Empirical Macroeconomics: The Effects of Monetary Policy

Aviel Nagel

Jeffrey Parker

Department of Economics
Reed College
2003

Research for this chapter was undertaken during the summer of 2002 under a Goldhammer Collaborative Research Grant from Reed College. It is a chapter from a book in progress by Parker surveying empirical methods and results in macroeconomics.

1 Conventional wisdom about money's effects

Two fundamental propositions about the effect of the quantity of money on the economy predate the emergence of monetary economics as a recognized discipline of study. The first is that increases in the quantity of money that are not accompanied by corresponding increases in real output eventually lead to inflation. The second is that a shortage of money can depress the amount of trade and productive activity. Both of theses propositions were recognized by the classical quantity theorists in the early 20th century and by the monetarists of the 1950s and 1960s.

The classic empirical documentation of these ideas was Milton Friedman and Anna Schwartz's *Monetary History of the United States, 1867–1960*. Friedman and Schwartz's book is a monumental contribution to monetary economics in several dimensions. They used records of banks and government agencies to construct measures of the money supply and reserves going back into the 19th century. Based on these measures and on a richly detailed analysis of contemporary writings, they meticulously document numerous episodes where disturbances emanating from the monetary sector were followed by recessions in real macroeconomic activity.

The best-known section of *Monetary History* is Chapter 7 on "The Great Contraction, 1929–33." This chapter argues that a series of identifiable financial-sector shocks acted to contract the money supply and lead the U.S. economy into its greatest depression. They identify the stock market crash in October 1929, two banking crises in October 1930 and March 1931, Britain's departure from the gold standard in September 1931, and final (and even more severe) banking crisis from the last months of 1932 through February 1933 as a punishing sequence of blows that combined to lower the money supply by over a third and commercial-bank deposits by nearly half.

Friedman and Schwartz's research methodology is mathematically simple but logically powerful. For each episode, they examine the movements in monetary and macroeconomic variables in detail. In each case, they consider the effects that the monetary shock may have had alongside other possible explanations of the movements of the aggregate variables. In order to assess various possible channels of causality, they consider not only the timing of various events, but also the narrative descriptions in contemporary accounts by policymakers, bankers, business people, and journalists. This "narrative approach" allows one to use non-quantitative information in the process of identification and analysis. In their words, "[w]e can go beyond the numbers alone and, at least on some occasions, discern the antecedent circumstances whence arose the particular movements that become so anonymous when we feed the statistics into the computer." (Friedman and Schwartz (1963)) Although more recently the econometric mainstream has tended to rely exclusively on "statistics fed into computers," we discuss below a recent and controversial revival of the narrative method of identifying monetary shocks.

To give but one example of the narrative approach, we shall examine their analysis of the Panic of 1907.[1] They begin by characterizing the behavior of indicators of real activity during the period from May 1907 to June 1908. Because national income data were not collected until well into the twentieth century, they measure real activity by a set of variables associated with the volume of goods production and trade: production of various commodities for which data are available, freight-car loadings, and failures of com-

[1] See Friedman and Schwartz (1963).

mercial firms. They point out that most indicators declined little until a bank panic occurred in October of 1907. At that point, all indicators of real activity show that the economy contracted sharply.

The money supply began declining in May 1907, reversing its usual upward trend and sliding 2.5 percent by September. Then, with the onset of the banking crisis, the monetary contraction accelerated, lowering the money supply a further 5 percent by February of 1908. Friedman and Schwartz break down the sources of the decline in the money supply before October as a combination of a gold outflow (responsible for about 1 percentage point of the decline) and an increase in the demand for reserves by the banking system, which both drained currency out of public circulation and reduced the volume of deposits.

In the middle of October 1907, losses at a series of banks that had speculated on risky mining stocks caused heavy withdrawals of funds by nervous depositors. These banks, in turn, owed money to some of the largest banks and trust companies in New York, which began to experience difficulties. By the end of October, a full-fledged run on the banking system ensued, with depositors scrambling to try to liquidate their bank deposits and banks desperately seeking sufficient liquidity to provide the needed funds. Large New York banks quickly suspended the convertibility of deposits into currency and specie, and other banks around the country soon followed suit. The suspension of convertibility stopped the runs on banks and allowed banks to gradually restore their depleted reserves.

Friedman and Schwartz go on to analyze in detail the month-to-month movements in bank reserves and deposits and in the public's demand for currency. By comparing the 1907-08 episode to the later (and much more severe) bank panics of the Great Depression, they conclude that the early suspension of convertibility by affected banks, which stopped the runs on banks and moderated the decline in the money supply, likely kept the contraction in real activity from becoming as severe as the decline in the 1930s.

The detailed analysis of movements in monetary measures over nearly a century of highly varied U.S. history and comparison of these movements with the contemporaneous changes in economic variables lead Friedman and Schwartz to some broad conclusions about the effects of money on the economy. In their final chapter, they summarize their findings as follows:

> 1. Changes in the behavior of the money stock have been closely associated with changes in economic activity, money income, and prices.
>
> 2. The interrelation between monetary and economic change has been highly stable.
>
> 3. Monetary changes have often had an independent origin; they have not been simply a reflection of changes in economic activity.
>
> 4. In monetary matters, appearances are deceiving; the important relationships are often precisely the reverse of those that strike the eye. (Friedman and Schwartz (1963))

By and large, the "close associations" and "stable interrelations" that they found supported the predictions of the modern quantity theory of money. For example, they note that the only period in their sample in which prices rose steadily and substantially was 1897 to 1914, which coincided with the highest rate of monetary growth of any (peacetime) period they studied. This supports the long-run connection between monetary growth and inflation.

The four periods with the highest degree of macroeconomic stability were also periods in which the money supply grew smoothly and the financial system operated without major disruption. In contrast, major disturbances in the banking system or policy-induced monetary contractions occurred at the onset of all major business contractions in their century-long sample period. Moreover, they found no examples of a monetary or financial disturbance of comparable magnitude that did not lead to a recession.

Friedman and Schwartz's attribution of causality from money to business cycles was criticized by James Tobin (1970), who claimed that they had committed the *post hoc ergo propter hoc* fallacy—assigning a causal relationship to two events on the basis of which happened first. In a rejoinder, Friedman (1970) argued that temporal precedence was only one of several criteria from which they inferred the direction of causality and that the case for a significant role for money in leading to business cycle fluctuations was clear independent of the precise timing of changes.

2 Monetary policy in macroeconometric models

As noted in Chapter 2, macroeconomists of the 1950s and 1960s devoted a great deal of effort to estimating large-scale empirical models of complete macroeconomic systems. One goal of these models was to evaluate the effects of alternative macroeconomic policies, including monetary policy, on the overall economy. Modern econometricians have pointed out several serious difficulties with this approach to policy analysis, but the results of these studies helped form a conventional wisdom about the effects of monetary policy that has probably continued to influence the evolution of more recent studies.

Most of these models were based on a Keynesian IS/LM framework with a Phillips curve added on to determine inflation.[2] Changes in monetary policy (usually specified as an exogenous shift in nonborrowed reserves) affected the money supply, which changed interest rates to balance money demand with supply. The change in interest rates then affected spending on investment and consumption, which caused changes in output and, eventually, in prices.

Table 1 shows the simulated effect of a $1 billion increase in nonborrowed reserves on several macroeconomic variables using four macroeconometric models.[3] For reference, the total stock of nonborrowed reserves was $20–30 billion in the 1960s, so an increase of $1billion would be an increase of 3–5 percent. While the results vary considerably across the models, several features are notable. First, the models agree more about the short-term effects of policy than about the long-term effects. All four models predict increases in real output in the first two years after a monetary expansion. However, the models' predictions at the five-year horizon vary from an even greater expansion to no effect to a sizable contraction in output. GNP in 1958 dollars ranged from about $500 to $800 billion in the 1960s, so a figure of 5–8 in the table would reflect a 1 percent increase, depending on the year used as a base.

[2] An exception was the "St. Louis model," developed at the Federal Reserve Bank of St. Louis based on a monetarist conception of the macroeconomy.

[3] These results are taken from Klein and Burmeister (1974), which summarizes a comparison study of the properties of various macro models. The models above were chosen for inclusion in the table based on the comparability of their reported monetary-policy multipliers. Some figures presented in the original were renormalized for comparison purposes.

Table 1. Simulated effects of $1 billion increase in nonborrowed reserves

	Wharton Model[a]	DRI Model[d]	Hickman-Coen Model	Brookings Model[e]
Effect on real output (billions of 1958 dollars)				
After 1 year	+3.9	+6	+1.7	+1.8
After 2 years	+6.2	+9	+1.7	+3.2
After 5 years	+6.7[b]	–3	+0.4	–1.2
Effect on short-term interest rate (percentage points)[c]				
After 1 year	–0.29	–0.75	NA	NA
After 2 years	–0.33	–0.5	NA	NA
After 5 years	–0.51[b]	–0.5	NA	NA
Effect on price level (percentage points)				
After 1 year	–0.1	0	+0.4	0.0
After 2 years	–0.1	+0.1	+0.6	+0.2
After 5 years	+0.1[b]	+0.4	+0.1	0.0

Sources: Wharton Model: Duggal, Klein, and McCarthy (1974), Tables 9, 11, and 12. DRI Model: Eckstein, Green, and Sinai (1974), Figures 4 and 10. Hickman-Coen Model: Hickman, Coen, and Hurd (1974), Table 3. Brookings Model: Schink (1974), Table 1.

Notes:
[a]Monetary policy multipliers reported for the Wharton model are the average on the values reported for high-employment and low-employment initial conditions.
[b]After 4 years.
[c]4-6 month commercial paper for Wharton; 90-day Treasury bills for DRI; no interest rate results reported for Hickman-Coen or Brookings models.
[d]Effects for DRI model are reported with less precision because they were read from a figure rather than computed from numbers in a table. Price changes were approximated as the difference between reported changes in nominal and real GNP, converted to percentage changes assuming an initial level of $700 billion.
[e]Monetary policy effects for the Brookings model were inferred as the difference between fiscal policy simulations with and without "monetary accommodation," which was defined as a $500 million increase in nonborrowed reserves.

The two models for which interest-rate effects are reported show a substantial and sustained decline in short-term interest rates as a result of the monetary expansion. The magnitude and duration of this estimated effect seem large in comparison with today's models.

Finally, there is no consensus at all in the models about the effects of monetary policy on prices. In the Wharton and Brookings models, there seems to be no effect whatsoever, while the Hickman-Coen model predicts a small short-run effect and the DRI model has a small effect that increases with time. The deviation of these predictions from the now-accepted notion of long-run monetary neutrality is striking.

An iconoclast among the 1980s econometric macro models was the St. Louis Federal Reserve model. Unlike the predominantly Keynesian models reported in Table 1, the St. Louis model had a simple monetarist structure, relating output, prices, and other variables to movements in the money supply.

3 Econometric interpretation of the money/output correlation: Does money cause output?

It is a commonly cited truism of econometrics that "correlation does not imply causality." Positive correlation between changes in the money supply and changes in real economic activity is an accepted econometric fact. However, this correlation could arise through monetary changes causing changes in output (the conventional Keynesian and monetarist interpretation), through changes in output causing changes in the quantity of money (the interpretation of real-business-cycle proponents), or through a third variable causing changes in both money and output.

Without making additional assumptions (such as the identifying assumptions discussed in Chapter 3 and below), econometrics cannot help us distinguish among these alternatives. The concept of Granger causality was developed in the years following 1969 to attempt to provide a reasonable framework for assessing causality in time series setting. The seminal work of Sims (1972) set off a wave of econometric research using Granger causality to examine the causal relationship between money and output.

3.1 Granger causality

Granger (1969) introduced the concept of time-series causality that bears his name. His definition 1 defines causality in an econometric sense as follows:

> Causality. If $\sigma^2(X \mid U) < \sigma^2(X \mid U - Y)$,, we say that Y is causing X, denoted by $Y \Rightarrow X$. We say that Y is causing X if we are better able to predict X using all available information than if the information apart from Y had been used. (Granger (1969))

In this definition, U refers to the set of all information available that might be useful in predicting X, $U - Y$ is this set with values of Y excluded, and $\sigma^2(X \mid U)$ is the variance of the forecast error achieved in predicting X based on U. Thus, Y is deemed to cause X if a lower forecast error can be achieved by including Y in the set of variables used to predict X than if it is omitted.

Two main methods have been used to implement Granger causality tests. Early tests, including Sims's seminal work, used a single equation with leads and lags. More recently, most causality tests have been conducted using vector autoregression models. [4]

3.1.1 Testing Granger causality in a VAR

In the vector autoregression implementation, the Y variables in Granger's definition are explicitly lagged values only. In the special (but common) case where only two variables are considered, the universal information set U consists of lagged values of both X and Y, so $U - Y$ is just lagged values of X. In this case, testing for Granger causality involves testing whether lagged values of both X and Y are more successful in predicting X than lagged values of X by themselves.

In terms of equations, the bivariate model can be written

[4] Sims (1972) and Hosoya (1977) demonstrate that these two techniques are equivalent tests under broad conditions. Granger originally proposed performing these test in the frequency domain using the cross-spectra of the two variables. Frequency-domain tests are rarely used now. The earliest applications to the money/income relationship, starting with Sims (1972), used the lead/lag method. The earliest application of the VAR model to money/income causality appears to be Cuddington (1981).

$$
\begin{aligned}
X_t &= \alpha_{XX}(L)X_{t-1} + \alpha_{XY}(L)Y_{t-1} + e_{X,t}, \\
Y_t &= \alpha_{YX}(L)X_{t-1} + \alpha_{YY}(L)Y_{t-1} + e_{Y,t}.
\end{aligned}
\tag{1}
$$

The expressions $\alpha_{ij}(L)$ are polynomials in the lag operator (as discussed in the chapter on lag distributions) and the e terms are error terms. There are four possible causality outcomes in (1): no causality, X causes Y but not vice versa, Y causes X but not vice versa, and "feedback" or causality in both directions. These possibilities are shown in Table 2 along with the coefficient restrictions in (1) that are implied. These coefficient restrictions are usually tested with standard F tests of whether the set of coefficients in the relevant α lag polynomial are collectively zero.

Table 2. Coefficient restrictions in VAR Granger causality tests

		Does $X \Rightarrow Y$?	
		Yes	No
Does $Y \Rightarrow X$?	Yes	$\alpha_{XY} \neq 0$ $\alpha_{YX} \neq 0$	$\alpha_{XY} \neq 0$ $\alpha_{YX} = 0$
	No	$\alpha_{XY} = 0$ $\alpha_{YX} \neq 0$	$\alpha_{XY} = 0$ $\alpha_{YX} = 0$

The VAR implementation highlights two crucial assumptions of the Granger causality test. First, Granger causality cannot distinguish causality that is strictly contemporaneous. If Y_t affects X_t but the effect is completely exhausted in the current period, then past values of Y_t may be unhelpful in predicting X_t and thus the Granger procedure would detect no causality. Thus, Granger causality test are predicated on an assumption that causal relationship in time series data are never strictly contemporaneous—all causal effects occur over time.

The second assumption underlying Granger causality is that the present cannot cause the past. Granger causality tests are designed to detect correlation between current values of X and past values of Y. If such correlation is found, it is interpreted as a causal effect running from Y to X. While it may seem intuitively reasonable that past events are not influenced by current ones, many important counterexamples exist, especially when one variable is easily predictable.[5] For example, sales of heavy clothes typically rise during the fall *before* the onset of cold weather. Granger causality might interpret this empirical data as indicating that the clothes purchases cause the cold weather! Clearly, a lot of common sense is needed to avoid misinterpreting Granger causality tests.

3.1.2 The lead/lag method of testing for Granger causality

Although most of the recent literature has used the VAR method, the earliest tests of causal relationships used the lead/lag method. This method uses a single equation to test the time-series structure of the relationship between "prewhitened" versions of the two variables.[6] The hypothetical *causing* variable is on the left-hand side of this regres-

[5] This is directly related to the "post hoc ergo propter hoc" fallacy for which Tobin (1970) criticized Friedman and Schwartz.

[6] A prewhitened series is one that has been purged of serial correlation. Each variable is replaced by its "innovation," the component of the variable that cannot be predicted from its own past values. Sims prewhitened by replacing each variable Z_t by $\log Z_t - 1.5 \log Z_{t-1} + 0.5625 \log Z_{t-2}$.

sion and both leads and lags of the variable being caused are on the right. In other words, to test whether Y causes X, one estimates a regression of the following form with the tilde indicating the prewhitened version of each variable:

$$\tilde{Y}_t = \beta_0 + \sum_{i=0}^{p} \beta_i \tilde{X}_{t-i} + \sum_{j=1}^{q} \gamma_j \tilde{X}_{t+j} + \varepsilon_t . \tag{2}$$

Lag and lead lengths p and q must be specified in advance. Sims shows that you can test Granger causality by testing the null hypothesis that all of the γ coefficients in (2) are zero, which is a simple F test.

In addition to the caveats discussed above for the VAR approach, the lead/lag approach also requires that the data be prewhitened. Since different authors might use different time-series filters to prewhiten the series, this adds an additional point of controversy to causality analysis.

3.2 Sims's bivariate test of money and income

Sims (1972) uses equation (2) to test for causality between nominal GNP and two measures of the money supply (M1 and the monetary base) using quarterly data from 1949 to 1968. Table 3 shows Sims's reported F statistics for the hypothesis that the future lags in (2) are zero. Both tests of GNP causing money have very small test statistics that are statistically insignificant. The tests at the bottom of the table show that both monetary variables appear to have strong causal effects on GNP.

Table 3. Sims's test statistics for causality between GNP and money

Causality test	*F* statistic
GNP ⇒ M1	0.36
GNP ⇒ Monetary base	0.39
M1 ⇒ GNP	4.29*
Monetary base ⇒ GNP	5.89*

*Significant at the 0.05 level.
Source: Sims (1972), Table 3.

Sims's evidence of unidirectional causality from money to income was interpreted as support for a traditional monetarist or Keynesian hypothesis that monetary policy has real effects.[7] However, a flood of alternative evidence soon muddied the picture. For example, Feige and Pearce (1979) show that changing the prewhitening filter switching to the VAR methodology for Sims's sample leads to a conclusion of no causality in either direction. Sims himself (Sims (1980)) finds much weaker evidence for causality from money to income when interest rates are added as an additional variable in a VAR system. Eichenbaum and Singleton (1986) find further that results depend crucially on whether the data series are made stationary by differencing or by adding a time trend.

[7] Though note that Sims (1972) and the papers that immediately followed used nominal rather than real GNP as the income variable. Thus, a causal effect of money on GNP could be on prices rather than on real activity. Later studies typically used either real GNP (or GDP) or industrial production.

Stock and Watson (1989) summarize the voluminous early literature as "researchers using only slightly different specifications have reached disconcertingly different conclusions using postwar U.S. data." They attribute the disagreement to the effects of three aspects of the specification: (1) how the data are made stationary, (2) whether or not interest rates are included in the specification, and (3) the choice of sample period.

Sims did his test with future prediction. He showed equivalence to Granger definition. Hosoya (1977) shows more general than that. First study I've found that uses VAR methodology is Cuddington (1981). No. Feige and Pearce (1979) cite Sargent (1976) as the first application and compare three methods, (Sargent prewhitens by spectral techniques). Feige & Pearce also have long citation list of early causality studies on p. 522. Barth and Bennett (1974) do early test for Canada.

4 Rational expectations: Anticipated and unanticipated money

The gradual increase in steady inflation from low levels in the early 1960s to near double digits by 1980 identified devastating gaps in mainstream macroeconomics, both theoretical and empirical. On the theoretical side, early work by Friedman (1968) and Phelps (1968) on the effect of expected inflation on the Phillips curve led to the new classical macroeconomic models of Lucas (1972), Lucas (1975), Sargent and Wallace (1975), and Sargent and Wallace (1976).

The key feature of the new classical macroeconomics was the distinction between changes in monetary policy that agents in the economy anticipate correctly and changes that take people by surprise. These models predict that correctly anticipated monetary shocks should have neutral effects: proportional effects on prices and other nominal variables and no effect on real variables. Unanticipated monetary changes could have real effects in these models.

The development of the new classical models focused the attention of macroeconomists on the formation of expectations by economic agents. Lucas closed this gap theoretically by integrating the idea of "rational" expectations, developed by John Muth (1961). According to the rational expectations hypothesis, individuals form optimal forecasts of macroeconomic variables based on all of the information they have at the time the forecast is made. They are usually assumed to have knowledge of the actual structure of the macroeconomy and to incorporate this knowledge in forming their expectations.

In order to test the new classical models, rational expectations had to be implemented empirically. The earliest and most common approach was to assume that individuals' expectations of a variable could be represented as the optimal prediction from a least-squares regression of that variable on a set of "information" variables. Any variable that could be observed by agents at the time the expectation was formed (i.e., lagged variables) and that would be expected (theoretically or empirically) to be informative could be used as an information variable.

The above definition of an information variable would exclude few of the thousands of entries in a macroeconomic database. However, degrees-of-freedom considerations dictate that only a few information variables can be included in any single expectations equation. This allowed considerable latitude for variation in the choice of information variables, and a corresponding variation in econometric results.

Robert Barro was the first to use this method to test the monetary policy implications of the new classical theory. Barro (1977) looks at the effects of anticipated and unanticipated money on unemployment in the United States; a companion paper, Barro (1978), examines the effects on real output and prices.

Formally, Barro's approach can be written as two equations. The first is the equation for money growth that is used to proxy for expectations.

$$DM_t = \alpha_0 + \sum_{i=1}^{K} \alpha_i X_{i,t} + \varepsilon_t, \tag{3}$$

where DM_t is the rate of growth of the money supply (Barro uses M1), $X_{i,t}$ for $i = 1, 2, \ldots, K$ are variables that help predict money growth and that are in agents' information set at time t, and ε_t is an error term that represents the unpredictable component of money growth. Note that the X variables must either be lagged variables or must be variables that can be observed very quickly, so that agents know their current-period values before forming expectations about current money growth.

The second equation uses equation (3) to break down money growth into anticipated and unanticipated components, and tests for the effects of each on the variable of interest: unemployment, output, or the price level. If we denote this variable by y and the non-monetary variables that help determine y by Z_i for i = 1, 2, … , J, then we can write this equation as

$$y_t = \beta_0 + \sum_{j=1}^{J} \beta_j Z_{j,t} + \sum_{s=0}^{S} \gamma_s \left(\alpha_0 + \sum_{i=1}^{K} \alpha_i X_{i,t-s} \right) + \sum_{s=0}^{S} \theta_s \varepsilon_{t-s} + \eta_t. \tag{4}$$

where the terms in the third summation measure the effects of unanticipated changes in money growth and the second summation is anticipated changes.

The new classical theory concludes that for unemployment and output, all of the γ coefficients in (4) should be zero, while the θ coefficients should be nonzero. The θ values should be negative in the unemployment equation (unexpected increases in money growth should lower unemployment) and positive in the real output equation (they raise output).

Barro did his estimation in two steps. First he estimated a simple version of equation (3) for money growth using annual data from 1941 to 1973. [8] The X variables in Barro's specification are two lagged values of money growth; the current value of $FEDV$, a measure of the deviation of real federal expenditures from their "normal" (adaptively smoothed) value; and a lagged value of UN, which is the log of $[U/(1 - U)]$, with U representing the unemployment rate. Equation (5) shows the results of this estimation.

$$\underset{}{DM_t} = \underset{(0.031)}{.087} + \underset{(0.15)}{0.24}\, DM_{t-1} + \underset{(0.13)}{0.35}\, DM_{t-2} + \underset{(0.015)}{0.082}\, FEDV_t + \underset{(0.010)}{0.027}\, UN_{t-1} \tag{5}$$

[8] The equation shown is from the former paper. The latter paper uses a similar equation, but the coefficients are slightly altered by the extension of the sample through 1976.

The R^2 for equation (5) is 0.90, which means that agents using this rule to form expectations would have anticipated 90 percent of the variation in money growth.

The second step of Barro's procedure was to estimate variants of equation (4) for unemployment, output, and prices replacing the unobservable ε with the residuals, which he called DMR_t, from equation (5) and using his estimates from equation (5) to replace the α coefficients in (4). The dependent variables were *UN*, the log function of the unemployment rate discussed above, the log of output, and the log of the price level.

To estimate equation (4), Barro had to specify the set of "control variables" Z for each equation. In the unemployment equation, he included a variable that measured the share of the 15–44 male population in the military in years where there was a military draft (zero in other years) and a variable measuring the minimum wage as a share of average private, non-farm wage rate, multiplied by the share of the work force covered by minimum-wage laws. In the output equation, his control variables were the military draft variable and a time trend. His preferred specification of the price equation contained a time trend, the share of federal spending in GNP, and the nominal AAA corporate bond interest rate.

The unemployment and output equations were estimated in the form of (4), selectively including and excluding the blocks of variables corresponding to anticipated money growth and unanticipated money growth to test their effects. The price equation was estimated in a form equivalent to (4):

$$\ln P_t = \beta_0 + \sum_{j=1}^{J} \beta_j Z_{j,t} + \gamma_0 DM_t + \sum_{s=1}^{S} \gamma_s \left(\alpha_0 + \sum_{i=1}^{K} \alpha_i X_{i,t-s} \right) + \sum_{s=0}^{S} \theta_s \varepsilon_{t-s} + \eta_t . \tag{6}$$

In equation (6), new classical theory predicts that $\gamma_0 = 1$ that $\gamma_s = 0$ for $s > 0$, and that $\theta_s < 0$. The negative coefficients on unanticipated money growth reflects the fact that unexpected increases in the money supply should lead to a less-than-proportional effect on prices in the short run.

Table 4 reports Barro's results. The estimated coefficients shown are those for the regression with anticipated money growth excluded from the equation, in other words, with the new classical hypothesis imposed. The test statistics reported at the bottom of the table are for tests of two hypotheses. The first tests the hypothesis that the γ terms are zero, which means that anticipated money growth has no effect. The second tests whether the θ coefficients are zero, or that unanticipated money growth does not affect the dependent variable.

The new classical theory under rational expectations claims that only unanticipated changes in the money supply should have an effect on real variables such as unemployment and output. The first two columns of Table 4 seem to support this conclusion. The F tests of whether the unanticipated-money variables can be excluded (i.e., whether the θ coefficients are zero) are rejected. However, the corresponding F tests indicate that including the anticipated-money measure fails to improve the fit of the equation significantly, so Barro cannot reject the new classical null hypothesis that anticipated money growth does not affect real variables (i.e., that the γ coefficients are zero).

Table 4. Barro's estimates of the effects of anticipated and unanticipated monetary policy (standard errors in parentheses)

Explanatory variable	Dependent variable		
	$\log[U/(1-U)]$	log of real output	log of price level
Constant	–3.07 (0.15)	2.95 (0.04)	–4.60 (0.26)
DM_t	–	–	1.02 (0.07)
DMR_t	–5.8 (2.1)	1.04 (0.21)	–0.74 (0.17)
DMR_{t-1}	–12.1 (1.9)	1.21 (0.22)	–1.48 (0.21)
DMR_{t-2}	–4.2 (1.9)	0.44 (0.21)	–1.79 (0.25)
DMR_{t-3}	–	0.26 (0.16)	–1.36 (0.23)
DMR_{t-4}	–	–	–0.72 (0.20)
DMR_{t-5}	–	–	–0.34 (0.16)
Military draft variable	–4.7 (0.8)	0.55 (0.09)	–
Minimum wage variable	0.95 (0.46)	–	–
Federal spending share	–	–	0.59 (0.14)
Nominal AAA bond rate	–	–	3.7 (1.1)
Time trend	–	0.0354 (0.0004)	–0.0108 (0.0020)
Sample period	1946–73	1946–76	1948–76
R^2	0.78	0.998	0.999
Durbin-Watson	1.96	1.8	1.8
Test of all $\gamma_s = 0$	$F^3_{19} = 1.4$	$F^4_{20} = 0.2$	$F^6_{13} = 1.7$ [a]
Test of all $\theta_s = 0$	$F^3_{19} = 15.7$	$F^4_{20} = 3.6$	$F^6_{13} = 7.9$
5% critical value for F tests	3.1	2.9	2.9

Sources:
Unemployment equation from Barro (1977), equation (4). Output and price equations from Barro (1978), equations (3) and (8), respectively.

Notes:
[a]This test is for the exclusion of current and five lagged values of anticipated money growth, given the current money growth (*DM*) enters with a coefficient constrained to unity.

The new classical model asserts that anticipated changes in money should change prices in equal proportion, but that unanticipated changes may have a smaller effect. The price equation reported in the right-hand column of Table 4 supports these effects. The actual current change in the money has an estimated coefficient (γ_0 in equation (6)) of 1.02 that is not statistically different from 1. Given that price is measured in log terms, this implies that a one percentage point increase in money growth leads to a one-percent increase in prices—the predicted proportional effect.

Given the unitary coefficient on current money growth, any additional monetary terms reflect deviations from an exact proportional effect. According to the new classical theory, additional anticipated money changes should have no marginal effect (given the proportional effect of current money growth already present in the equation), but unanticipated changes should enter with a negative sign (reflecting the less-than-proportional effect that these changes have). Barro finds $\theta_s < 0$ at all lags, indicating a negative (and statistically significant) effect of unanticipated changes, and no significant effect of anticipated changes ($\gamma_s = 0$ for $s > 0$ is not rejected).

Like the new classical theory itself, Barro's results were greeted skeptically by macroeconomists. A vast literature emerged estimating variations on Barro's studies with different information variables included in the money-growth (expectations) equation, different control variables included in the output, unemployment, and price equations, different sample periods, and different econometric procedures. Some studies concurred with Barro's conclusions in support of the new classical model, but in the majority of specifications both unanticipated and anticipated changes in the money supply seemed to affect real variables (though not necessarily in symmetric ways).

One difficulty of Barro's approach is a key identification restriction that must hold in order for the results to be valid. If all of the X variables that help determine money-growth expectations are also included in the set of Z variables that have independent effects on y, then the second set of terms in (4) will be collinear with the first set and the β and γ coefficients cannot be separately identified. Thus, Barro's test strategy depends crucially on being able to find one or more variables that affect money growth (so they should be in X) but have no separate effect on output or unemployment (and therefore can be excluded from Z). Moreover, as Sargent (1976) pointed out, lagged values of money growth cannot fulfill this need for identifying variables because the lag length S in equation (4) is never confidently known a priori and lagged money terms can never be ruled out as appearing in (4).

In Barro's equations, the *FEDV* variable is the key identifying variable that is assumed not to have an independent effect on unemployment and output. Given that many Keynesian models attribute strong aggregate-demand effects to fiscal policy, this identifying assumption in open to challenge. Moreover, Small (1979) argued that the *FEDV* variable did not properly capture the temporary effects of wartime changes in federal spending and that this shortcoming strongly affected Barro's results. Small also demonstrated problems with Barro's choice of Z variables in the unemployment equation.

One prominent study that finds results less supportive to the new classical model is Mishkin (1982). In addition to using a different set of variables for X and Z, Mishkin uses a more sophisticated econometric technique that combines both steps of Barro's analysis into one simultaneous, non-linear system. He substitutes for ε in equation to get

$$y_t = \beta_0 + \sum_{j=1}^{J} \beta_j Z_{j,t} + \sum_{s=0}^{S} \gamma_s \left(\alpha_0 + \sum_{i=1}^{K} \alpha_i X_{i,t-s} \right) + \sum_{s=0}^{S} \theta_s \left[DM_{t-s} - \left(\alpha_0 + \sum_{i=1}^{K} \alpha_i X_{i,t-s} \right) \right] + \eta_t . \quad (7)$$

Equation (7) contains only observable variables. It can be estimated by nonlinear least squares either separately or simultaneously with equation (3), imposing the restrictions that the α coefficients be the same in both equations. Testing the latter restriction allows a test of whether the expectations implicit in (7) correspond, as they would under rational expectations, to the optimal forecasts of (3).

Mishkin ran regressions for unemployment and the log of output using quarterly data from 1954–76. Four lags of money growth, the Treasury bill interest rate, and the high-employment Federal government budget surplus were used by agents as X variables to predict money growth. Only a time trend was included in Z as a control variable. Mishkin's results are highly sensitive to S, the length of the lag with which money is allowed to affect unemployment and output.[9] For short lags (8 quarters), the results failed to reject either the rational-expectations restrictions (equality of the estimated α coefficients between equations) or zero coefficients (γ) on anticipated money growth terms. However, for 20-quarter lags, both hypotheses are decisively rejected for both unemployment and output.

Subsequent studies used widely varying sets of X and Z variables, different lag lengths, and data from many countries to examine the effects of anticipated and unanticipated money growth. The results varied widely with some studies supported Barro's original results, some finding that *neither* anticipated nor unanticipated money growth affects real variables, and the majority finding real effects (though not necessarily identical effects) from both anticipated and unanticipated changes.[10]

5 Identifying monetary shocks and their effects

A great deal of attention in the modern empirical literature on monetary policy has been devoted to the identification of monetary shocks. Simply defined, monetary shocks are changes in the monetary policy instrument that are *not* normal, predictable responses to other variables in the economy. The rational expectations theory discussed above emphasizes the difference between expected and unexpected monetary policy moves, and under certain assumptions monetary shocks may be similar to unanticipated changes in money. However, the issue here is purely an econometric one that is present even in the case where unanticipated and anticipated monetary policy changes have identical effects.

Let's suppose, in contrast to the rational expectations theory, that all changes in the monetary instrument have similar effects on the economy. Why is it important to analyze the effects of *shocks* rather than just looking at responses to monetary policy in general? The answer is that only by looking at monetary shocks that are not just a response to other economic variables can we be sure that the results we observe are truly due to monetary policy.

To see why this is the case, consider the following situation. Suppose that the only changes that ever occurred in monetary policy were a result of increases in inflation—i.e., that there were no monetary shocks. We might observe that each time inflation rises,

---

[9] Mishkin also used a polynomial distributed lag technique to constrain the lag coefficients to lie on a fourth-order polynomial and to approach zero at the end of the lag.

[10] A useful summary table categorizing the results of many empirical studies of this question is Table 13.1 of Goodhart (1989).

the monetary authority would pursue contractionary policy and inflation would go back down. We could interpret this result as showing that monetary contraction lowers inflation. However, the same observations would be consistent with a theory that monetary policy has no effect on inflation whatever and that high inflation today leads by itself to lower inflation tomorrow independent of any monetary policy action. If we never observe high inflation without monetary contraction or monetary contraction without high inflation, we can never determine which one causes the lower inflation that follows.

Only when we observe a *deviation* of monetary policy from its usual response to inflation—a monetary shock—can we identify the separate effects of high inflation and monetary policy changes. Thus, identifying the monetary shock is crucial for *identification* of the effects of monetary policy, even if the actual effects of shocks are the same as the effects of responsive monetary policy.[11]

We shall see that economists have followed many different strategies for identifying monetary shocks. Although there are important differences in the results, for a sizable number of identification schemes, the effects of monetary policy on the economy seem to follow a similar pattern that is broadly consistent with the predictions of standard macroeconomic theory.

5.1 The VAR approach

Since it was introduced by Sims (1980), the vector autoregression (VAR) has dominated the analysis of the effects of monetary policy. As discussed in Chapter 7, a VAR consists of a set of regression equations, each of which expresses one variable as a linear function of its own lagged values and lagged values of the other variables in the set. Estimation of a VAR requires no assumptions about which variables might have immediate impacts on which other variables. This was one of the selling points that Sims used in popularizing VAR methods. However, as discussed in detail in Chapter 7, identification restrictions are necessary in order to interpret the relationship between the residuals of the VAR and the underlying "structural" shocks to the variables of the system. Only when these shocks have been identified can the estimated VAR be used to generate impulse response functions describing the effects on the system of a shock to any one of the variables.

To briefly review that argument, consider a simple two-variable VAR with (logs of) money and output as the two variables. The equations of the VAR are

[11] In some ways, this is a "natural experiment" analogous to the controlled experiments used to analyze the effectiveness of medical treatments. If everyone who gets a disease is given the same treatment, then there is no way to identify the effect of the treatment from the natural progress of the disease. Only by comparing the outcomes of treated and untreated (placebo) patients can the treatment effects be measured. The monetary shock can be thought of as the difference between two monetary policy responses to similar conditions. By comparing the macroeconomic outcomes under the two different monetary responses, we are able to isolate the effects on macroeconomic variables of the monetary policy "treatment."

$$y_t = \alpha_{y,0} + \sum_{i=1}^{p} \beta_{yy,i} y_{t-i} + \sum_{i=1}^{p} \beta_{ym,i} m_{t-i} + \eta_{y,t}$$
$$m_t = \alpha_{m,0} + \sum_{i=1}^{p} \beta_{my,i} y_{t-i} + \sum_{i=1}^{p} \beta_{mm,i} m_{t-i} + \eta_{m,t}. \quad (8)$$

The η terms are regression error terms that measure movements in the two variables that are not correlated with past values of either variable. In general, these errors cannot be interpreted directly as structural shocks to m or y. While the structural shocks to the two variables are (by assumption) not correlated with each other, the η terms are linear combinations of the two and will generally be correlated if either variable affects the other contemporaneously.

In order to interpret the η error terms as structural shocks, an identifying assumption must be applied specifying the nature of the contemporaneous effect between y and m. The four possible combinations of assumptions about contemporaneous effects between y and m are shown in Table 5. In order to interpret the correlation between $\eta_{m,t}$ and $\eta_{y,t}$, we usually assume that contemporaneous causality operates in only one direction, either from m to y or from y to m. This is called a "recursive" identification assumption and corresponds to the upper right and lower left cells in Table 5, which are described as "identified."

Table 5. Identifying assumptions and identification of shocks for two-variable VAR.

<table>
<tr><td colspan="2" rowspan="2">Pattern of assumed contemporaneous effects:</td><td colspan="2">Does m_t affect y_t?</td></tr>
<tr><td>Yes</td><td>No</td></tr>
<tr><td rowspan="4">Does y_t affect m_t?</td><td rowspan="2">Yes</td><td>Not identified</td><td>Identified</td></tr>
<tr><td>Not possible to infer anything about structural shocks.</td><td>$\eta_{y,t}$ is output shock. The part of $\eta_{m,t}$ that is orthogonal to $\eta_{y,t}$ is monetary shock.</td></tr>
<tr><td rowspan="2">No</td><td>Identified</td><td>Overidentified</td></tr>
<tr><td>$\eta_{m,t}$ is monetary shock. The part of $\eta_{y,t}$ that is orthogonal to $\eta_{m,t}$ is output shock.</td><td>Each η is a structural shock. $\eta_{y,t}$ and $\eta_{m,t}$ are uncorrelated or assumptions are rejected.</td></tr>
</table>

Using the lower left cell as an example, if money affects output in the current period but output does not have an immediate effect on money, then the only reason that the two current η terms are correlated is because η_y is responding to the monetary shock. Because output shocks in period t do not affect m_t, η_m is not affected by the structural output shock and reflects only the monetary shock. In this case, we can interpret η_m as the structural monetary shock, while the structural output shock can be calculated as the part of η_m that cannot be predicted by (is orthogonal to) η_y.

Making the opposite recursive assumption—that output affects money immediately but not vice versa—gives the opposite pattern of identification: η_y is the output shock and the part of η_m that is orthogonal to that is orthogonal to η_y is the monetary shock. Once the shocks have been identified by one of these assumptions, the impulse response functions showing the time pattern of each shock's effect on both variables can be computed. The impulse response function for the effect of the monetary shock on output is of central interest.

Which of the two possible identifying assumptions should one choose? If η_m and η_y are highly correlated, then the two assumptions are likely to lead to quite different time series for the identified shocks. These may, in turn, have very different patterns of impulse responses. Because these are two ways of interpreting a single set of regression results, nothing in the data can tell us which (if either) of these two identification schemes is correct. Instead, one must rely on theory or outside empirical evidence (such as micro-level studies) in order to assess which is the more plausible identification.

The simple two-variable case examined above generalizes readily to more than two variables. In the multivariate case, recursive identification involves ordering the variables in a causal sequence in which variables positioned lower on the list are assumed to have no immediate effect on variables above them. In his seminal VAR analysis for the U.S. and Germany, Sims (1980) used six variables and assumed the following causal ordering: money, output, unemployment, nominal wages, prices, and import prices. In other words, he assumed that none of the other variables had any effect within the current quarter on the money supply (which was placed at the top of the list), while the money supply could affect all other variables contemporaneously.

Sims summarizes his estimated effects of monetary policy for the two countries as follows:

> In the U.S. money innovations have very persistent effects on both money and other nominal variables. In Germany, money innovations, though larger, are much less persistent. The peak effect of the money innovation on real GNP is much bigger for the U.S. than for Germany. [Sims (1980)]
>
> [In both countries, money innovations tend temporarily to increase the real wage and real GNP and to reduce unemployment, with an opposite swing in these variables following. [Sims (1980)]

While these basic conclusions have been supported by many subsequent studies, a large literature has explored the implications of alternative VAR systems.

5.1.1 Points of controversy in VAR studies of monetary effects

Conducting a VAR study of the effects of monetary policy requires specification of many aspects of the analysis. Among these are

- Variable to measure monetary policy
- Recursive or other identification assumptions
- Variables to include other than the monetary policy variable
- Observation interval (annual, quarterly, monthly)
- Sample period
- Lag length

Although all are potentially important, most of the controversy has revolved around the first three questions: How to measure monetary policy, how to identify shocks to the chosen monetary policy variable, and what variables to include in the full VAR system.

Sims, in his seminal study and in subsequent work, favored using the narrow M1 definition of the money supply to measure monetary policy. However, because the money supply is affected not only by policy actions but also by the actions of banks and holders of currency and deposits, other economists have usually chosen a variable reflecting the state of the market for bank reserves, over which the central bank exerts more direct control. Common choices include the stock of nonborrowed reserves, the amount of borrowed reserves, and the interest rate on overnight, inter-bank loans (the federal funds rate in the United States). Some studies have included more than one of these variables and have used identification restrictions based on assumptions about the reserves market to identify monetary shocks as a composite of the error terms on the variables.

Recursive identification schemes have been the norm, but "structural VARs" incorporating non-recursive identification assumptions became more common starting in the 1990s. Such structural models incorporate assumptions about the interaction of the variables of the system that do not follow the strict causal ordering of a recursive system. Some such assumptions follow the "X has no immediate effect on Z" pattern, but others are more sophisticated. For example, some authors have used the assumption of long-run neutrality of monetary shocks as an identifying assumption: the monetary shock has no effect on real output in the long run.

The set of variables in the VAR usually includes at least one interest rate (which may be the monetary policy variable), at least one money or reserve variable (another possible choice for the monetary policy variable), real output, and a price index. Studies for countries other than the United States typically also include a real or nominal exchange rate. Unemployment is often included, as is an index of commodity prices, which has proved helpful in eliminating one of several anomalous results in more basic setups.

5.1.2 Selected VAR results

It is impossible to review comprehensively here the voluminous VAR literature on monetary policy. Recent papers that examine the literature in some detail include Christiano, Eichenbaum, and Evans (1999) and Leeper, Sims, and Zha (1996). This section will examine several widely cited studies that demonstrate typical results and highlight particular puzzles that have arisen.

The coefficients of VAR models are numerous (generally the number of lags times the square of the number of variables) and not subject to ready interpretation. The interpretation comes with the application of the identification assumptions to estimate impulse response functions showing how the various shocks affect each variable over time. The impulse responses are easier to read as graphs than as tables, so the standard presentation of VAR results is as a matrix of small graphs. Each column of the matrix describes the effects of one shock on the variables of the system, with the variables being affected changing across the rows.

5.1.2.1 Money supply as policy measure and the liquidity puzzle

The original VAR analysis of monetary policy in Sims (1980) measured monetary policy by M1. No interest rate was present in his six-variable VAR. The impulse responses in such systems with a monetary aggregate and no interest rate tend to reflect the standard effects of monetary policy: a sustained expansionary policy increases output fairly quickly and raises prices more slowly but more permanently.

However, in later work Sims added an interest rate to the VAR system, which changed the results significantly and presented a puzzle for interpretation.[12] First, when a short-term interest rate such as the federal funds rate or the three-month Treasury bill rate is included, the effects of interest rate shocks on real variables swamp those of shocks to the quantity of money. Second, while the effects of a shock to the interest rate have effects that look in many ways like the hypothesized effects of monetary policy, money shocks have quite different effects. Increases in the money supply seem to lead to very small increases, rather than large decreases, in interest rates, which contradicts the liquidity effect that monetary expansion is supposed to have. The anomalous effect of money on interest rates has been called the "liquidity puzzle."[13]

The liquidity puzzle has often been resolved by reinterpreting the causes of the shocks. Under the revised interpretation, the monetary policy shock is taken to be the change in interest rates and the shock to the quantity of money is interpreted as a disturbance to money demand.

If we think of the quantity of money and the interest rate as being determined by a market equilibrium mechanism to balance supply and demand forces, then we would not ordinarily expect that *either* changes in price or changes in quantity would be purely supply (policy) effects. Assuming that the demand for money is neither perfectly elastic nor perfectly inelastic, only under the (identifying) assumption that the policymaker supplies money perfectly inelastically would the quantity of money respond *only* to policy changes. Although this simplifying assumption is often made in the exposition of the IS/LM model, it may not be very realistic.

The first reason that the money supply may not reflect pure policy actions is that measures such as M1 and M2 include bank deposits, which are not under the direct control of the policymaker. The textbook theory of the "money-supply multiplier" shows that when the public's preference for currency relative to deposits changes or when banks' desired ratio of reserves to deposits changes, then the money supply will change. Only if the central bank anticipates these changes and offsets them by changing the volume of bank reserves by exactly the right amount will the central bank succeed in controlling the quantity of money exactly.

It would be difficult to exercise such close control over the quantity of money, and the evidence is strong that most central banks usually do not even try to do so, at least over short horizons. Since the mid-1980s, the Federal Reserve in the United States has expressed its monetary policy goals in terms of a target level for the federal funds interest rate. If the Fed supplies reserves perfectly elastically at its target level of the interest rate, then the interest rate becomes the policy variable and the quantity of money changes in response to shifts in money (and reserve) demand.

Interpreting the shock to the quantity of money as a money-demand shock solves the liquidity puzzle. If reserves are supplied elastically at the Fed's target interest rate, then an increase in money demand should have a small positive, if any, effect on interest rates, which is what Sims observed.

---

[12] See Sims (1986) and Sims (1992).

[13] For more details on the liquidity puzzle, see Reichenstein (1987) and Leeper and Gordon (1992).

Bernanke and Blinder (1992) find evidence that the Federal Reserve (at least in recent decades) does tend to supply reserves elastically at the target rate chosen at the most recent Federal Open-Market Committee meeting. They also examine the predictive power of five potential monetary policy variables for a monthly U.S. sample. Among the set M1, M2, a long-term Treasury bond interest rate, a short-term Treasury bill rate, and the federal funds interest rate, they find that the federal funds rate has the most predictive power in forecasting a collection of eight monthly macroeconomic variables such as industrial production, unemployment, and personal income. Based on this evidence, they favor a specification with the federal funds rate as their monetary policy instrument.[14]

5.1.2.2 Estimated effects of federal-funds-rate shocks

As noted above, interpreting the shock to the federal funds rate as a monetary policy shock yields effects on output, monetary variables, and interest rates that correspond to conventional wisdom about monetary policy effects. However, if an increase in the federal funds rate is a contractionary monetary shock, then it should have a strong negative effect on prices. What Eichenbaum (1992) characterized as a "price puzzle" emerged from early VAR systems based on the federal funds rate. In these models, a positive (i.e., contractionary) shock to the funds rate was found to *raise* prices significantly for a one-to-two-year period after the shock.

Sims (1992) explored the price puzzle using alternative specifications on data from five OECD countries. He hypothesized that the reason that positive shocks to the interest rate might seem to cause increases in prices is that both the interest rate increase and the price increase are responses to initial inflationary impulses arising from a variable that is not included in the VAR system. As a candidate to measure such inflationary impulses, he added a price index of commodities that are traded on commodity exchanges to the VAR. Commodity prices are observed very quickly and thus could trigger an immediate counter-inflationary response by the central bank. When this variable was included, it largely eliminated the price puzzle for the United States, Germany, and the United Kingdom.[15] Commodity prices have now become a fixture in monetary policy VARs in order to solve the price puzzle.

Figure 1 is taken from Figure 2 of Christiano, Eichenbaum, and Evans (1999) and shows the effects of a positive (contractionary) shock to the federal funds rate in a seven variable quarterly VAR. Recursive identification is used with output (Y), prices (Price), and commodity prices (Pcom) ordered first, followed by the federal funds rate, then nonborrowed reserves (NBR), total bank reserves (TR), and the money supply (M1 or M2). This ordering assumes that the monetary policy variable, the funds rate, responds to current values of output, prices, and commodity prices, but only with a lag to shocks in reserves and the money supply. With the exception of the bottom right panel, the money supply is measured by M1. The bottom right panel shows the response of M2 to a federal funds shock when the entire system is re-estimated using M2 instead of M1.

[14] It is worth noting that they did not include a reserve measure such as nonborrowed reserves in their set of potential monetary policy variables.

[15] The puzzle still existed in Sims's data for France and Japan even with commodity prices in the VAR.

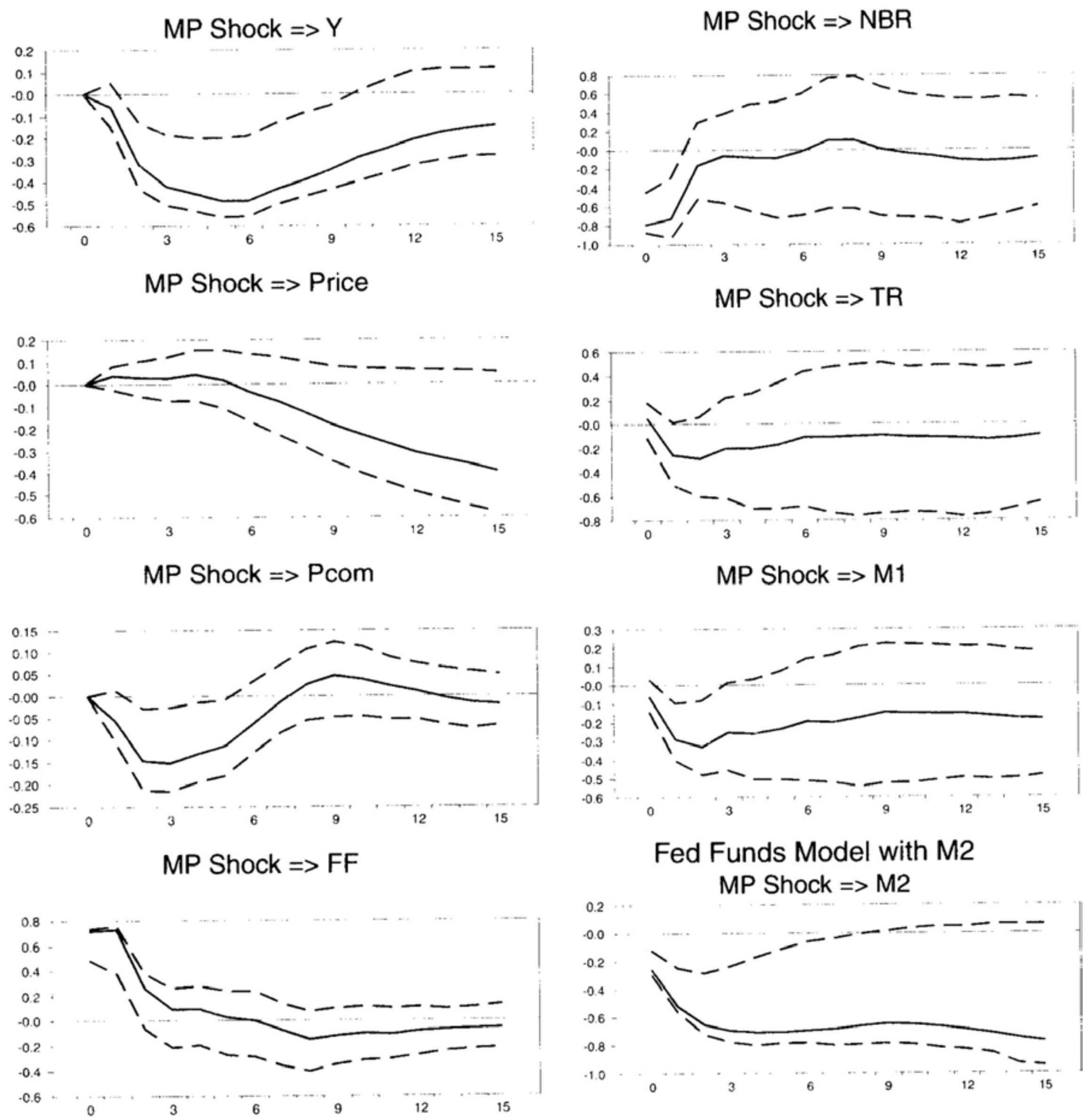

Figure 1. Effects of federal funds rate shock in quarterly model.

Each panel of Figure 1 shows the effect of a positive shock to the federal funds rate on one of the variables of the system over the following 15 quarters. A value of zero means that the shock has no effect on the variable—that the variable continues on the same path it would have followed had there been no monetary shock. A positive (negative) value indicates that the monetary shock would cause the variable to be above (below) its normal path. The solid line depicts the estimated effect, while the dashed lines show the bounds of a 95% confidence interval.

The top left panel shows output falling quickly and significantly following an increase in the funds rate, but that the effect on output appears to be temporary. Output has recovered most of the way to its original path by the end of the 15-quarter horizon shown. These results are consistent with the contractionary effect that such changes are usually assumed to have.

The second panel on the left shows that prices remain stable for over a year, then fall substantially and continue to be lower (and are still falling) through fifteen quarters. Notice that no price puzzle appears with the inclusion of commodity prices in the system. Commodity prices themselves fall much more quickly than the general price level, but do not seem to be permanently affected. Prices that are set in daily trading on commodity exchanges are known to be much more sensitive to economic conditions than common retail prices, which explains the more rapid response.

The bottom left panel shows the effect of a one-time federal funds rate shock on the future path of the funds rate itself. Its pattern suggests that the funds rate is likely to remain at the higher level for one quarter, then decline over the following two quarters back to its original level. The right-hand panels show that both reserve aggregates and both measures of the money supply have the expected negative response to an increase in the funds rate.

To summarize the results in Figure 1, the estimated effects of a contractionary monetary policy shock correspond quite closely with the conventional conclusions of Keynesian and monetarist models. The funds rate jumps and stays up for a few quarters. The paths of reserves and monetary aggregates fall. The path of output declines temporarily, while prices fall more slowly but more permanently.

5.1.2.3 Nonborrowed reserves as a monetary policy variable

We saw above that in VAR systems with a money measure (M1 or M2) and the federal funds rate, interpreting the funds rate shock as the monetary policy measure and the money quantity shock as a shock to money demand leads to results that support conventional theories. However, interpreting the money shock as monetary policy leads to a liquidity puzzle and other contradictory conclusions.

Some economists, notably Lawrence Christiano and Martin Eichenbaum, have advocated using nonborrowed reserves as a measure of monetary policy. There are several reasons for favoring nonborrowed reserves. On a theoretical basis, Eichenbaum (1992) notes that, among the numerous monetary aggregates available for the United States, "the *only* monetary aggregate which the Federal Open Market Committee can directly control is nonborrowed reserves." (Emphasis in original.) Moreover, in terms of empirical performance, "in sharp contrast to the results based on [the monetary base] or M1, inference about the effects of monetary policy on interest rates is very robust when … [nonborrowed reserves] is used in the analysis. Regardless of whether we work with [monetary or interest rate] rules, regardless of whether we work with monthly or quarterly data, and regardless of which postwar sample period we work with, the same result emerges. Unanticipated expansionary policy shocks drive down short-term interest rates for substantial periods of time. Measured in this way, expansionary monetary policy shocks also generate increases in real GNP." (Christiano and Eichenbaum (1992))

Figure 2, also taken from Figure 2 of Christiano, Eichenbaum, and Evans (1999), shows the effects of a contractionary nonborrowed reserves shock on the variables of a similar VAR system with the same recursive identification ordering: the policy instrument is placed after output, prices, and commodity prices, and before other financial sector variables.

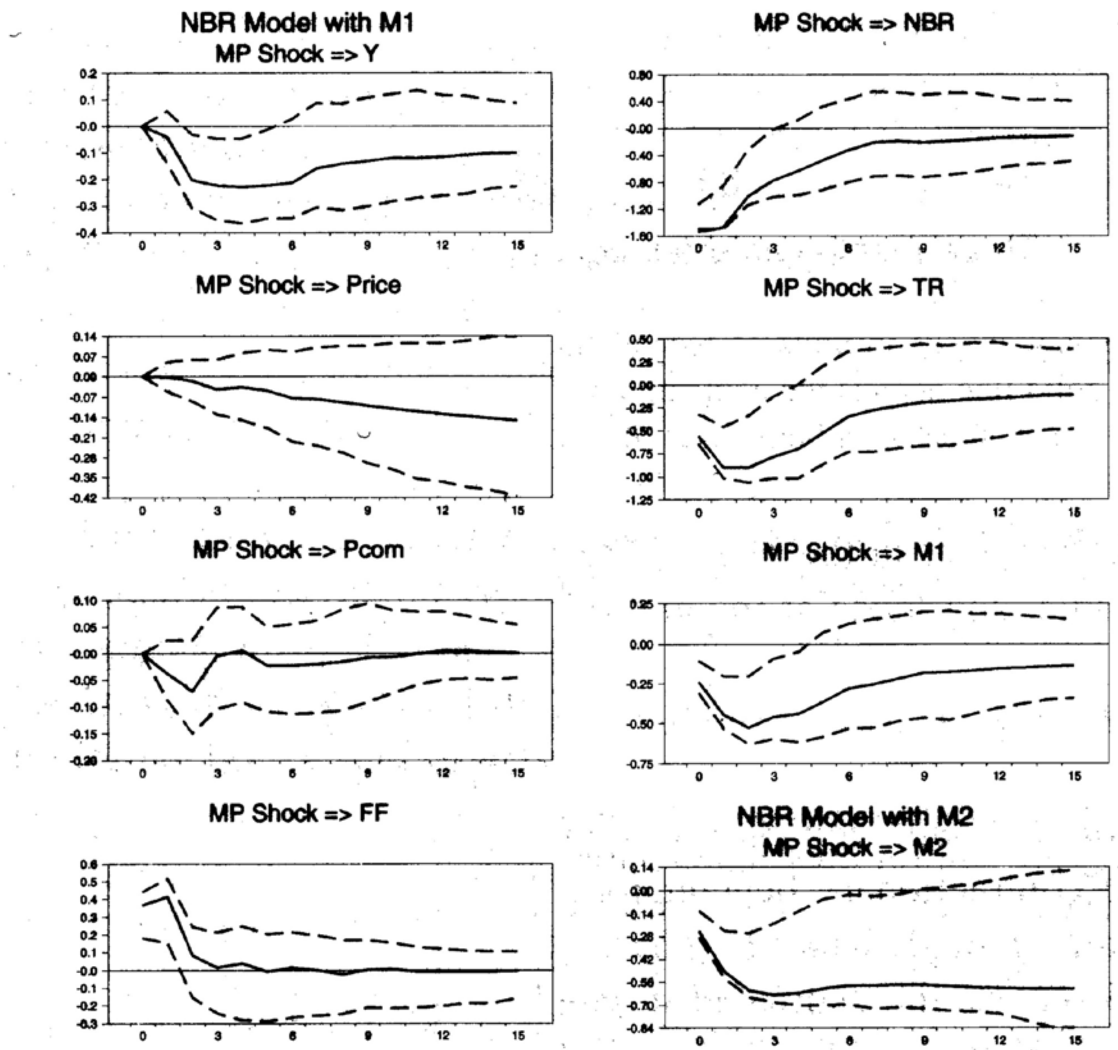

Figure 2. Effects of shock to nonborrowed reserves in quarterly model.

The results in Figure 2 are strikingly similar to those in Figure 1. Based on this evidence, one can tell an empirically credible story that the Federal Reserve has effected monetary policy over the postwar sample by manipulating the market for bank reserves. If one looks at the quantity of nonborrowed reserves as the direct monetary instrument, as in Figure 2, then the federal funds rate responds quickly and in the opposite direction to reserve changes. If one considers the funds rate as the direct monetary variable, as in Figure 1, then nonborrowed reserves change quickly and in the opposite direction as the change in the target interest rate. Although the Federal Reserve has changed its emphasis on quantity targets vs. interest rate targets through the postwar period, the results of Christiano, Eichenbaum, and Evans (1999) suggest that it may be reasonable to think about either policy regime as shifting banks along a fairly stable and downward-sloping demand curve for reserves.[16]

Whichever way one views monetary policy, the effects on non-monetary variables in Figure 1and Figure 2 are similar. Contractionary monetary policy reduces real output and raises interest rates temporarily and lowers prices with a lag.

5.1.2.4 Other identification assumptions in VAR models

Recursive identification schemes are simple to implement and understand, but they do not incorporate all of the knowledge that economists have about the market for bank reserves. It is also impossible to build into a recursive model a hybrid policy rule that attaches importance to both interest rates and reserve aggregates.

In an influential paper, Strongin (1995) based identification restrictions on more detailed information about the reserve market. Strongin argues that neither the federal funds rate nor any single reserve aggregate is a true measure of monetary policy. He claims that "[a]s a matter of actual practice, the Federal Reserve accommodates innovations in the banking system's demand for reserves." [Strongin (1995)] Thus, the initial impact of a tightening or loosening of monetary policy would be felt in the proportions of borrowed vs. nonborrowed reserves. He proposes the ratio of nonborrowed to total reserves as an indicator of monetary policy.

Strongin finds that his measure of monetary policy is successful in eliminating some of the anomalies present in the earlier literature. In particular, his specification is not subject to perverse liquidity effects or to inconsistencies in Granger causality outcomes. However, when commodity prices are added to the VAR system, Christiano, Eichenbaum, and Evans (1999) show that the effects of monetary shocks using Strongin's measure are similar to those of the federal funds and nonborrowed reserves measures.

Cochrane (1998) uses the VAR methodology to examine the evidence on anticipated vs. unanticipated changes in monetary policy. His results suggest that the two kinds of monetary changes may have very different effects. He estimates large real effects comparable to those discussed above for unanticipated monetary shocks, but the effects of anticipated monetary actions are much smaller.

5.1.3 Some objections to VAR models of monetary policy

Some economists have argued that the VAR approach is not an effective test of the effects of monetary policy. For example, Runkle (1987) computes confidence intervals for the impulse response functions and variance decompositions presented by Sims (1980) and finds that the confidence intervals are so large as to raise serious questions about the validity of Sims's conclusions.

Another problem with the VAR framework is that (like other standard econometric procedures) it assumes that the structure of the economy is unchanged over the sample period. We know that the Federal Reserve has emphasized control of different policy instruments during various periods.[17] Such changes presumably alter the structure of the economic responses to monetary policy that are measured by the VAR coefficients. This is problematic for the VAR method unless either the changes are infrequent enough

[16] Meulendyke (1988) discusses the Fed's various targeting regimes during the 1970s and 1980s. In the 1990s, the Fed seems to have focused most strongly on the federal funds rate.

[17] See Meulendyke (1988) for a discussion of alternative instruments up to the mid-1980s.

to allow separate estimation for each sub-period or changes in the structure of the model are built into the VAR.

Finally, Rudebusch (1998) assesses the plausibility of the actual time series of monetary shocks identified by VAR models. He evaluates these shocks by examining whether they tend to be stable, whether they agree with common perceptions of monetary tightness or laxness at any given time, and whether they correspond to measures of unexpected federal funds rate changes derived from futures markets.

He finds that VAR estimates of monetary shocks fail all three tests. Small changes in the specification of the VAR lead to very different estimated shocks, so the shock estimates themselves do not appear to be robust. Moreover, the estimated shocks do not correspond to conventional wisdom (or often the Fed's professed intent) about times when monetary policy was expansionary or contractionary and about the response of policy variables to economic conditions. Finally, the time series of unexpected changes in the federal funds rate derived as the difference between the rate and the futures-market rate from the previous period does not correspond closely with the VAR-estimated shocks.[18]

5.2 The narrative approach to identifying shocks

As discussed above, the VAR approach to econometric identification of monetary policy shocks has yielded mixed results. Many diverse strategies have been used to try to identify monetary shocks in the postwar data, often with divergent outcomes. Because the identification of monetary shocks is crucial to the analysis of the effects of monetary policy, Romer and Romer (1989) augmented traditional macroeconomic data with "narrative" information about the intentions of the Federal Open Market Committee members.

Following a methodology used successfully by Friedman and Schwartz (1963), Romer and Romer examined the published minutes of FOMC meetings to find dates on which the committee chose explicitly to take a contractionary policy stance in order to lower inflation. They found six "Romer dates" on which such policy actions were undertaken: October 1947, September 1955, December 1968, April 1974, August 1978, and October 1979.

Having located these dates by non-econometric means, the Romers proceeded to analyze the behavior of macroeconomic variables during the period after the policy actions. Specifically, they focused on two measures of real economic activity that are observable on a monthly basis: the Federal Reserve index of industrial production and the unemployment rate. For each variable, they performed regressions of the variable on two years (24 months) of its own lagged values. Using these regressions, they generated benchmark forecasts for each Romer date describing how each variable's existing momentum and trajectory might have evolved had their been no policy action. Finally, they compared the actual path of each variable to its projected path.

[18] For a survey of alternative specifications and criticisms, see section 4.4 of Christiano, Eichenbaum, and Evans (1999).

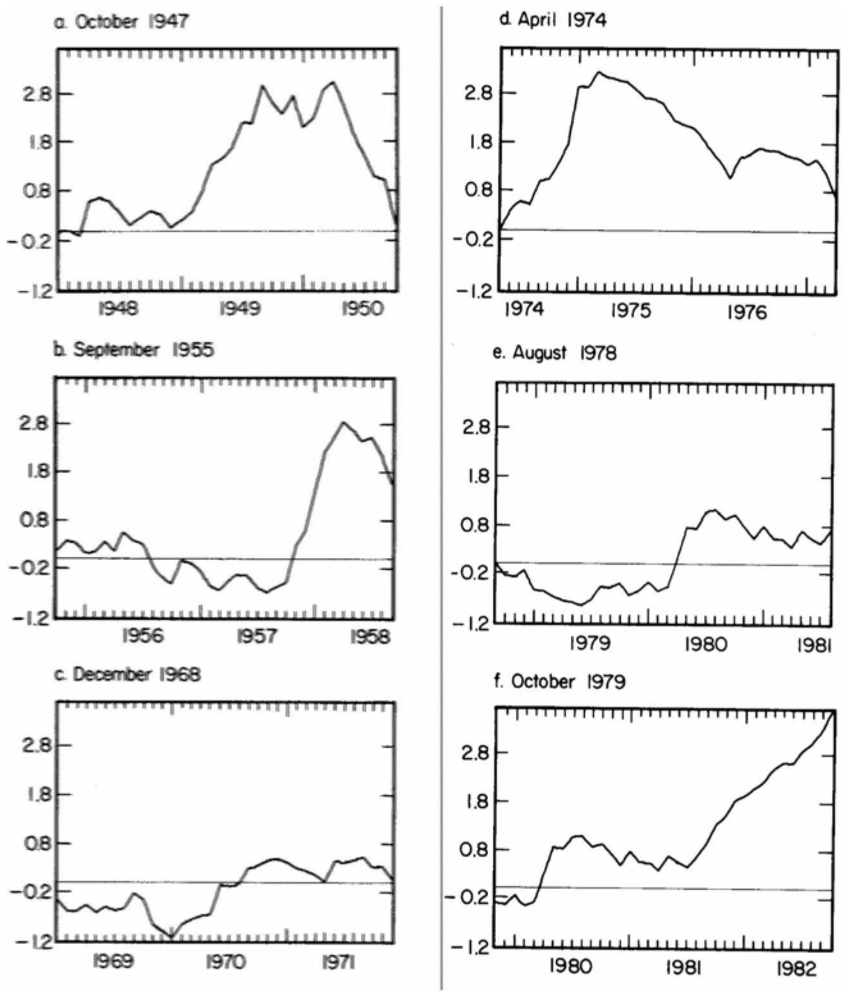

Figure 3. Unpredicted changes in unemployment after Romer dates.

Figure 3 (Romer and Romer's Figure 2) shows the difference between the actual and predicted paths of unemployment after each of the six Romer dates. In most cases, there is a substantial increase in unemployment within about two years after the beginning of the Fed's contractionary policy. The pattern of changes in industrial production (not shown here) shows a related pattern of decreasing within a few quarters of the contractionary shock.

In addition to the evidence based on forecast errors, Romer and Romer perform an additional test. To assess the statistical significance of their measured monetary shocks, they construct a dummy variable that equals one in the months corresponding to Romer dates. This dummy variable is then added to the industrial production and unemployment autoregressions. The effect of the Romer-date dummy is statistically significant and the corresponding impulse response function shows a strong, lagged response of industrial production (negative) and unemployment (positive) to monetary contractions.

6 Channels of monetary policy transmission: Interest rates and the credit channel

The traditional "interest rate channel" through which monetary policy is believed to affect the economy works through a simple IS/LM framework. Contractionary policy leads to a rise in the real interest rate and an increase in the user cost of capital, this causes investment and durable good consumption to drop, finally leading to aggregate demand falling. However, most empirical evidence suggests that monetary policy's direct interest-rate effect is a temporary affect on short-term rates. Such temporary changes in short-term rates should have little corresponding effect on long-term rates. Since long-term rates are the ones that should affect durable-goods spending, monetary policy should lead to mild changes (if any) in variables that depend on this interest-rate. Moreover, the empirical literature on investment spending has struggled to find any empirical association between even long-term interest rates and investment spending. Despite these difficulties, we have seen in earlier sections that monetary contractions appear to have large effects on output.

In particular, Bernanke and Gertler (1995) estimate the reactions of investment and components of consumption to federal funds rate shocks in a VAR system. Their results are shown in Figure 4 (their Figure 3). All components of consumption and investment seems to fall after a monetary contraction, with business fixed investment having the smallest and slowest response.

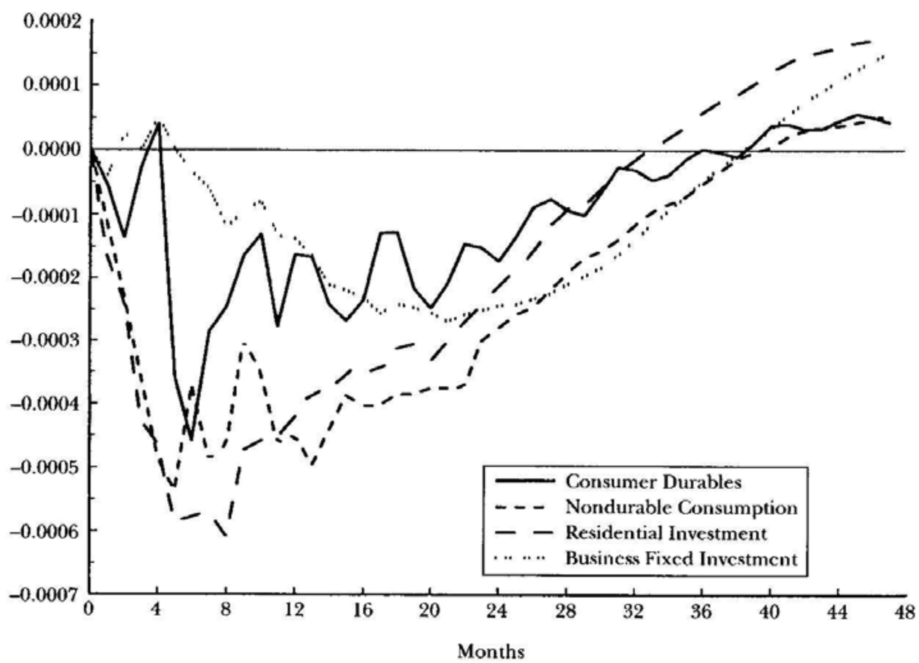

Figure 4. Bernanke and Gertler's estimated responses of expenditure components to monetary contraction.

Given that the individual links in the interest-rate channel seem shaky, some economists have questioned whether it can explain the strong monetary-policy effects discussed above and shown again in Figure 4. It is hard to understand how a small change in the short-term interest rate has led to such a large change in variables that are presumably sensitive to the long-term rate. It is because of this puzzle that economists look for other ways that monetary policy may affect spending. The so-called "credit channel" of monetary transmission proposes a complementary mechanism by which changes in credit conditions resulting from monetary shocks may lead to reinforcing spending declines. We now examine some of the empirical evidence on the credit channel.[19]

6.1 Capital Market Imperfections

In a "perfect" credit market, all the relevant information about the supply of and demand for credit would be summarized in the equilibrium price of credit: the interest rate. In such a world, the tightness or looseness of credit availability would be measured with perfect accuracy by the interest rate. A high real interest rate would indicate costly credit and only those whose need for credit justified the high cost would borrow. Similarly, a low interest rate would show readily available credit. There would be no need to seek other indicators of credit conditions.

Many economists, for example, Jaffee and Russell (1976) and Stiglitz and Weiss (1981), have noted that imperfect information leads to problems of moral hazard and adverse selection in credit markets. These information problems impede the ability of the interest rate to clear the credit market. In the simple model, increases in interest rates are more likely to drive safe borrowers from the market than risky borrowers. If lenders cannot accurately distinguish the riskiness of borrowers, then they cannot screen out risky ones and an increase in interest rates will make their pool of borrowers riskier. Lenders may resist this by keeping interest rates below the market-clearing level and rationing credit based on collateral or credit history to avoid adversely selecting riskier borrowers. In a model such as this, interest rates may not fully convey the tightness of credit conditions.

Is such a situation likely? Bernanke (1983) argues that this type of tailspin is exactly what happened during the Great Depression. The early 1930s saw a widespread fear of bank failures that led to bank runs and rapid withdrawal of deposits. With actual bank failures and less money in the banking system, credit became more expensive. To prove that credit market imperfections have an effect distinct from that of money, Bernanke estimated an output equation for the 1919-41 period. When financial crisis variables are included the fit of the equation improves substantially, leading Bernanke to tentatively conclude a confirmation of the credit hypothesis. If Bernanke's analysis is correct and imperfections play a role in the credit market, we would expect the demand for loans to always exceed the supply of loans.

Further direct evidence of credit-rationing effects come from King (1986), who estimates a structural model of loan supply and demand that allows for the existence of excess demand. He finds evidence of excess loan demand (credit rationing) for many years of his postwar sample. Spikes of excess demand are particularly strong in the late 1960s, middle 1970s, and around 1979, which are periods that Romer and Romer identify as significant monetary contractions.

6.2 Balance-Sheet and Bank-Lending Channels

The credit channel of monetary transmission may operate through either or both of two mechanisms, the balance-sheet channel and the bank-lending channel. The ability of

[19] For an overview of the credit channel, see Bernanke and Gertler (1995).

a firm to borrow or obtain credit depends on two main factors, the state of the credit market in general and the state of the firm's balance sheet and other financial statements. The balance sheet reports the firm's collateral, value of its outstanding debt, and its total net worth. Other financial statements report profits and cash flow.

Each of these variables may be used by lenders as an indicator of the firm's ability to repay loans. A creditor is interested in the collateral of a company, since a default on a loan would result in the handing over of assets (collateral) to the lender. More collateral means easier borrowing. A lender is concerned with the ability of the borrower to repay loans, so cash flow and existing value of debt are closely watched. Total net worth is also a general signal about the solvency of a company; a larger company that has a proven track record should have an easier time raising funds than a newer and smaller firm.

Monetary policy directly affects several of these financial variables. Contractionary policy, which increases interest rates, raises the value of a firm's debt, with a higher level of debt the firm must spend more of its cash servicing the debt, so their cash flow and borrowing ability falls. Higher interest rates depress the value of a company's assets and so they have less collateral from which to borrow. Higher interest rates may also lead to lower expected future sales, due to an overall expected downturn in the economy.

Bernanke and Gertler (1995) try to determine the effects of contractionary policy on cash flow. If the credit channel operates through balance sheet changes, like those described above, a higher interest rate should lead to lower cash flow. They test this hypothesis by creating a VAR that has several components of cash flow, including interest payments (cash used to service debts), profits, gross income and employee compensation (costs of the firm). Using quarterly data from 1965-1994, they identify a contractionary shock as an increase in the federal funds rate. Their Figure 5, shown below as Figure 5 gives evidence in support of the balance-sheet theory. A decline in profits is due in a large part to higher interest payments giving support to the importance of debt. Income falls faster than costs, showing how cash flow becomes an issue after contractionary policy.

The balance-sheet channel described above predicts credit restriction to certain firms from all sources. Although many of the affected firms rely heavily on bank lending, banks have no special role—they act just like any other lender. The bank-lending channel focuses on constraints on bank lending due to changes in bank's balance sheets that occur due to contractionary monetary policy. The selling of open-market securities by the Fed is what technically creates a monetary policy contraction. Banks reserves fall after the open-market sales, which (assuming that they have no undesired excess reserves) requires them to adjust their balance sheets in some way to restore their reserve position relative to their deposits.

This adjustment could be accomplished in several ways. If banks are drained of their reserves, they may attempt to raise money from other sources in order to maintain their lending. One possible source of funds is large certificates of deposit (CDs), which currently carry no reserve requirement and thus can provide a source of funds for banks even when reserves are short. If banks seek more CDs when monetary policy is tight, then CD rates ought to increase at these times relative to other interest rates. Bernanke and Gertler follow the movement of the CD rate relative to the T-bill rate. They show that increases in the T-bill rate (periods of contraction) are met by even larger increases in the CD rate.

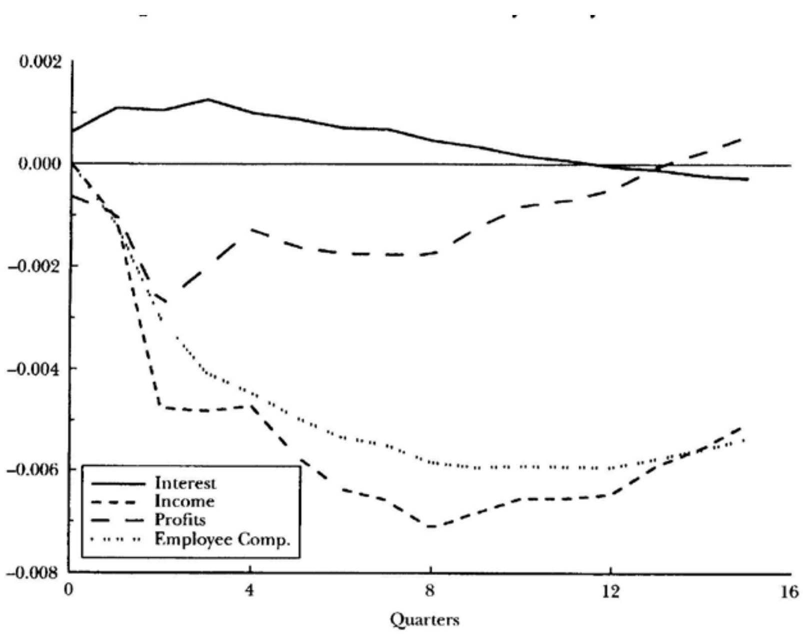

Figure 5. Bernanke and Gertler's estimates of effects of monetary contraction on cash-flow variables.

If banks do not change the mix of deposits to accommodate lower reserves, then they must reduce other assets to attempt to attract additional reserves. The other assets available are marketable securities and loans. The bank-lending channel emphasizes the possibility that monetary contraction will have a direct effect on banks' supply of loans.

Bernanke and Blinder (1992) test whether tight money reduces lending. They use a monthly VAR from 1959-1978 with six lags of the federal funds rate, unemployment rate, log of CPI, log of deposits, log of securities, and log of loans. Monetary shocks are identified as unpredicted changes in the federal funds rate.

They find that after a monetary contraction, banks' holdings of deposits and securities dropped immediately, while bank loans fell only after six months. One can interpret this lagged response of loans as either supporting or refuting a bank-lending channel. Bernanke and Blinder argue that this evidence supports bank-lending effects, since there is a persistent decline in loans.

Romer and Romer (1990), on the other hand, read this as failure of the bank lending explanation. According to their interpretation, banks sell securities in order to maintain their volume of loans in the face of the decline in deposits. Loans then fall late as the general downturn in the economy reduces the demand for them.

Romer and Romer support their story by examining the behavior of loans after the "Romer dates" on which the Fed intentionally pursued contractionary policy. They find that the pattern of bank loans during monetary contractions is not significantly different than at other times, suggesting that bank loans are not affected in a special way by monetary policy.

6.3 External Finance Premium

In a world of perfect information, there is no difference in the cost of raising capital internally or externally. The justly famous theorem of Modigliani and Miller (1958) shows that with fully efficient capital markets firms are indifferent at the margin between issuing bonds, issuing new stock, or using internal retained earnings to finance investment.

When information asymmetries raise problems of moral hazard and adverse selection, firms may be unable to convince potential lenders of the true value of their proposed investments. This makes external finance (through borrowing) more expensive than internal finance (using retained earnings). This difference in costs is called the external finance premium.

Earlier analysis saw contractionary monetary policy lead to similar fall in output and bank loans. Causality between output and bank loans could run either direction. The credit-channel interpretation is that the unavailability of bank loans reduces investment by bank-dependent firms, which lowers aggregate demand and output. An alternative interpretation would be that the reduction in output comes "first," which lowers the demand for bank loans, which then decline in volume. Romer and Romer (1990), discussed above, conclude from their results that the latter interpretation is more likely, thus questioning the importance of the credit channel.

Kashyap, Stein, and Wilcox (1993) (KSW) test the effects of credit-channel variables in several ways. First, they examine the reaction of bank loans and of the commercial paper market to monetary contractions, using the four Romer dates since the beginning of their sample in 1963, augmented by the credit crunch of 1966, which they argue is a significant credit contraction associated with monetary policy. They show that after monetary contractions, the mix of bank loans and commercial paper financing shifts significantly away from loans into the commercial paper market. Figure 6 (panel c of their Figure 2) shows the time path of new bank loans as a share of total new short-term external finance (the "mix") following the five monetary contractions. The solid straight line shows the average growth rate of the mix variable and the bold solid line is the average response after the four Romer dates.

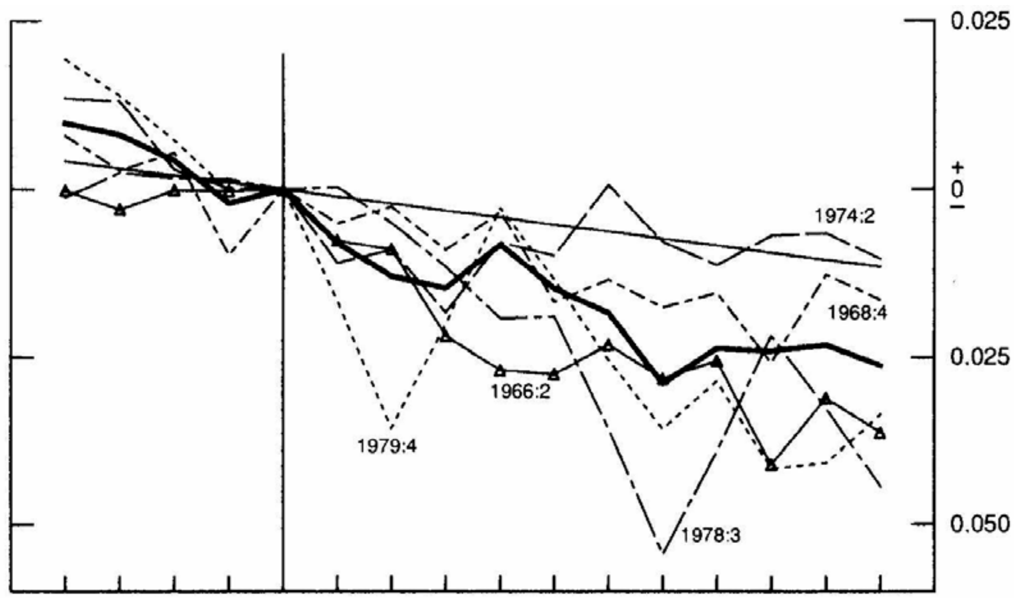

Figure 6. The mix of bank loans and commercial paper after monetary contractions.

As shown in Figure 6, the external finance mix tends to swing strongly away from bank-intermediated loans and toward direct placement of commercial paper during monetary contractions. In KSW's interpretation, large firms with established credit succeed in raising capital through the commercial paper market, whereas smaller and less-known firms who cannot participate in that market lose access to credit altogether. This is consistent with a bank-lending channel of monetary policy.

To formalize the significance of the effect of monetary contractions on credit variables, KSW perform Granger causality tests. They find the monetary contractions have a statistically significant effect in predicting subsequent movements in two key credit variables: the mix defined above and the spread between the prime interest rate and the rate on commercial paper. A response in this interest rate spread may indicate the presence of a bank-lending effect because if banks are constraining their lending strongly, we should see the prime interest rate rise relative to the open-market rate on commercial paper.

The tests discussed above provide evidence that monetary contractions affect bank-lending measures. But do those bank-lending measures then have an effect on spending on new capital? To test this second stage of possible bank-lending effects, KSW examine whether two measures of bank-credit availability affect various categories of business investment spending. Table 6 (from KSW's Table 2) shows the probability values associated with an F test of the hypothesis that the credit variable (mix or price-CP spread) has no effect on the investment variable. At conventional significance levels, a probability value smaller than 0.05 indicates rejection of the hypothesis, which means that the credit variable helps explain subsequent changes in investment.

Table 6. Tests for explanatory power of mix and prime-CP spread in investment equations

Category	Mix			Prime-CP Spread		
	Accelerator	Neoclassical	Q	Accelerator	Neoclassical	Q
Nondurable inventories	0.173	0.009	—	0.022	0.001	—
Durable inventories	0.002	0.049	—	0.500	0.302	—
Producers' durable equip	10^{-6}	0.003	0.026	10^{-7}	0.003	0.027
Nonresidential structures	0.460	0.391	0.175	0.537	0.011	0.037

KSW tested the credit variables in three traditional investment models: an accelerator model in which investment depends on the change in real output, a neoclassical model in which investment is determined by the cost of capital, and a q model in which the value of equities relative to the replacement cost of capital determines investment.

The results summarized in Table 6 mainly support the existence of a bank-lending channel, especially for the important category of producers' durable equipment. Both the mix and spread variables have a significant effect on durable equipment investment for all three investment models, which suggests that the changes in bank credit conditions that seem to occur after monetary contractions do indeed appear to affect business investment.

KSW also run tests investigating the causal effects of their mix and spread variables on other monthly indicators of macroeconomic activity such as industrial production, personal income, employment, and unemployment. Although the results are not universally statistically significant for each of the sub-sample periods they examine, there is substantial evidence (shown in their Table 3) that the bank-credit variables have significant effects on many macroeconomic indicators.

6.4 Difference between Large and Small Firms

KSW's evidence focused on the credit market as a whole. However, a major theme of the credit-channel literature is that some firms are more credit-constrained than others. Access to direct capital markets such as commercial paper is limited to large firms with established reputations. Other firms are dependent on bank loans, whose availability may also depend on firm characteristics such as availability of collateral assets and general financial health.

A large literature has emerged since 1990 investigating the effects of financing constraints on different classes of firms distinguished by size, dividend policy, or other aspects of their financial condition. Many of these have focused on the most easily measured characteristic of firms: size. Size is not a perfect indicator of credit-market access, but there is considerable evidence that small firms are more often credit constrained than their larger counterparts. Table 7, taken from Table 1 of Gertler and Hubbard (1988), shows that small firms tend to retain a larger share of their earnings, suggesting less access to external financing in general. Small firms also tend to obtain a much larger share of their long-term funding from banks, which is consistent with the hypothesis that they lack access to direct credit markets.[20] Gertler and Gilchrist (1994) report similar bank-dependence of small firms for short-term lending.

Table 7. Sources of funds for manufacturing firms, 1970–84.

Net plant size of firm	Share of earnings retained	Long-term debt from banks
< $10 million	79%	67.3%
$10 – $50 million	76%	71.6%
$50 – $100 million	68%	71.0%
$100 – $250 million	63%	52.4%
$250 million – $1 billion	56%	40.8%
> $1 billion	52%	14.7%

Because small firms tend to lack access to credit markets, they depend on internal finance through retained earnings and on bank loans for most of their investment and working capital. Internal funds are squeezed during economic downturns as revenues slide relative to costs. Access to bank lending may be impaired directly by the impact of a monetary contraction on bank lending and indirectly through the deterioration of firms' indicators of financial creditworthiness, such as profits, cash flow, and net worth. Thus, if the credit channel plays an important role in the transmission of monetary shocks, we should see the investment of small firms being affected more strongly than

[20] The standard data set for examination of firms' financial activities is the *Quarterly Financial Report of Manufacturing, Mining, and Trade Corporations* (QFR) collected and published by the Bureau of the Census.

larger ones. In particular, during periods of tight money all firms continue to need credit (and some may desire additional credit). As bank loans dry up, large firms issue commercial paper and bonds, while small firms may be left without means to raise capital.

Gertler and Gilchrist (1994) examine the behavior of sales, inventories, and short-term debt of small and large firms after monetary contractions measured by Romer dates. Figure 7 (taken from Gilchrist and Gertler's Figure IV) shows the estimated effects of a monetary contraction for 16 ensuing quarters. Sales decline for both kinds of firms, though more strongly for the small firms. The remarkable differences lie in the behavior of inventories and short-term debt.

Given a (presumably) temporary decline in sales, firms would like to accommodate this decline by allowing inventories to build up and financing the reduction in sales revenue with credit. This seems to be exactly what happens to large firms, as both inventories and short-term debt initially rise after a monetary contraction. However, the behavior of small firms is very different. Small firms reduce inventory dramatically in the face of the sales decline. Their short-term debt also declines strongly despite a presumed increase in the desire to borrow. This evidence of an asymmetric impact of monetary contractions on large and small firms is strongly consistent with the credit-channel theory of monetary transmission.

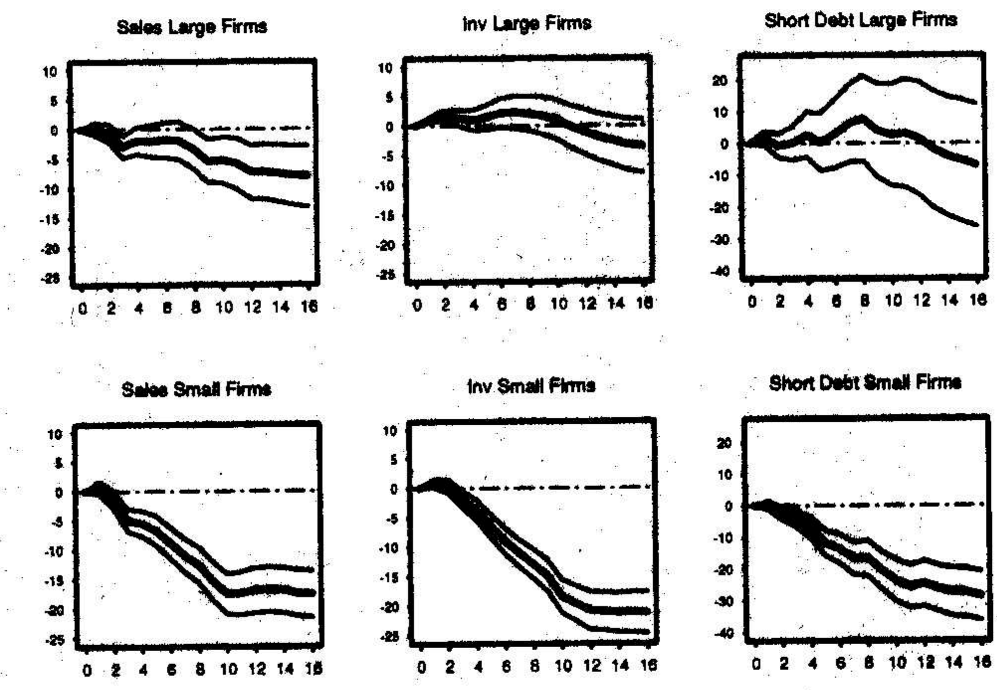

Figure 7. Gertler and Gilchrist's estimated effects of monetary contraction on small and large firms' sales, inventories, and short-term debt.

7 Summary and Conclusion

This chapter has examined evidence on the effects of monetary policy using a wide variety of empirical methods. The vast majority of the studies examined here suggest that monetary policy, measured either as changes in a monetary aggregate or as changes in the federal funds interest rate, have substantial short-run effects on real variables such as output and unemployment.

This result was clearly enunciated by Friedman and Schwartz (1963) and by the Keynesian macroeconometric modelers of the 1960s. The neoclassical theories of the rational-expectations school held that only unanticipated changes in monetary policy should have real effects. Evidence on this hypothesis and leans toward rejection of the most extreme version of the theory. But the results of Cochrane (1998) suggest that the distinction between anticipated and unanticipated monetary policy actions may be useful.

Most recent analysis has used vector autoregressions to estimate monetary effects. Although the VAR methodology was introduced to avoid the extreme identification assumptions implicit in early macroeconometric models, some identifying assumptions must still be made in order to analyze monetary effects using impulse responses. Many models have been estimated using a variety of assumptions both about identification and about the proper choice of monetary instrument. Despite the wide range of assumptions, the estimated effects of monetary shocks on output and prices are relatively homogeneous. Output responds most strongly six to eighteen months after a monetary shock, with price effects following later.

Finally, some economists have argued that monetary policy affects the economy through its effects on banks and credit conditions more than through the money supply and interest rates directly. Some supportive evidence suggests that credit-constrained firms are unusually strongly affected by monetary contractions.

References Cited

Barro, Robert J. 1977. Unanticipated Money Growth and Unemployment in the United States. *American Economic Review* 67 (1):101-115.

———. 1978. Unanticipated Money, Output, and the Price Level in the United States. *Journal of Political Economy* 86 (4):549-580.

Barth, James R., and James T. Bennett. 1974. The Role of Money in the Canadian Economy: An Empirical Test. *Canadian Journal of Economics* 7 (2):305-311.

Bernanke, Ben S. 1983. Nonmonetary Effects of the Financial Crisis in Propagation of the Great Depression. *American Economic Review* 73 (3):257-276.

Bernanke, Ben S., and Alan S. Blinder. 1992. The Federal Funds Rate and the Channels of Monetary Transmission. *American Economic Review* 82 (4):901-21.

Bernanke, Ben S., and Mark Gertler. 1995. Inside the Black Box: The Credit Channel of Monetary Policy Transmission. *Journal of Economic Perspectives* 9 (4):27-48.

Christiano, Lawrence J., and Martin Eichenbaum. 1992. Identification and the Liquidity Effect of a Monetary Policy Shock. In *Political Economy, Growth, and Business Cycles*, edited by A. Cukierman, Z. Hercowitz and L. Leiderman. Cambridge, Mass.: MIT Press.

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans. 1999. Monetary Policy Shocks: What Have We Learned and to What End? In *Handbook of Macroeconomics*, edited by J. B. Taylor and M. Woodford. Amsterdam: Elsevier Science.

Cochrane, John H. 1998. What Do the VARs Mean? Measuring the Output Effects of Monetary Policy. *Journal of Monetary Economics* 41 (2):277-300.

Cuddington, John T. 1981. Money, Income, and Causality in the United-Kingdom: An Empirical Reexamination. *Journal of Money, Credit and Banking* 13 (3):342-351.

Duggal, Vijaya G., Lawrence R. Klein, and Michael D. McCarthy. 1974. The Wharton Model Mark III: A Modern IS/LM Construct. In *Econometric Model Performance: Comparative Simulation Studies of the U.S. Economy*, edited by L. R. Klein and E. Burmeister. Philadelphia: University of Pennsylvania Press.

Eckstein, Otto, Edward W. Green, and Allen Sinai. 1974. The Data Resources Model: Uses, Structure, and Analysis of the U.S. Economy. In *Econometric Model Performance: Comparative Simulation Studies of the U.S. Economy*, edited by L. R. Klein and E. Burmeister. Philadelphia: University of Pennsylvania Press.

Eichenbaum, Martin. 1992. Interpreting the Macroeconomic Time Series Facts: The Effects of Monetary Policy: Comments. *European Economic Review* 36 (5):1001-11.

Eichenbaum, Martin, and Kenneth J. Singleton. 1986. Do Equilibrium Real Business Cycles Theories Explain Postwar U. S. Business Cycles? *NBER Macroeconomics Annual* 1:91-134.

Feige, Edgar L., and Douglas K. Pearce. 1979. The Causal Relationship Between Money and Income: Some Caveats for Time Series Analysis. *Review of Economcs and Statistics* 61 (4):521-533.

Friedman, Milton. 1968. The Role of Monetary Policy. *American Economic Review* 58 (1):1-17.

———. 1970. Comment on Tobin. *Quarterly Journal of Economics* 84 (2):318-327.

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

Gertler, Mark, and Simon Gilchrist. 1994. Monetary Policy, Business Cycles, and the Behavior of Small Manufacturing Firms. *Quarterly Journal of Economics* 109 (2):309-340.

Gertler, Mark, and Robert Glenn Hubbard. 1988. Financial Factors in Business Fluctuations. In *Financial Market Volatility*. Kansas City: Federal Reserve Bank of Kansas City.

Goodhart, C. A. E. 1989. *Money, Information and Uncertainty*. 2nd ed. Cambridge, Mass.: MIT Press.

Granger, C. W. J. 1969. Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. *Econometrica* 37 (3):424-438.

Hickman, Bert G., Robert M. Coen, and Michael D. Hurd. 1974. The Hickman-Coen Annual Growth Model: Structural Characteristics and Policy Responses. In *Econometric Model Performance: Comparative Simulation Studies of the U.S. Economy*, edited by L. R. Klein and E. Burmeister. Philadelphia: University of Pennsylvania Press.

Hosoya, Yuzo. 1977. On the Granger Condition for Non-Causality. *Econometrica* 45 (7):1735-1736.

Jaffee, Dwight M., and Thomas Russell. 1976. Imperfect Information, Uncertainty, and Credit Rationing. *Quarterly Journal of Economics* 90 (4):651-66.

Kashyap, Anil K., Jeremy C. Stein, and David W. Wilcox. 1993. Monetary Policy and Credit Conditions: Evidence from the Composition of External Finance. *American Economic Review* 83 (1):78-98.

King, Stephen R. 1986. Monetary Transmission: Through Bank Loans or Bank Liabilities? *Journal of Money, Credit and Banking* 18 (3):290-303.

Klein, Lawrence R., and Edwin Burmeister, eds. 1974. *Econometric Model Performance: Comparative Simulation Studies of the U.S. Economy*. Philadelphia: University of Pennsylvania Press.

Leeper, Eric M., and David B. Gordon. 1992. In Search of the Liquidity Effect. *Journal of Monetary Economics* 29 (3):341-369.

Leeper, Eric M., Christopher A. Sims, and Tao Zha. 1996. What Does Monetary Policy Do? *Brookings Papers on Economic Activity* (2):1-63.

Lucas, Robert E., Jr. 1972. Expectations and the Neutrality of Money. *Journal of Economic Theory* 4 (2):103-124.

———. 1975. An Equilibrium Model of the Business Cycle. *Journal of Political Economy* 83 (6):1113-44.

Meulendyke, Ann Marie. 1988. A Review of Federal Reserve Policy Targets and Operating Guides in Recent Decades. *Federal Reserve Bank of New York Quarterly Review* 13 (3):6-17.

Mishkin, Frederic S. 1982. Does Anticipated Monetary Policy Matter? An Econometric Investigation. *Journal of Political Economy* 90 (1):22-51.

Modigliani, Franco, and Merton H. Miller. 1958. The Cost of Capital, Corporation Finance and the Theory of Investment. *American Economic Review* 48 (3):261-297.

Muth, John F. 1961. Rational Expectations and the Theory of Price Movements. *Econometrica* 29 (3):315-335.

Phelps, Edmund S. 1968. Money-Wage Dynamics and Labor-Market Equilibrium. *Journal of Political Economy* 76 (4, Part 2):678-711.

Reichenstein, W. 1987. The Impact of Money on Short-Term Interest Rates. *Economic Inquiry* 25 (1):67-82.
Romer, Christina D., and David H. Romer. 1989. Does Monetary Policy Matter? A New Test in the Spirit of Friedman and Schwartz. In *NBER Macroeconomics Annual 1989*, edited by O. J. Blanchard and S. Fischer. Cambridge, Mass.: MIT Press.
———. 1990. New Evidence on the Monetary Transmission Mechanism. *Brookings Papers on Economic Activity* 1990 (1):149-198.
Rudebusch, Glenn D. 1998. Do Measures of Monetary Policy in a VAR Make Sense? *International Economic Review* 39 (4):907-31.
Runkle, David E. 1987. Vector Autoregressions and Reality. *Journal of Business and Economic Statistics* 5:437-442.
Sargent, Thomas J. 1976. A Classical Macroeconomic Model for the United States. *Journal of Political Economy* 84 (2):207-238.
———. 1976. The Observational Equivalence of Natural and Unnatural Rate Theories of Macroeconomics. *Journal of Political Economy* 84 (3):631-640.
Sargent, Thomas J., and Neil Wallace. 1975. 'Rational' Expectations, the Optimal Monetary Instrument, and the Optimal Money Supply Rule. *Journal of Political Economy* 83 (2):241-254.
———. 1976. Rational Expectations and the Theory of Economic Policy. *Journal of Monetary Economics* 2 (2):169-183.
Schink, George R. 1974. The Brookings Quarterly Model: As an Aid to Longer Term Economic Policy Analysis. In *Econometric Model Performance: Comparative Simulation Studies of the U.S. Economy*, edited by L. R. Klein and E. Burmeister. Philadelphia: University of Pennsylvania Press.
Sims, Christopher A. 1972. Money, Income, and Causality. *American Economic Review* 62 (4):540-552.
———. 1980. Comparison of Interwar and Postwar Business Cycles: Monetarism Reconsidered. *American Economic Review* 70 (2):250-257.
———. 1980. Macroeconomics and Reality. *Econometrica* 48 (1):1-48.
———. 1986. Are Forecasting Models Usable for Policy Analysis? *Federal Reserve Bank of Minneapolis Quarterly Review* 10 (1):2-16.
———. 1992. Interpreting the Macroeconomic Time Series Facts: The Effects of Monetary Policy. *European Economic Review* 36 (5):975-1000.
Small, David H. 1979. Unanticipated Money Growth and Unemployment in the United States: Comment. *American Economic Review* 69 (5):996-1003.
Stiglitz, Joseph E., and Andrew Weiss. 1981. Credit Rationing in Markets with Imperfect Information. *American Economic Review* 71 (3):393-410.
Stock, James H., and Mark W. Watson. 1989. Interpreting the Evidence on Money-Income Causality. *Journal of Econometrics* 40 (1):161-181.
Strongin, Steven. 1995. The Identification of Monetary Policy Disturbances: Explaining the Liquidity Puzzle. *Journal of Monetary Economics* 35 (3):463-97.
Tobin, James. 1970. Money and Income: Post Hoc Ergo Propter Hoc? *Quarterly Journal of Economics* 84 (2):301-317.