Gordon H., Jr. "Measuring Monetary Policy," working paper (August 1994), Federal Reserve Bank of Kansas City.

Shapiro, Matthew D. "Federal Reserve Policy: Cause and Effect," in N. Gregory Mankiw, ed., *Monetary Policy.* University of Chicago Press for the National Bureau of Economic Research, 1994, pp. 307-34.

Sims, Christopher. "Interpreting the Macroeconomic Time Series Facts: The Effects of Monetary Policy," *European Economic Review* (May 1992), pp. 975-1000.

Stiglitz, Joseph E., and Andrew Weiss. "Credit Rationing in Markets with Imperfect Information," *The American Economic Review* (June 1981), pp. 393-410.

Strongin, Steven. "The Identification of Monetary Policy Disturbances: Explaining the Liquidity Puzzle," Federal Reserve Bank of Chicago *Working Paper No. 92-27* (November 1992).

Zeldes, Stephen P. "Consumption and Liquidity Constraints: An Empirical Investigation," *Journal of Political Economy* (April 1989), pp. 305-46.