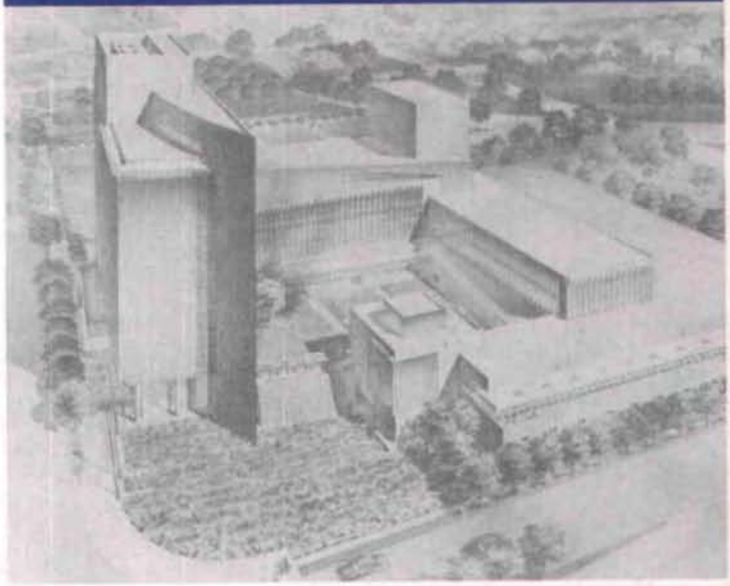
FEDERAL RESERVE BANK OF DALLAS
Fourth Quarter 1994

Economic Review



Would the Addition of Bond or Equity Funds Make M2 a Better Indicator of Nominal GDP?

John V. Duca

Understanding the Price Puzzle

Nathan S. Balke and Kenneth M. Emery

Indicators of the General Price Level and Inflation

Zsolt Becsi

Economic Review

Federal Reserve Bank of Dallas

The *Economic Review* is published by the Federal Reserve Bank of Dallas. The views expressed are those of the authors and do not necessarily reflect the positions of the Federal Reserve Bank of Dallas or the Federal Reserve System.

Subscriptions are available free of charge. Please send requests for single-copy and multiple-copy subscriptions, back issues, and address changes to the Public Affairs Department, Federal Reserve Bank of Dallas, P.O. Box 655906, Dallas, TX 75265-5906, (214) 922-5257.

On the cover: an architectural rendering of the Federal Reserve Bank of Dallas.

Contents

John Duca assesses the possibility that adding bond mutual funds, equity mutual funds, or both to M2 would improve this monetary aggregate's ability to forecast nominal GDP growth. He finds that M2B (M2 plus bond funds) and M2+ (M2 plus bond and stock funds) are statistically significant in explaining past nominal GDP growth. Duca further shows that M2B and M2+ each yield better forecasts of nominal GDP growth since 1990 than does M2, but to a lesser extent when the federal funds rate and the ten-year Treasury note yield are included in his forecasting model. Because bond and equity mutual funds are less directly influenced by the Federal Reserve than M2, Duca cautions that, relative to M2, M2B and M2+ are likely to be less controllable by the Federal Reserve.

Given these findings, Duca argues that M2B and M2+ show promise as information variables that the Federal Reserve may use along with other economic indicators in setting monetary policy. Recent forecast results and anecdotal information suggest that if equity funds continue to become more substitutable for nontransactions deposits, M2+ may prove to be increasingly helpful in this capacity.

Recent developments in measuring the stance of monetary policy have highlighted an interesting puzzle—namely, that an unexpected tightening in monetary policy leads to an increase rather than a decrease in the price level. In this article, Nathan Balke and Kenneth Emery present evidence on the price puzzle and discuss possible explanations for it.

Balke and Emery find that the most plausible explanation is that, during the 1960s and '70s, monetary policy was not implemented in a way that fully offset inflationary supply shocks. During this period, monetary policy would tighten in response to a supply shock but not by enough to prevent inflation from rising. In the data, therefore, contractionary policy is positively correlated with inflation. Since the early 1980s, however, the price puzzle has disappeared for either one, or both, of two reasons: the Federal Reserve has placed greater emphasis on achieving price stability, or there have been fewer inflationary supply shocks to the economy.

Contents

This article examines whether price indexes, such as the CPI, the PPI, and the implicit price deflator for GDP (PGDP), tell a consistent story about the general price level and inflation rate. To this end, Zsolt Becsi analyzes the time series properties of these indexes. He finds that the PGDP has a stable long-term relationship with both of the other price indexes. Some evidence suggests that PGDP and CPI inflation have common long-run trends, while PPI inflation has no discernible stable long-run relationship with either PGDP or CPI inflation.

Some theories suggest that the price level relevant for monetary policy is broader than price indexes of final goods and services such as the PGDP. This article investigates whether the PGDP captures movements in other price or inflation series. There is weak evidence that the PGDP shares common trends with the price levels and inflation rates of some intermediate goods and assets. Overall, these results suggest that PGDP makes a good indicator of the general price level for monetary policy because it reflects shocks to a broad range of other series.

John V. Duca
Research Officer
Federal Reserve Bank of Dallas

Would the Addition of Bond or Equity Funds Make M2 a Better Indicator of Nominal GDP?

For some time, the Federal Reserve has sought to keep inflation low to foster maximum sustainable growth.[1] Given the costs of reducing inflation, the Federal Reserve has, since the early 1980s, pursued a policy of preventing inflation from rising.[2] Because monetary policy affects the economy with a lag, implementing this forward-looking, low inflation strategy requires that the Federal Reserve accurately forecast and gauge price pressures.

One way to keep inflation low is to keep nominal gross domestic product (GDP) growing at a moderate pace that, at most, only slightly exceeds the long-run growth rate of inflation-adjusted output.[3] To keep nominal spending growth at such a pace, the Federal Reserve looks at economic indicators to track and forecast nominal GDP. One notable indicator is the monetary aggregate M2, whose relationship to nominal GDP may be breaking down in the 1990s, partly because households are shifting assets away from M2 deposits into bond and equity mutual funds. This article assesses whether M2 would be a better indicator of nominal GDP growth if it were expanded to include bond and, possibly, equity mutual fund assets.[4]

The use of money as an indicator of nominal spending can be justified by the equation of exchange:

$$(1) \qquad M \times V = P \times T = Y,$$

where M = money, V = velocity (GDP/M), T = inflation-adjusted transactions (measured by inflation-adjusted GDP), P = the price level, and Y = nominal GDP. Holding nominal GDP constant, people typically reduce their money holdings as the gap between the yield on nonmonetary assets (for example, U.S. Treasury securities) and deposit rates widens. Consequently, as this spread, or opportunity cost of holding money, increases, the velocity of money rises.

I would like to thank Jean Zhang and Chih-Ping Chang for excellent research assistance; John Benvenuto and the Investment Company Institute for providing data on mutual funds; and Nathan Balke, Evan Koenig, and especially Ken Emery for comments and suggestions. Any remaining errors are my own.

[1] *See Rudebusch and Wilcox (1994) and Wynne (1993). High inflation lowers long-term growth by increasing uncertainty. High uncertainty not only limits long-term contracting and investment but also reduces efficiency by hindering consumers' and firms' search for the lowest prices, which, in turn, hinders market forces from shifting resources to the lowest cost producers. Because the U.S. tax code does not index capital gains and depreciation for inflation, high inflation also lowers long-run growth through raising the real after-tax cost of capital, thereby reducing investment.*

[2] *Reducing inflation from high levels has often been accompanied by recessions as consumers and firms often need to experience economic slack before reducing their wage and price demands to levels in line with low inflation.*

[3] *For example, if long-run output growth is 2.5 percent under low inflation, then over the long run, 4.5-percent nominal GDP growth implies 2-percent inflation, using the implicit GDP price deflator.*

[4] *Brayton and Tinsley (1993) find that using the federal funds rate as an instrument to hit nominal GDP as an intermediate target outperformed trying to hit price level or money intermediate targets in terms of stabilizing the price level. This article assesses the relative performance of M2 variants, not as potential intermediate targets but as information variables that could be used to help forecast short-run nominal GDP growth.*

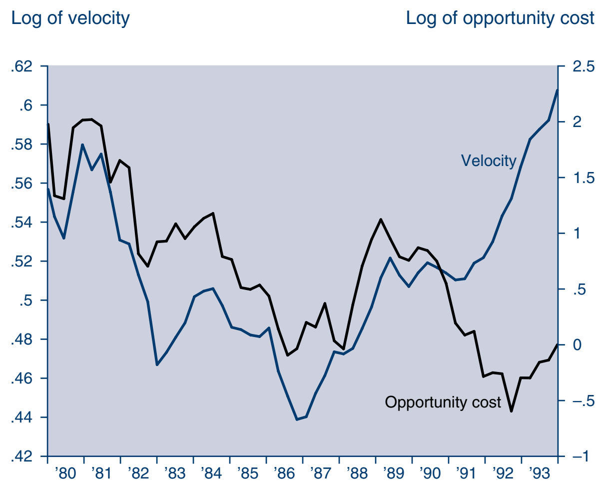
If interest rate variables can reliably predict velocity, then nominal GDP can be inferred from money and interest rates. This is an important implication for policy-making because estimates of nominal GDP are available after a considerable lag and are subject to sizable revisions, whereas good information on interest rates and monetary aggregates is available with very little lag.

When M1's velocity was predictable, M1 was used as an indicator of nominal GDP. However, this relationship began breaking down in the mid-1970s when unusually weak M1 underpredicted nominal GDP. Moreover, the link between M1 and nominal GDP became somewhat looser after the deregulation of deposits in the early 1980s, which made the demand for M1 very interest rate sensitive. Consequently, M1 has been used less and less as an indicator of nominal GDP.

Up through 1990, evidence had mounted that the demand for M2 was more predictable than the demand for M1 (see Hetzel and Mehra 1989 and Moore, Porter, and Small 1990). Partly as a result, M2 became a more popular indicator of nominal GDP and of inflation (see Hallman, Porter, and Small 1991). However, since the early 1990s, M2 growth has been unusually weak and has been underpredicting nominal GDP growth. As shown in Figure 1, this breakdown occurred in the early 1990s when M2's velocity began diverging sharply from a conventional measure of its opportunity cost. This unusual weakness is confirmed in econometric models of M2, as documented by Anderson and Collins (1994), Duca (forthcoming), and Feinman and Porter (1992).

One common explanation for this estimated shortfall, or "missing M2," is that households shifted their assets from M2 deposits into bond and, possibly, stock mutual funds (see Anderson and Collins 1994 and Duca, forthcoming). If such portfolio shifts are too difficult to accurately model, then one option is to redefine M2 to include bond and, possibly, equity funds. Indeed, something similar happened in the early 1980s, when M2 was redefined to include money market mutual funds, or MMMFs (Simpson 1980 and Duca 1993a, 4). More recently, Duca (forthcoming) has found that M2 is less explainable in money models compared with an M2 aggregate that is redefined to include bond funds. In addition, Becsi and Duca (forthcoming) and Duca

Figure 1
M2 Velocity and Its Opportunity Cost



SOURCE: Board of Governors, Federal Reserve System.

(1994) have found that expanded M2 aggregates that include either bond or bond and equity funds easily outperformed M2 in forecasting inflation in recent years using the P-star inflation model of Hallman, Porter, and Small (1991). The current study extends this research by assessing the ability of such expanded aggregates to forecast nominal GDP growth relative to that of M2.

This article is organized as follows. The second section intuitively reviews what occurs when money demand relationships break down and discusses why households may substitute bond and equity mutual funds for M2 deposits. The next two sections assess the relative ability of different versions of M2 to explain future nominal GDP growth. The following section assesses the stability of such models. The article concludes by discussing the policy implications of the findings.

Why money demand breaks down and the recent role of mutual funds

The recent breakdown in the relationship between M2's velocity and conventional measures of its opportunity cost likely reflects that these measures have not tracked the decline in the attractiveness of M2 deposits relative to other financial assets. Possible explanations for this include that other asset yields have become more important, that government regulations have made M2 less attractive, and that the private sector has made bond and equity funds more attractive (Duca 1993b, forthcoming).

In the past, unusual weakness in money growth has been associated with declines in bank competitiveness (see Duca 1993b). One relevant example is the "missing M1" of the mid-1970s, when the interaction of high interest rates and regulations impaired the ability of banks to offer deposit and credit services to firms. In response, many firms substituted repurchase agreements and cash management for non-interest-bearing demand deposits. In addition, many large firms shifted away from bank loans toward commercial paper. This shift reduced compensating demand deposit balances that were held in proportion to firms' bank loans. At the same time, households shifted out of deposits into money market mutual funds, which paid interest rates above the deposit rate ceilings at banks. By expanding the number of households that could directly or indirectly invest in commercial paper, money funds made commercial paper cheaper than bank loans for low risk firms and opened a new channel through which short-term credit could flow from households to firms.

In response to these episodes, the Federal Reserve redefined M2 in 1980 to internalize shifts between bank and non-bank-like deposits so as to create a better economic indicator. Over time, M2 has evolved to include new instruments, most notably, money funds and their bank counterpart, money market deposit accounts (MMDAs).[5] Because of redefinitions, much of M2's apparent value as an indicator before the early 1980s is misleading. In recent years, bond and equity mutual funds have grown rapidly at the expense of money funds and small time deposits, both of which are components of M2.

Bond and equity funds. Bond and equity funds are substitutable for M2 deposits and for direct bond and equity investments. Because they are mutual fund shares, they offer investors lower risk compared with direct holdings of securities because the funds are diversified and professionally managed. Many funds are also in asset management accounts that provide liquidity by giving investors credit lines and by allowing investors to shift assets among equity, bond, and checkable money funds at little or no cost. Bond funds are good substitutes for M2 for two other reasons. First, because most bond fund assets are invested in U.S. government and other high-grade bonds, they generally have low credit risk. Second, bond funds typically offer higher expected returns than M2, owing to the longer maturity of assets that bond funds hold. However, this longer maturity creates a price risk for investors because bond prices fluctuate. Compared with bond funds, equity funds offer higher expected returns and higher risk, which may make them less substitutable for M2 deposits. Thus, it is unclear, a priori, whether M2 plus bond and equity funds (M2+) is a better indicator of nominal GDP than M2 plus bond funds (M2B).

How the recent missing M2 period reflects a bypassing of banks. How can a bypassing of the banking system through bond and equity funds lead to an episode of missing money? Suppose a firm raises $100 by issuing bonds bought by a bond fund. The bond fund pays the firm with $100 from selling mutual fund shares to a household, which obtains the $100 by withdrawing $100 from a small time deposit. Using the $100 raised from issuing a bond, the firm pays down $100 in bank loans. Note that any rise in checking accounts used to make any of these transfers is temporary because the rise in checking accounts runs off after the transfers are completed.

On the firm's balance sheet, total liabilities are unchanged as the $100 decline in loans matches the $100 rise in bonds. Total household assets are also unchanged because the $100 decline in small time deposits matches the $100 rise in bond funds. The bond fund, however, sees a $100 increase in both assets and liabilities, while banks see a $100 decline in loans and deposits. Thus, M2 falls by $100, while the sum of bond funds and M2 (M2B) is unchanged.

In recent years, many firms have shifted from bank loans toward bonds and equity for finance, partly because the spread of the prime rate over

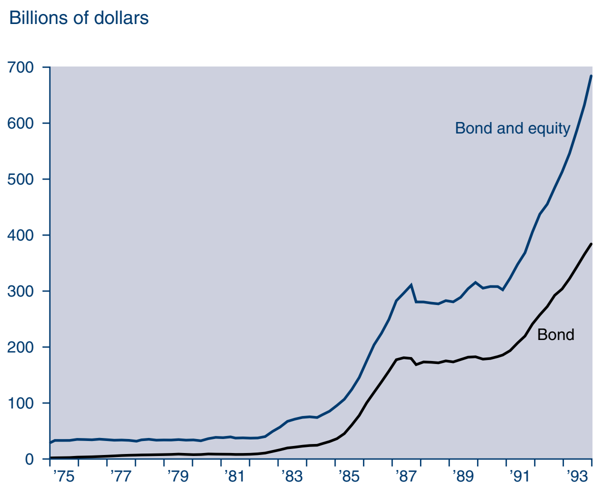
[5] *M2 includes currency, demand deposits, savings deposits (passbook savings plus MMDAs), noninstitutional MMMFs, small time deposits, overnight repurchase agreements, and overnight Eurodollar deposits.*

short-term rates has risen as banks passed on the higher cost of the new risk-based capital standards. At the same time, households have shifted out of M2 to bond and equity funds. Essentially, bond and equity funds provide another channel through which long-term finance can flow from households to firms.

Bond and equity fund growth. Adding either bond funds or both bond and equity funds to M2 may help restore M2 as an economic indicator by internalizing shifts between bank deposits and bond and equity fund assets. Figure 2 plots bond funds and bond plus equity funds held by households. As with M2, both series exclude Individual Retirement Account (IRA) and Keogh assets along with institutional holdings. (For details on M2+ and M2B, see Collins and Edwards 1994 and Duca, forthcoming, respectively.)

In the mid-1980s, households flocked to bond and equity funds as the eligibility restrictions on IRAs and Keogh accounts were loosened. As more households learned about these funds when opening IRAs, many shifted assets into non-IRA/Keogh fund accounts as well. Balance sheet data suggest that more of these fund inflows came from direct holdings of bonds and equities than from M2. After 1986 tax reform tightened IRA and Keogh rules, bond and equity funds were about flat during the late 1980s. More recently, these funds have surged, this time more at the expense of M2 than of directly held securities. Excluding IRA and Keogh assets, adding bond funds or both bond and equity funds to M2 produces an adjusted M2 that has grown faster than M2 in recent years.

Figure 2
Household Bond and Equity Mutual Fund Assets



SOURCE: Investment Company Institute.

Empirical results using lags of only money and nominal GDP

This section simply analyzes the ability of M2, M2B, and M2+ to indicate near-term nominal GDP growth. After the basic empirical model is presented, results from regressions and from out-of-sample forecasts are discussed.

Basic empirical model. Nominal GDP growth (y) is estimated from regressions using four lags of itself, four lags of money growth (m), and the one-quarter lag of the long-run relationship between the logs of nominal GDP and money (EC):

$$(2) \qquad y_t = \beta_o + \sum_{i=1}^{4} \beta_i y_{t-i} + \sum_{i=1}^{4} \gamma_i m_{t-i} + \alpha EC_{t-1},$$

where β_o is a constant, β_i is the estimated effect of nominal GDP growth in quarter $t-i$, γ_i is the estimated effect of money growth in quarter $t-i$, and α reflects the impact of deviations of nominal GDP from its long-run equilibrium relationship to the level of money holdings.[6]

Essentially, the error-correction term accounts for information relating the log levels of output and money and in doing so, prevents the model from letting nominal output levels drift too far away from the level of money (see Hafer and Kutan 1992 for a related discussion). For estimating equation 2, the EC term is based on the equation of exchange (equation 1) and the assumption that the long-run velocities of M2, M2B, and M2+ are stable throughout the sample period. In particular, the average velocity of these aggregates are substituted into the equation of exchange to obtain

$$(3) \qquad EC = \log(\mathit{nominal\ GDP}) - \log(\mathit{money}) - \log(\mathit{average\ velocity}).$$

[6] *The models used in Orphanides, Reid, and Small (1994) to compare M2 and M2+ do not include error-correction terms as discussed by Duca (1994). Models tested by Feldstein and Stock (1994) also suffer from this criticism.*

EC can be thought of as the gap between nominal spending and its equilibrium level as implied by money balances. Thus, for example, a positive value of *EC* implies that nominal GDP growth will decline to restore equilibrium, all else being equal. For this reason, *EC* is expected to have a negative sign.

By contrast, the sum of coefficients on lags of money growth should be positive, as implied by the equation of exchange (equation 1). In theory, the sum of coefficients on lagged nominal GDP growth could be positive or negative. However, in practice, movements in nominal GDP growth tend to persist for some time, reflecting swings in real growth and in inflation.[7]

Regression results. Using data on the levels of M2, M2B, and M2+ that go back to first-quarter 1959, equation 2 is estimated over the in-sample period 1960:2–94:1. For each run, one of the three definitions of M2 is used in defining the error-correction (*EC*) term and lagged money growth variables.

As shown in models 1 through 3 of Table 1, the fit (corrected R^2) of equation 2 is highest for M2, somewhat lower for M2B, and lower yet for M2+. For each aggregate, lags of money growth are jointly significant according to F-statistics. T-statistics indicate that the error-correction terms for M2 and M2B are marginally significant, while that for M2+ is insignificant. Together, all terms involving money (the *EC* and lagged money growth terms) are jointly significant for each aggregate, and, as expected, the sum of coefficients on lagged money growth is positive and the error-correction term is negative. This pattern is also obtained when the error-correction terms are based on estimated cointegrating vectors using the Johansen and Juselius (1990) procedure (models 4 through 6). These findings indicate that each M2 aggregate helps explain future movements in nominal GDP growth over the full sample period.

To help control for short-run velocity movements induced by changes in relative rates of return, models 7 through 12 add four lags of opportunity cost measures to equation 2. To control for substitution with short-term investments, one type of opportunity cost term (*SOC*) is based on the log of the spread between the three-month Treasury bill rate and the average return on money. To account for shifts with longer term investments, a second type of opportunity cost term (*LOC*) is based on the log of the spread between the ten-year Treasury note yield and the average return on money. Federal Reserve Board data on M2 average rates of return are used for M2 and in constructing weighted average rates of return for M2B and M2+. The weighted average rates of return for M2B and M2+ assume that the return on bond funds is approximated by the ten-year Treasury yield and the return on stock funds, by the annualized percentage change in the S&P 500 index of stock prices.[8] The long-term opportunity cost terms are jointly significant, while the short-term cost terms are not. The positive sign on the sum of the *LOC* coefficients could reflect that the velocity of the M2 aggregates rises with *LOC*.[9]

The qualitative results for models 7 through 12 differ slightly from those of models 1 through 6 on two counts. First, the relative R^2s of the M2+ models improve greatly, and the error-correction (*EC*) terms are significant in the M2+ models, perhaps reflecting that the opportunity cost terms partly control for the impact of capital gains and

[7] *The persistence of nominal GDP movements depends partly on monetary policy. As an extreme example, if the Federal Reserve removed all but the most temporary movements in nominal GDP growth around a constant moderate growth rate, then changes in nominal GDP growth would unwind in one quarter, and the one-quarter lag of nominal GDP growth would be negatively correlated with the current growth rate. In practice, temporary shocks to nominal GDP growth often last longer than one quarter, and there have been some persistent swings that make this correlation positive. Nevertheless, the example implies that the Federal Reserve's shift toward stabilizing nominal GDP growth or inflation will make the sum of coefficients on lags of nominal GDP growth less positive. Indeed, this sum declines in size as the sample is extended from 1983 onward, which is consistent with Emery (1994), who finds that changes in the inflation rate became less persistent after the early 1980s; this may reflect that Federal Reserve efforts to keep inflation low after 1983 made any deviation of inflation from this modest pace rather short-lived.*

[8] *While not an ex ante rate, the stock price change may be a reasonable proxy. Whenever the level of a spread was less than 0.5 percent, a Taylor-log approximation of the log of the spread was used.*

[9] *Alternatively,* LOC *is positively correlated with the gap between long-term and short-term interest rates, which is positively related to nominal GDP growth. However, this latter correlation is statistically insignificant.*

Table 1

Predicting Nominal GDP With Money and Opportunity Costs[1] (1960:2–94:1)

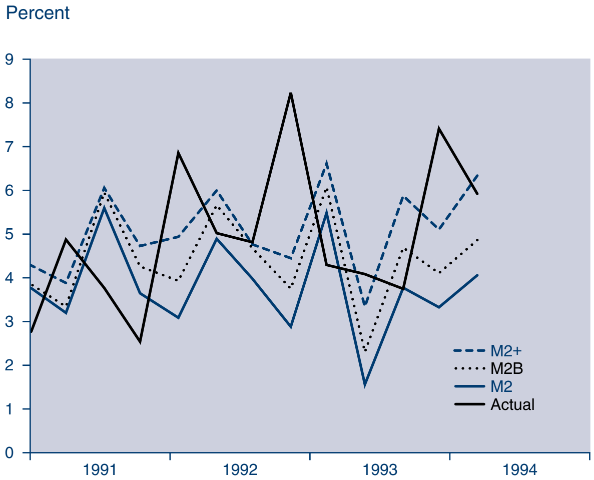
Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12
GDP	.3616	.4142^{+}	.4642*	.3559	.4142^{+}	.4577*	.0908	.0915	.1105	.0753	.0915	.1019
	(1.71)	(2.22)	(2.70)	(1.77)	(2.22)	(2.64)	(.24	(.25)	(.28)	(.23)	(.25)	(.27)
EC_{t-1}	–.0546^{+}	–.0442^{+}	–.0322				–.0727*	–.0676*	–.0578*			
	(–1.87)	(–1.81)	(–1.56)				(–2.22)	(–2.56)	(–2.37)			
EEC_{t-1}				–.0601*	–.0442^{+}	–.0293				–.0672*	–.0676*	–.0531*
				(–2.23)	(–1.81)	(–1.46)				(–2.41)	(–2.56)	(–2.40)
M2	.3129*			.3368**			.3424**			.4305**		
	(2.83)			(3.56)			(3.22)			(4.48)		
M2B		.4042**			.4042*			.4377**			.4377**	
		(4.49)			(4.49)			(4.76)			(4.76)	
M2+			.4012**			.4072**			.4398**			.4673**
			(4.51)			(4.59)			(4.99)			(5.31)
SOC							–.0120	–.0109	–.0111	–.0100	–.0109	–.0105
							(1.76)	(1.12)	(1.26)	(1.29)	(1.12)	(1.14)
LOC							.0168*	.0143*	.0171**	.0112^{+}	.0143*	.0148**
							(2.52)	(2.76)	(3.05)	(2.06)	(2.76)	(2.98)
F: *M* & *EC*[2]	5.18**	4.83**	4.35**	5.52**	4.83**	4.28**	5.89**	5.10**	4.91**	6.11**	5.10**	4.98**
R^2	.2224	.2135	.2005	.2312	.2135	.1987	.2569	.2590	.2552	.2624	.2590	.2559

[1] Sums of coefficients are provided for lags of GDP, money, and interest rate variables with F-test statistics for the joint exclusion set of lags given in parentheses. Underneath the coefficient on each *EC* term is its t-statistic in parentheses. Estimated constants are omitted to conserve space.

[2] F-statistic on the error-correction term (*EC*) and lagged money growth.

* (**,$^{+}$) denotes significance at the 95-percent (99-percent, 90-percent) confidence level.

Figure 3
Nominal GDP Growth Forecasts Using Money



losses on stock fund assets. Second, the M2B model has a negligibly higher R^2 than that of the M2 model when quantity theory based-EC terms are used (models 7 through 9), whereas M2 has an edge when estimated EC terms are used instead (models 10 through 12).

Forecasts. In-sample results overstate the ability of M2 to forecast nominal GDP growth in the 1990s. This can be shown by forecasting nominal GDP growth starting in first-quarter 1991 using coefficients from models 1 through 3 estimated over 1960:2–90:4 and actual values of all right-hand side variables since then. As shown in Figure 3, M2 underforecasts nominal GDP growth, while M2B and M2+ perform well.[10] Forecasts using M2B and M2+ yielded average errors of –.22 and +.25 percentage points at an annual rate, respectively, compared with –2.64 percent for M2 (*Table 2*). In addition, the sums of squared errors are 74 and 78 percent lower for M2B and M2+, respectively, than for M2. Although not statistically significant, these differences are economically meaningful and suggest that M2 has recently been distorted by portfolio shifts into bond and equity funds that are implicitly taken into account by M2B and M2+.

Regression results using lags of money, interest rates, and nominal GDP

A number of researchers have investigated interest rate variables as alternative indicators of economic activity that contain information beyond that in monetary aggregates (for example, Friedman and Kuttner 1992).[11] Motivated by this research, this section addresses the issue of how the three M2-type aggregates perform in the presence of interest rate indicators.

Empirical model. Several sets of regressions and simulations are run based on adding lags of interest rate variables (x) to equation 2:

$$(4) \qquad y_t = \beta_o + \sum_{i=1}^{4} \beta_i y_{t-i} + \sum_{i=1}^{4} \gamma_i m_{t-i} + \sum_{i=1}^{4} \delta_i x_{t-i} + \alpha EC_{t-1},$$

where δ_i denotes the coefficient reflecting the effect of the ith lag of x.

For each definition of money, four interest rate variables are assessed using equation 4: (1) the federal funds rate (*FF*), (2) the constant maturity yield on ten-year Treasury notes (*10YRT*), (3) the spread between the yield on ten-year Treasury notes and the federal funds rate (*YCURVE*), and (4) the spread between the six-month prime commercial paper rate and the six-month Treasury bill rate (*PAPERBILL*).[12] In addition to running a set of three money regressions for each interest rate variable, an extra set of regressions is run that

[10] *Forecasts (not shown) were also done using models 7 through 9. Once again, M2B and M2+ yielded smaller sums of squared errors than M2, while, on average, M2 tended to underpredict nominal GDP growth to a greater degree than M2B or M2+. However, each of these models yielded worse forecasts than corresponding models without lags of* SOC *and* LOC*. Models adding lags of* SOC *but not of* LOC *yielded forecasts that were similar to those in Figure 3.*

[11] *Another motivation for including interest rates is that they may help control for movements in velocity because they are correlated with M2 opportunity cost movements. However, because deposit deregulation has altered the correlation of opportunity cost and interest rate variables, controlling for velocity movements is better handled by adding terms like* SOC *and* LOC.

[12] *For background on these variables, see Bernanke and Blinder (1992) on the federal funds rate, Stock and Watson (1989) on the yield curve spread, and Bernanke (1990) and Friedman and Kuttner (1992) on the paper-bill spread.*

Table 2

Nominal GDP Forecast Results (Forecasts over 1991:1–94:1 based on a 1960:2–90:4 Insample Period)

Lagged Money and Nominal GDP with an Error-Correction Term[1]

Aggregate	Average annualized error[2] (Percent)	S.S.E.
M2	–2.64	.00159
M2B	–.22	.00042
M2+	+.25	.00035

Lagged Money, Interest Rates, and Nominal GDP with an Error-Correction Term[3]

Aggregate	Average annualized error (Percent)	S.S.E.
M2	–1.16	.00079
M2B	–.53	.00060
M2+	+.16	.00048

[1] The error-correction term is based on a constant velocity and the quantity theory of money.

[2] Annualized, average percentage point error. Negative entries denote underpredictions of nominal GDP growth.

[3] The error-correction terms are based on cointegrating relationships estimated over the 1960:2–90:4 in-sample period.

includes the federal funds rate and the ten-year Treasury yield.

Because *PAPERBILL* and *YCURVE* have no trends, their lags can be added to equation 4 without altering the EC term. However, because the ten-year Treasury yield and the federal funds rate do have trends,[13] it is not valid to simply add lags of these last two interest rate variables without changing the EC term and without first transforming the lagged rate terms into changes. To handle this problem, four lags of the ten-year Treasury yield and the federal funds rate are used ($\Delta 10YRT$ and ΔFF), and the EC term is redefined to control for the long-run relationship between nominal output, money, and interest rates.[14] The EC terms are from estimates of the long-run relationships for the following sets of variables for each definition of M2: (1) log of nominal GDP, log of money, and *FF*; (2) log of nominal GDP, log of money, and *10YRT*; and (3) log of nominal GDP, log of money, *FF*, and *10YRT*.[15]

Regression results. Results for models 13 through 24 (*Table 3*) indicate that the lags of the paper-bill spread (*PAPERBILL*), changes in the federal funds rate (ΔFF), and changes in the ten-year Treasury yield ($\Delta 10YRT$) are each jointly significant at the

[13] *Chi-squared statistics from Dickey–Fuller unit root tests rejected that the levels of the federal funds rate (17.7—trend—and 17.8—no trend—at four lags) and ten-year Treasury yield (15.2—trend—and 15.5—no trend) were stationary at the 4-, 6-, 9-, and 11-percent significance levels, respectively.*

[14] *Qualitative results were similar using the EC term from equation 3.*

[15] *Using the Johansen and Juselius (1990) procedure, the EC term was based on the estimated cointegrating vector for each combination of interest rates, money, and nominal GDP that had the highest degree of significance according to test results on the rank of the cointegration space. For each combination, only one cointegrating relationship had a significance level of 5 percent.*

Table 3

Predicting Nominal GDP With Money and Interest Rates[1] (1960:2–94:1)

Variable	Model 13	Model 14	Model 15	Model 16	Model 17	Model 18	Model 19	Model 20	Model 21	Model 22	Model 23	Model 24
GDP	.4159	.4350^{+}	.5405**	.4037	.4489^{+}	.5612**	.0974	.0989	.4510*	.1718	.1922	.5991**
	(2.02)	(2.29)	(4.16)	(1.74)	(2.22)	(3.55)	(.56)	(.58)	(2.80)	(.53)	(.63)	(3.49)
EC_{t-1}	–.0627*	–.0463^{+}	–.0424*	–.0527	–.0505	–.0335	–.0360**	–.0155*	–.0010	–.0043^{+}	–.0045^{+}	–.116
	(–2.15)	(–1.87)	(–2.02)	(–1.73)	(–1.76)	(–1.43)	(–2.95)	(–2.57)	(–.41)	(–1.97)	(–1.97)	(–1.58)
M2	.2292*			.2893^{+}			.6539*			.6025**		
	(.81)			(2.09)			(6.74)			(3.64)		
M2B		.3649^{+}			.3819*			.7443**			.5887**	
		(2.06)			(3.44)			(6.26)			(3.53)	
M2+			–.00002			–.00002			–.000002			–.000002
			(1.33)			(1.07)			(.86)			(1.22)
ΔFF										.0005^{+}	.0004*	–.0036**
										(2.43)	(2.50)	(3.83)
$\Delta 10YRT$							.0089*	.0111*	–.0003^{+}	.0063^{+}	.0084*	.0047
							(2.76)	(3.09)	(2.05)	(2.07)	(2.47)	(1.55)
PAPERBILL	.0019*	.0034*	.0031**									
	(2.60)	(2.56)	(5.24)									
YCURVE				.0005	–.0004	.0009^{+}						
				(1.25)	(1.13)	(2.02)						
F: *M* & *EC*[2]	3.28**	2.96*	1.99^{+}	4.14**	3.74**	1.41	5.59**	5.01**	.70	2.92*	2.83*	1.01
R^2	.2537	.2450	.2169	.2217	.2105	.1387	.2628	.2482	.1182	.2776	.2750	.2210

[1] Sums of coefficients are provided for lags of GDP, money, and interest rate variables with F-test statistics for the joint exclusion set of lags given in parentheses. Underneath the coefficient on each *EC* term is its t-statistic in parentheses. Estimated constants are omitted to conserve space.

[2] F-statistic on the error-correction term (*EC*) and lagged money growth.

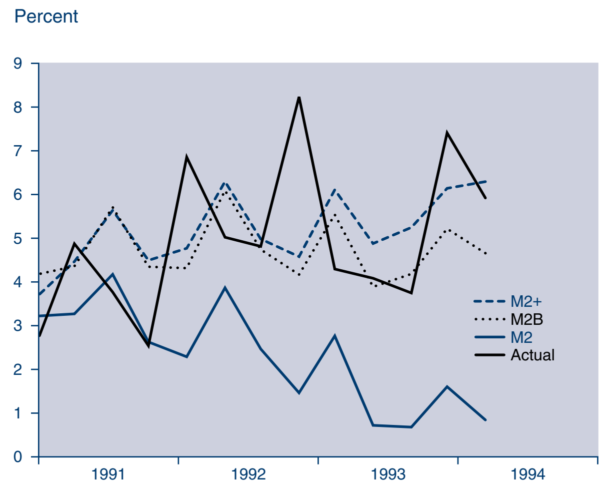
* (**,$^{+}$) denotes significance at the 95-percent (99-percent, 90-percent) confidence level.

10-percent level or better. Thus, these interest rate variables, unlike the yield curve (*YCURVE*), help forecast nominal GDP growth above and beyond the information content in the M2-type aggregates. The models including both the federal funds rate and the ten-year Treasury yield (models 22 through 24) have the best fit (corrected R^2s) among the models for each broad monetary aggregate.

The relative performance of M2, M2B, and M2+ is similar to earlier results, with M2 yielding a slightly higher R^2 than M2B and with M2+ performing the worst among the three. However, in models including both the federal funds rate and the ten-year Treasury yield, the full-sample R^2s of the M2B and M2 models (models 22 and 23) are closer. For all three aggregates, the sum of coefficients on lagged money growth is positive, lags of money growth are jointly significant, the error-correction term is insignificant but correctly signed, and the lags of money and the error-correction term are jointly significant in models 22 through 24. Overall, the results indicate that M2, M2B, and M2+ provide useful information in predicting movements in nominal GDP growth.

Forecast results. Based on in-sample fit, the models containing both the federal funds rate and ten-year Treasury yields are used to evaluate the relative forecast performance of the three M2-type aggregates. As before, the forecasts of nominal GDP growth start in first-quarter 1991 using coefficients from equation 2 estimated over 1960:2–90:4 and actual values of all right-hand side variables since then.[16] Figure 4 illustrates that M2 underpredicts nominal GDP growth, while M2B and M2+ yield somewhat better forecasts. As shown in Table 2, forecasts using M2B and M2+ have average errors of –.53 and +.16 percentage points at an annual rate, respectively, compared with –1.16 percentage points for M2. The sums of squared errors are .00060 and .00048 for M2B and M2+, respectively, or 24 percent and 39 percent lower than that for M2 (.00079), respectively. Together with Figure 3, these findings imply that M2 noticeably underpredicts nominal GDP growth when used alone but to a lesser extent than when used along with interest rates. One explanation for this pattern may stem from a tendency for interest rate models to overpredict nominal GDP in recent years, while M2 growth tends to underpredict it. As a result, these tendencies may be offset when both types of variables are included.

Figure 4
Nominal GDP Growth Forecasts Using Interest Rates and Money



Evidence on stability

This section assesses whether the three M2 aggregates have been relatively stable predictors of nominal GDP growth since the early 1980s.

Rolling regression joint exclusion tests. One way of assessing the stability of the forecasting models is to test whether all the money variables in them can be excluded using different sample periods. The rolling regression approach is used here, where the initial sample used is 1960:2–83:1 and each subsequent sample period adds one further observation. First-quarter 1983 is chosen as the starting point on the grounds that M2 was redefined for a second time in 1983 to include MMDAs and that the last change in monetary operating procedures occurred in late 1982.

Models 1 through 3 were chosen for these F-tests because their error-correction terms do not need to be reestimated for each sample, unlike

[16] *Forecasts using M2B and M2+ that omitted interest rates have better fits than those corresponding runs that include these two interest rates. This may reflect a change in the information content of interest rates over time.*

models 16 through 18, which would require 144 searches for a unique error-correction term. For each model, Figure 5 plots the F-statistic on the joint exclusion of the error-correction term and the four lags of money growth. For all three aggregates, these terms are always jointly significant. However, the joint significance statistics for M2B and M2+ decline in the mid-1980s while that for M2 declines in the early 1990s.

Chow tests. As a further check, Chow tests are run on models 1 through 3 over 1984:1–93:4 to test if the model residuals become unusually large. The F-statistics from these tests are plotted in Figure 6.[17] Chow tests cannot reject stability for any of the aggregates.

Dummy variable tests. Because Chow and joint exclusion tests do not necessarily rule out minor shifts in the relationship between M2 aggregates and nominal GDP, a series of dummy variable tests are run for the two periods of rapid growth in bond and equity funds: the mid-1980s and early 1990s.

A dummy variable equal to 1 after fourth-quarter 1991 (*DUM92*) is added to models 1 through 3 and 22 through 24 to test for a shift in the constant that persists after fourth-quarter 1991; this definition is consistent with Figure 3, which shows M2 underforecasting nominal GDP growth since early 1992. *DUM92* is significant only in the noninterest rate model using M2 (*Table 4*), with a positive sign on *DUM92* implying that the dummy variable helps offset the tendency of M2 to underpredict nominal output in the early 1990s. This finding is consistent with earlier forecasting results showing that M2 noticeably underpredicts nominal GDP growth when used alone but not when used along with interest rates.[18]

As discussed earlier, M2B and M2+ were likely distorted in the mid-1980s by inflows that reflected shifts away from direct holdings of securities (see Duca 1992). This may explain why models using M2 have slightly better full-sample fits than corresponding models that use M2B even though the models using M2B perform better in recent years.

Figure 6
Chow Tests for Residual Stability

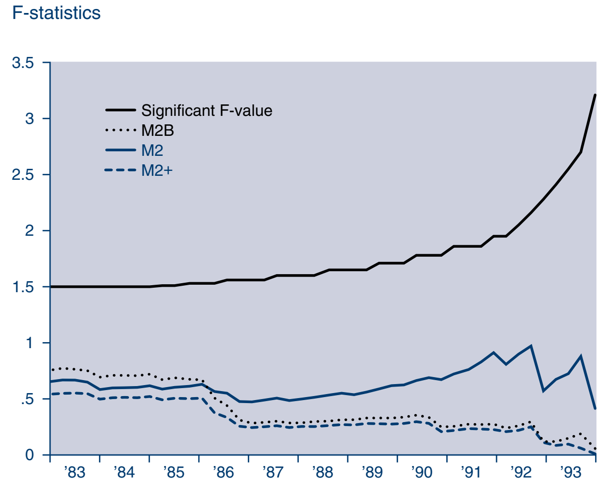
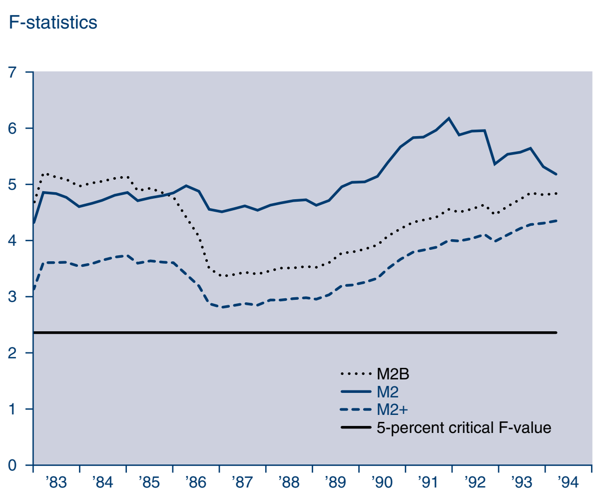


Figure 5
Joint Exclusion Tests for All Money Variables



[17] Although the critical F-values plotted in Figure 6 are not technically correct, Andrews' (1993) correction would raise the critical values, which would not affect the qualitative results since the lack of stability is rejected using the uncorrected critical F-values.

[18] This finding is consistent with other runs (not shown) that added variables interacting DUM92 with lagged money growth and the error-correction term. These variables were insignificant, with the exception that the product of DUM92 and EC was significant at the 10-percent level in the noninterest rate model using M2 without terms interacting DUM92 and lags of money growth.

Table 4
Dummy Variable Tests

Dummy Variable Tests for an Early-1990s Shift in the Constant

Added variable	Money only			Money and interest rates		
	M2	M2B	M2+	M2	M2B	M2+
DUM92	.0078*	.0022	.0008	.0046	.0027	.0009
	(2.04)	(.63)	(.21)	(1.09)	(.68)	(.25)
$\bar{R}^2$	.2416	.2097	.1944	.2747	.2657	.2576

Dummy Variable Tests for a mid-1980s Shift in the Constant

Added variable	Money only			Money and interest rates		
	M2	M2B	M2+	M2	M2B	M2+
D8587	–.0063*	–.0097**	–.0084*	–.0051	–.0068+	–.0064+
	(–1.98)	(–2.95)	(–2.55)	(–1.46)	(–1.94)	(–1.81)
$\bar{R}^2$	.2401	.2589	.2340	.2805	.2858	.2775

T-statistics are in parentheses.
** (*,+) denotes significance at the 1-percent (5-percent, 10-percent) level.

To test this hypothesis, a dummy variable (*D8587*), equal to 1 over 1985:1–87:1 and 0 otherwise, is added to models 1 through 3 and 22 through 24 to test for a temporary shift in the constant occurring over 1985:1–87:1. In models containing interest rates, *D8587* is significant only in models using M2B or M2+ and with a negative effect that helps control for how portfolio shifts from non-M2 assets bolstered bond and equity funds relative to M2 (*Table 4*). In models without interest rates, *D8587* is significant in models using M2, M2B, and M2+, suggesting that M2 may have also been bolstered by shifts away from direct security holdings. With or without interest rates, the models using M2B have higher R^2s than those of corresponding models using M2 or M2+, with the latter having similar R^2s.[19]

These results imply that the links between nominal GDP and M2B shifted in the mid-1980s, and the shifts are best modeled as a temporary shift in the constant term. Furthermore, the larger estimated impact of *D8587* in the models using M2B and M2+ relative to corresponding models using M2 supports the view that bond and equity fund inflows in the mid-1980s partly reflected shifts away from direct security holdings.

Conclusion

Recent instability in M2 and portfolio shifts into bond and equity funds have raised the issue

[19] *A dummy (*MFCUM*) was added to test for an increasing shift in the constant term during the mid-1980s that then levels out and becomes permanent. Mimicking movements in the shares of bond funds in M2B and bond and equity funds in M2+,* MFCUM *equals 0 before 1985:1, 1 in 1985:1, rises by 1 each quarter through 1987:1, and equals 9 after 1987:1.* MFCUM *was (marginally) significant only in the models using M2B. For models using M2B and M2+, adding* D8587 *increased model fit (*R^2*s) more than adding* MFCUM *instead.*

Alternatively, M2B and M2+ could have had a different relationship to nominal GDP growth in the mid-1980s. To test this, other runs (not shown) added variables to models 1 through 3 and 22 through 24 interacting D8587 *with lags of money growth and the error-correction term. These interactive variables were insignificant, with the exception that they were jointly significant at the 10-percent level in the noninterest rate model using M2B. This finding highlights the importance of including interest rate information.*

of whether M2 should be more broadly defined to include either bond or bond and equity funds. Two criteria for addressing this issue are whether the broader aggregates are more controllable and whether they are better information variables that can be used to forecast nominal variables.

With respect to controllability, adding assets less directly influenced by the Federal Reserve to M2 would likely make M2 less controllable. However, because broader M2 aggregates internalize portfolio shifts that may be induced by Federal Reserve actions affecting interest rates, broader M2 aggregates may be less volatile if the impact of such shifts outweighs the impact of variation in securities prices on the value of bond and stock fund assets. This is an empirical issue that requires further research and more data.

As for judging which M2 aggregates are better information variables, several criteria include whether a more broadly defined M2 aggregate is more explainable in money demand models, yields better inflation forecasts, and is a better near-term indicator of nominal GDP growth. Previous work has shown that the demand for M2B may be more explainable than that of M2 (Duca, forthcoming) and that M2B and M2+ yield more accurate forecasts of inflation in the early 1990s than does M2 within the P-star framework (Becsi and Duca, forthcoming, and Duca 1994).

This article focuses on the relative ability of these aggregates to predict nominal GDP growth and finds that M2B and M2+ have outperformed M2 in recent years. However, when money is used in conjunction with short- and long-term interest rates, this recent advantage is not as large. Interestingly, although M2B yields better in-sample fits than M2+, M2+ has performed better recently, consistent with reports that stock funds are being increasingly used as substitutes for some M2 deposits.

These findings imply that along with M2, M2B and M2+ should be monitored as information variables. Nevertheless, given current changes in asset behavior and that past financial innovations have altered asset portfolios, the link between broad monetary aggregates and economic activity is vulnerable to shifts. Such shifts can stem not only from technological change or new financial products but also from shifts in monetary policy and household preferences that alter time series relationships. This was true not only for M2 in the early 1990s but also for M2B and M2+ in the mid-1980s. These considerations point to the need for further research and to the wisdom of not relying on any single monetary aggregate as the sole guide to setting monetary policy.

References

Anderson, Richard, and Sean Collins (1994), "Modelling U.S. Households' Demands for Liquid Wealth in an Era of Financial Change," mimeo, Federal Reserve Bank of St. Louis (St. Louis, May).

Andrews, Donald W. K. (1993), "Tests for Parameter Instability and Structural Change With Unknown Change Point," *Econometrica* 61 (July): 821–56.

Balke, Nathan A., and Kenneth M. Emery (1994), "The Federal Funds Rate as an Indicator of Monetary Policy: Evidence From the 1980s," Federal Reserve Bank of Dallas *Economic Review*, First Quarter, 1–16.

Becsi, Zsolt, and John V. Duca (forthcoming), "Adding Bond Funds to M2 in the P-Star Model of Inflation," *Economics Letters*.

Bernanke, Ben S. (1990), "On the Predictive Power of Interest Rates and Interest Rate Spreads," *New England Economic Review*, November/December, 51–68.

———, and Alan S. Blinder (1992), "The Federal Funds Rate and the Channels of Monetary Transmission," *American Economic Review* 82 (September): 901–21.

Brayton, Flint, and Peter A. Tinsley (1993), "Interest Rate Policies for Price Stability," FEDS Working Paper no. 93-22, Board of Governors of the Federal Reserve System (Washington, D.C., July).

Collins, Sean, and Cheryl L. Edwards (1994), "Redefining M2 to Include Bond and Equity Mutual Funds," Federal Reserve Bank of St. Louis *Review*, November/December.

Duca, John V. (forthcoming), "Should Bond Funds Be Included in M2?" *Journal of Banking and Finance*.

——— (1994), "Mutual Funds and Monetary Policy," Federal Reserve Bank of St. Louis *Review*, November/December.

——— (1993a), "Monitoring Money: Should Bond Funds Be Added to M2?" Federal Reserve Bank of Dallas *Southwest Economy*, June, 1–8.

——— (1993b), "Regulation, Bank Competitiveness, and Episodes of Missing Money," Federal Reserve Bank of Dallas *Economic Review*, Second Quarter, 1–24.

——— (1992), “The Case of the Missing M2,” Federal Reserve Bank of Dallas *Economic Review* (Second Quarter): 1–24.

Emery, Kenneth M. (1994), “Inflation Persistence and Fisher Effects: Evidence of a Regime Change,” *Journal of Economics and Business* (August): 141–52.

Feinman, Joshua, and Richard D. Porter (1992), “The Continued Weakness in M2,” FEDS Working Paper no. 209, Board of Governors of the Federal Reserve System (Washington, D.C., September).

Feldstein, Martin, and James H. Stock (1994), “Measuring Money Growth When Financial Markets Are Changing,” NBER Working Paper Series, no. 4888 (Cambridge, Mass.: National Bureau of Economic Research, October).

Friedman, Benjamin M., and Kenneth N. Kuttner (1992), “Money, Income, Prices, and Interest Rates,” *American Economic Review* 82 (June): 472–92.

Hafer, R. W., and Kutan, A. M. (1992), “On the Money-Income Results of Friedman and Kuttner,” Southern Illinois University at Edwardsville Working Paper no. 92-0303 (Edwardsville, Illinois).

Hallman, Jeffrey J., Richard D. Porter, and David H. Small (1991), “Is the Price Level Tied to the M2 Monetary Aggregate in the Long Run?” *American Economic Review* 81 (September): 841–58.

Hetzel, Robert L., and Yash P. Mehra (1989), “The Behavior of Money Demand in the 1980s,” *Journal of Money, Credit, and Banking* 21 (November): 455–63.

Johansen, Soren, and Katarina Juselius (1990), “Maximum Likelihood Estimation and Inference on Cointegration—with Applications to the Demand for Money,” *Oxford Bulletin of Economics and Statistics* 52 (May): 169–210.

Milbourne, Ross (1986), “Financial Innovation and the Demand for Liquid Assets,” *Journal of Money, Credit, and Banking* 18 (November): 506–11.

Moore, George R., Richard D. Porter, and David H. Small (1990), “Modeling the Disaggregated Demands for M1 and M2 in the 1980’s: The U.S. Experience,” in *Financial Sectors in Open Economies: Empirical Analysis and Policy Issues*, P. Hooper, K. H. Johnson, D. L. Kohn, D. E. Lindsey, R. D. Porter, and R. Tryon, eds. (Washington, D.C.: Board of Governors of the Federal Reserve System), 21–105.

Orphanides, Athanasios, Brian Reid, and David H. Small (1994), “Empirical Properties of a Monetary Aggregate that Adds Bond and Stock Funds to M2,” Federal Reserve Bank of St. Louis *Review*, November/December.

Rudebusch, Glenn R., and David Wilcox (1994), “Productivity and Inflation: Evidence and Interpretations,” manuscript, Board of Governors of the Federal Reserve System (Washington, D.C., May).

Simpson, Thomas D. (1980), “The Redefined Monetary Aggregates,” *Federal Reserve Bulletin* (February): 97–114.

Small, David H., and Richard D. Porter (1989), “Understanding the Behavior of M2 and V2,” *Federal Reserve Bulletin* (April): 244–54.

Stock, James, and Mark Watson (1989), “New Indexes of Coincident and Leading Economic Indicators,” in *NBER Macroeconomics Annual 1989*, O.J. Blanchard and S. Fischer, eds. (Cambridge, Mass.: The MIT Press), 351–94.

Wynne, Mark A. (1993), “Price Stability and Economic Growth,” Federal Reserve Bank of Dallas *Southwest Economy*, May/June, 1–5.

Nathan S. Balke
Associate Professor of Economics
Southern Methodist University and
Research Associate
Federal Reserve Bank of Dallas

Kenneth M. Emery
Senior Economist
Federal Reserve Bank of Dallas

Understanding the Price Puzzle

Few economists expected the bond market's negative reaction to the Federal Reserve's initial move to raise the federal funds rate during early 1994. To alleviate potential inflationary pressures, on February 4, 1994, the Federal Reserve increased the federal funds rate by 25 basis points to 3.25 percent. In response, long-term bond yields promptly increased 50 basis points over the following four weeks. At the time, bond market participants attributed much of the run-up in yields to worries that inflation would increase during the next year, thus eroding the value of fixed-income securities.

While many economists were caught off guard by the bond market's reaction, historical data on the federal funds rate and subsequent inflation behavior perhaps explain the reaction. Historical data show a positive relationship between inflation and the federal funds rate, the rate over which the Federal Reserve has the most control.[1] Thus, one explanation for the bond market's behavior is that increases in the federal funds rate have historically been associated with subsequent increases in inflation. If history is any guide, bond market participants were right to be worried.

The positive relationship between the federal funds rate and inflation has become known as the "price puzzle" (Bernanke and Blinder 1992; Christiano, Eichenbaum, and Evans 1994, forthcoming; and Sims 1992). It is a puzzle because an unexpected tightening of monetary policy (that is, an unexpected increase in the federal funds rate) is expected to be followed by a decrease in the price level, rather than an increase.

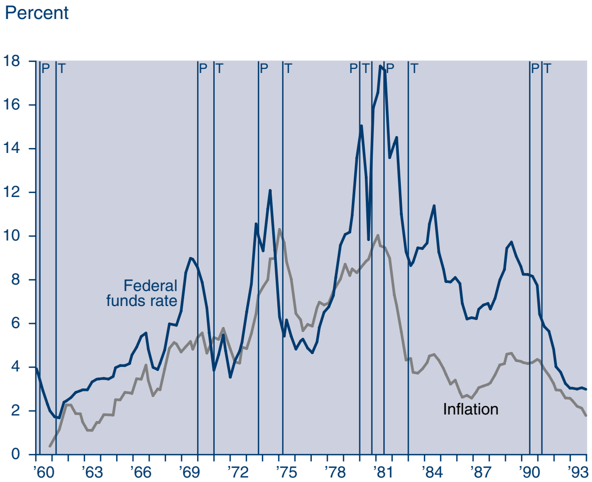
In this article, we document the positive correlation between federal funds rate increases and subsequent increases in prices. The strength of this correlation does not appear to be uniform over the postwar period. In previous work (Balke and Emery 1994), we found that relationships that had held in the 1960s and 1970s broke down in the 1980s. As a result, we also evaluate whether the price puzzle is present to the same degree in all periods. We find that evidence of the price puzzle is substantially stronger during the 1960s and 1970s than during the 1980s. In the 1980s, the correlation between the federal funds rate and future inflation is close to zero but is still not negative, as traditional theory would predict.

We also evaluate possible explanations for the price puzzle. These involve the Federal Reserve systematically responding to signals of higher future inflation by raising the federal funds rate, but not by enough to fully offset the subsequent inflation. Indeed, a plausible explanation appears to be that, during the 1960s and 1970s, the Federal Reserve responded to supply shocks by raising the federal funds rate but not by enough to prevent the aggregate price level from changing. Thus, a positive correlation between the federal funds rate and inflation arises. Since the early 1980s, however, the price puzzle has moderated. We suggest two possible reasons: the Federal Reserve has put more emphasis on achieving price stability and, hence, has responded more

The authors thank John Duca and Evan Koenig for helpful comments and suggestions. We also thank Chih-Ping Chang for his capable research assistance. Any remaining errors are the authors' responsibility.

[1] *The Federal Reserve influences this rate by buying or selling U.S. Treasury securities to the private sector, thus controlling the amount of reserves in the banking system. The federal funds rate is simply the market interest rate that banks must pay to borrow reserves overnight.*

Figure 1
Federal Funds Rate and Inflation



SOURCE: Board of Governors, Federal Reserve System.

vigorously to inflationary shocks, or there simply have been fewer large inflationary shocks to the economy.

This article is organized as follows. In the first section, we document the positive correlation between federal funds rate changes and subsequent inflation. In the second section, we present possible explanations for the price puzzle. Both explanations involve the Federal Reserve's systematic response to inflationary shocks. The third section examines whether these explanations are capable of explaining the price puzzle. Finally, in the fourth section, we conclude by interpreting the empirical results.

Documenting the price puzzle

The price puzzle arises because increases in the federal funds rate tend to be followed by increases in inflation. In this section, we document the existence of this positive correlation.

As shown in Figure 1, the federal funds rate and year-over-year inflation, as measured by the gross domestic product deflator, appear to be positively related: periods of high inflation coincide with relatively high federal funds rates. Note that during the 1960s and 1970s, both inflation and the federal funds rate have trended upward, whereas since 1982, the trend in inflation has been flat, while the federal funds rate has trended downward.

More formal evidence of the price puzzle is given in Table 1. Table 1 presents results from simple regressions of the federal funds rate against the average annualized rate of inflation over two subsequent years for the sample 1960:1 through 1993:4. Whether or not a time trend is included in the regressions, the results confirm a positive relationship between the federal funds rate and future inflation. For the full sample and the 1960–79 subsample, the federal funds rate is highly significant. Interestingly, for the 1982:4–93:4 sample, the federal funds rate does not contain statistically significant explanatory power for subsequent inflation as reflected in an insignificant t-statistic on the federal funds rate and a negative adjusted-R^2 for the regression.[2] This suggests that the price puzzle may not be as evident in the post-1982 period.

The federal funds rate, monetary policy, and the price puzzle

Historically, changes in the quantity of money have often served as a measure of monetary policy. The main problem with money, however, is that it often changes for reasons that have nothing to do with monetary policy. For example, most measures of money are influenced by the behavior of both banks and individuals, which, in turn, are influenced by economic conditions. In other words, the observed data on money represent a confluence of both supply factors (monetary policy actions) and demand factors (such as private-sector portfolio shifts). The problem with using money as a measure for the stance of monetary policy is that it does not reflect mainly Federal Reserve actions.

Recently, several economists have argued that movements in the federal funds rate may be a better indicator of changes in monetary policy than are changes in the quantity of money (McCallum 1983, Laurent 1988, Bernanke and Blinder 1992,

[2] *Throughout the paper, we do not separately examine the 1979:4 through 1982:3 sample because during this period, in contrast to the rest of the sample, the Federal Reserve did not target the federal funds rate in its implementation of monetary policy.*

Table 1
Federal Funds Rate Regressed on Subsequent Inflation[1]

	Trend not included		Trend included	
Sample	Coefficient on federal funds rate	adj. R^2	Coefficient on federal funds rate	adj. R^2
1960:1–93:4	.26 (2.11)*	.15	.30 (2.38)*	.15
1960:1–79:3	.75 (4.91)*	.56	.17 (2.72)*	.92
1982:4–93:4	.09 (.59)	.01	.05 (.43)	0

[1] Subsequent inflation equals the average annualized rate of inflation over the subsequent eight quarters as measured by the gross domestic product deflator.

T-statistics are in parentheses. To correct for possible heteroscedasticity as well as serial correlation, the White consistent covariance matrix with a Newey–West serial correlation correction and with a window width of twelve lags was estimated.

* Significant at the 5-percent level.

Quarterly data.

SOURCE: Board of Governors, Federal Reserve System.

and Goodfriend 1992). This view is based on the observation that, with the exception of the 1979–82 period, the Federal Reserve has implemented monetary policy by targeting the federal funds rate over short periods of time.

There is a potential shortcoming in directly using the federal funds rate as a measure of the stance of monetary policy. Movements in the rate reflect both the Federal Reserve's response to economic developments, as well as Federal Reserve actions that are independent, or exogenous, of these developments.[3] To assess the impact of exogenous monetary policy actions on the economy, several studies have used empirical models called *vector autoregressions* (VARs). (See the box entitled "Vector Autoregressions.") Basically, these models attempt to isolate the movements in the federal funds rate that are uncorrelated with changes in the other variables in the model and, thereby, represent purely exogenous movements in the federal funds rate, or exogenous monetary policy actions.

Using the VAR methodology, Bernanke and Blinder (1992), Sims (1992), and Christiano, Eichenbaum, and Evans (1994, forthcoming) have found that movements in the federal funds rate are largely consistent with the view that the funds rate is a good proxy for the stance of monetary policy. In this work, the Federal Reserve tightens policy in response to unexpected increases in both inflation and output. Additionally, unexpected, or exogenous, monetary policy actions are shown to have modest effects on real output. However, even in these VARs the price puzzle remains: exogenous monetary policy tightenings are followed by increases in the price level.

Consider a five-variable VAR similar to that examined by Christiano, Eichenbaum, and Evans (1994). This VAR includes real GDP (Y), the GDP deflator (P), the federal funds rate, nonborrowed reserves, and total reserves.[4] Figure 2 displays the response of output and prices to a so-called ex-

[3] *The rule that relates policy actions to developments in the economy is often referred to as the* feedback rule.

[4] *With the exception of the federal funds rate, all variables are in logarithms.*

Figure 2
Impulse Response of *Y* and *P*
To Federal Funds Rate Innovation, 1960–93

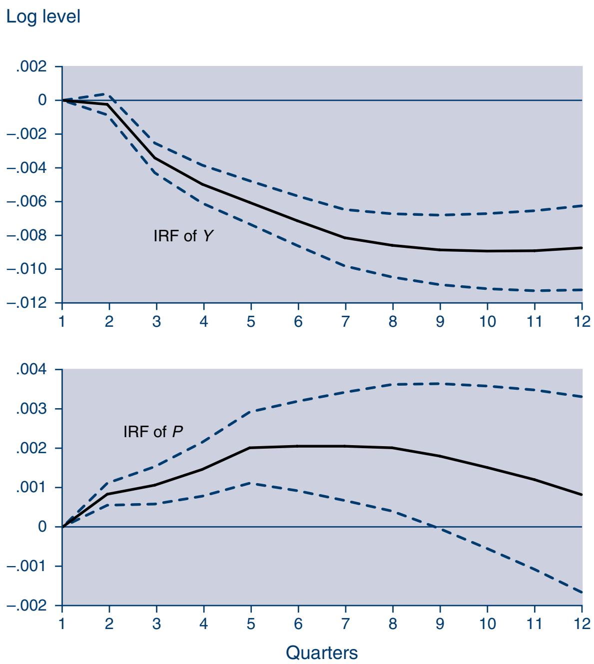
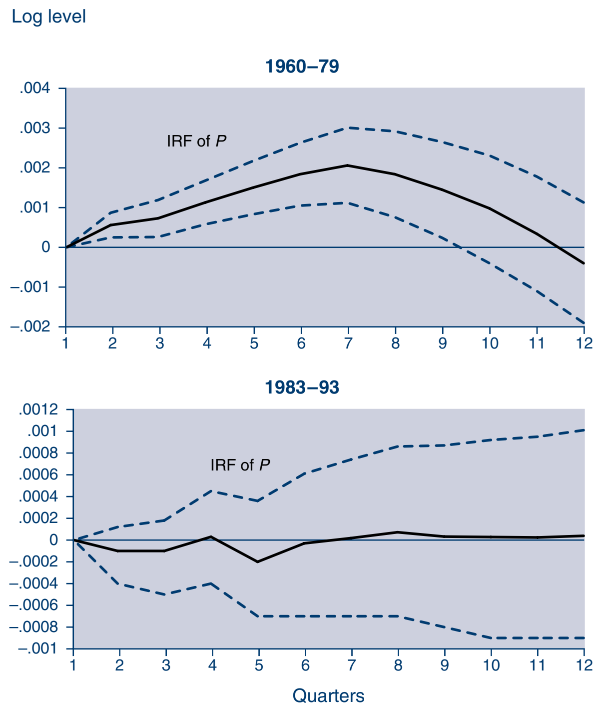


Figure 3
Impulse Response of *P*
To Federal Funds Rate Innovation



ogenous increase in the federal funds based on the entire 1960–93 sample.[5] While output falls in response to a monetary contraction, prices rise. Thus, the price puzzle is present.

When the full sample is broken into subsamples, the VAR evidence of a price puzzle becomes mixed.[6] Figure 3 plots the response of prices to a federal funds shock in the 1960–79 period and the post-1982 period. As shown in the figure, the price puzzle is present in the 1960–79 period; an exogenous increase in the federal funds rate results in a substantial increase in the price level. By contrast, the price puzzle is not as evident in the post-1982 period. During this period, a federal funds rate innovation does not cause prices to systematically rise; the effect on prices, though negative, is small and not statistically different from zero.

An alternative to the VAR approach

Rather than using the federal funds rates as an indicator of monetary policy or trying to identify monetary interventions econometrically, Romer and Romer (1989) examine the historical record to determine dates when a contractionary monetary policy action was taken.[7] To evaluate whether the price puzzle still exists using the Romer–Romer dates as proxies for monetary contractions, we run a regression of inflation (or prices) against four lags of inflation and the current value and eight

[5] *The federal funds rate comes third, after output and prices, in the causal ordering implied by the Choleski decomposition. The dotted lines in Figures 3 through 12 are one-standard error confidence bands.*

[6] *We tested whether a VAR estimated over 1960:1–79:3 was the same as a VAR estimated over 1982:4–93:4 (see Doan 1992) and could reject the null hypothesis of equality. This result was robust to whether the VAR was estimated in levels or first differences.*

[7] *The Romer–Romer dates are: 1968:4, 1974:2, 1978:3, 1979:4. The date 1988:3 is also included as a monetary contraction, based on work by Oliner and Rudebusch (1992).*

Vector Autoregressions

Vector autoregressions (VARs) are time series models that use only past values of the variables of interest to make forecasts. For instance, a three-variable VAR system of output, prices, and the federal funds rate can be expressed as

$$Y_t = \beta_1 + \beta_2 \Sigma\, Y_{t-i} + \beta_3 \Sigma\, P_{t-i} + \beta_4 \Sigma\, f_{t-i} + \epsilon_{\beta t}$$
$$P_t = \gamma_1 + \gamma_2 \Sigma\, Y_{t-i} + \gamma_3 \Sigma\, P_{t-i} + \gamma_4 \Sigma\, f_{t-i} + \epsilon_{\gamma t}$$
$$f_t = \delta_1 + \delta_2 \Sigma\, Y_{t-i} + \delta_3 \Sigma\, P_{t-i} + \delta_4 \Sigma\, f_{t-i} + \epsilon_{\delta t},$$

where Y, P, and f are output, the price level, and the federal funds rate, respectively. β is an intercept term, t is a time subscript, Σ is the summation sign, and ϵ is an error term. Thus, each of the three variables is expressed as a linear function of past values of itself and past values of other variables in the system.

In practice, the estimated error terms from each equation are correlated so that it is incorrect to assume that, for instance, ϵ_{ft} represents an independent surprise movement in the federal funds rate. To better interpret the dynamic relationships present in the data, the residuals from the VAR are broken up into linear combinations of independent (orthogonal) shocks. A common orthogonalization is to assume that the VAR system is *recursive* so that there is a chain of causality among surprises in the variables during any given period. For example, a possible recursive system of the VAR above is one in which output responds to an exogenous shock, the price level responds to the contemporaneous output shock and an exogenous price shock, while the federal funds rate responds to output and price level shocks contemporaneously as well as to an exogenous federal funds rate shock. In effect, a new set of surprises, or shock terms for each variable, are created that are now uncorrelated with each other. The transformation of the original shocks into recursive, orthogonal shocks is called the *Choleski decomposition*.

The Choleski decomposition is controversial because if the VAR is used to draw economic inferences, then the recursive restriction imposed on the system should be supported by economic theory. If the identifying assumption of recursivity is not justified, then the estimated parameters will be a mixture of both structural and reduced-form parameters.

For more on VARs, see Todd (1990), Runkle (1987), Sims (1986), Cooley and LeRoy (1985), and Hakkio and Morris (1984).

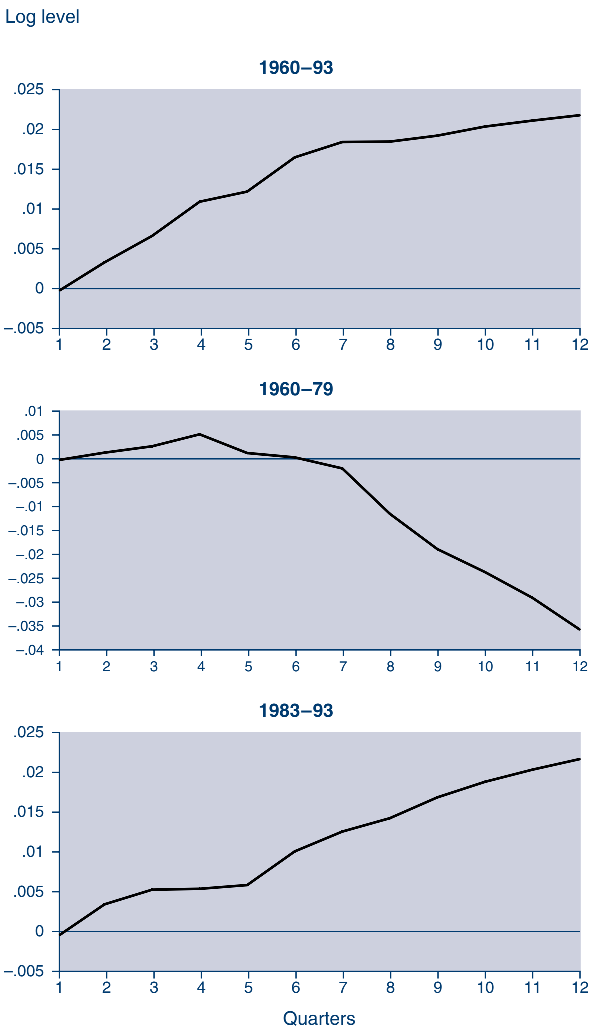
lags of a dummy variable that takes the value of 1 during a Romer–Romer monetary contraction date and zero otherwise. Figure 4 plots the dynamic response of the price level in response to the Romer–Romer dummy. For the full sample and the post-1982 sample, we see evidence of the price puzzle; prices tend to rise following a Romer–Romer monetary contraction. For the 1960–79 sample, the response of prices is initially flat and then falls six quarters after a Romer–Romer contraction. This suggests that the Romer dates may be a better proxy for monetary contractions in the early part of the sample than in the later part.

Overall, there appears to be a positive correlation between federal funds rate increases and future inflation. However, the extent of this correlation is sensitive to sample periods examined. With the exception of the Romer–Romer dates, the price puzzle is much more evident in the pre-1980 sample than in the post-1982 sample. This result holds for sophisticated VAR or intervention analysis as well as for simple regressions. In the next section, we explore some explanations for the positive correlation.

Explanations of the price puzzle

One explanation of the price puzzle suggested by Sims (1992) is that the Federal Reserve systematically responds to expectations of higher future inflation by raising the federal funds rate but by not enough to prevent inflation from actually rising. The result is that increases in the federal funds rate are followed by increases in inflation. The Sims explanation thus involves a forward-looking Federal Reserve that, nonetheless, fails to effectively prevent the anticipated future inflation. It also implies that the positive correlation between an apparently contractionary monetary policy intervention and future prices stems, in part, from failing to properly identify exogenous changes in the federal funds rates. Sims suggests that the Federal Reserve has information about future inflation that is not present in the simple VARs

Figure 4
Impulse Response Functions
Of *P* to Romer–Romer Dates



described in the previous section, and as a result, innovations in the federal funds rate from these VARs partly reflect the systematic response to inflationary shocks and are not truly exogenous.

An alternative, but similar, explanation is that the Federal Reserve reacts to supply shocks by raising the federal funds rate. A temporary, negative supply shock, for example, would have the effect of raising real interest rates, decreasing output, and increasing prices (at least in the short run). The Federal Reserve responds to the supply shock by raising the federal funds rate but by not enough to extinguish the inflationary consequences of the supply shock. Note that the supply shock explanation can explain both the price puzzle and the negative response of output to a positive federal funds rate innovation, even if monetary policy has no effect on the real economy. Furthermore, the degree to which the monetary authority is willing to "extinguish" the price increase might depend on the weight it places on price stability in its objective function relative to output stabilization; the greater the weight on price stability, the more aggressively the monetary authority reacts to the supply shock and, hence, the smaller the price puzzle.

Interestingly, for both explanations, the fact that the price puzzle appears muted in the 1980s suggests that the Federal Reserve's reaction function may have changed. The Federal Reserve may have become more forward looking in its inflation fighting effort (trying to stay ahead of the curve) and has more effectively preempted inflationary pressures. Alternatively, it may have placed more weight on price stability when reacting to supply shocks.[8]

Is the price puzzle solved?

These explanations for the price puzzle revolve around the Fed's response to inflationary pressures. Thus, to effectively evaluate these explanations within a VAR framework, one must introduce a variable into the system that contains information about future inflation or supply shocks that is not already contained in the existing VAR. Furthermore, evaluating these explanations involves examining the Fed's reaction function; namely, how does the federal funds rate respond to possible inflationary shocks? Much of the previous literature has failed to note the Fed's reaction function.

Christiano, Eichenbaum, and Evans (1994, forthcoming) have argued that the price puzzle is

[8] *Possibly, though, supply shocks may have been relatively small or positive since 1982 so that the Fed could focus more on its price stability objective. This explanation implies an asymmetric objective function in which the Fed is less willing to extinguish the inflationary consequences of negative supply shocks (and suffer the negative output consequences) than positive supply shocks.*

Figure 5
Impulse Response of *Y* and *P*
To Federal Funds Rate Innovation, 1960–93
(Commodity Prices Included)

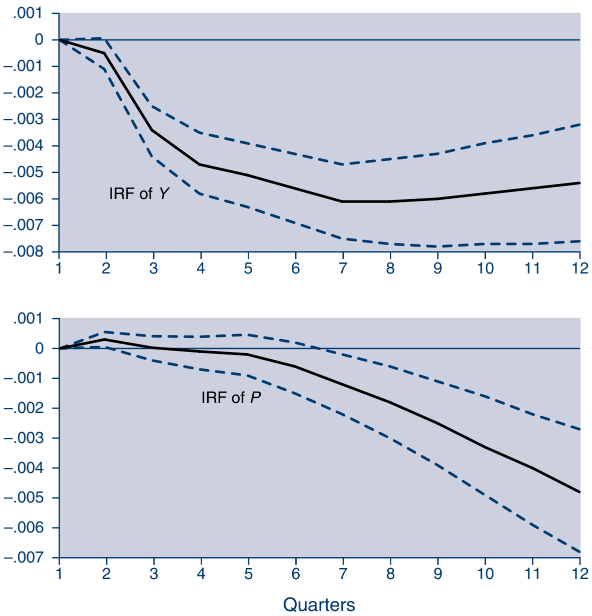
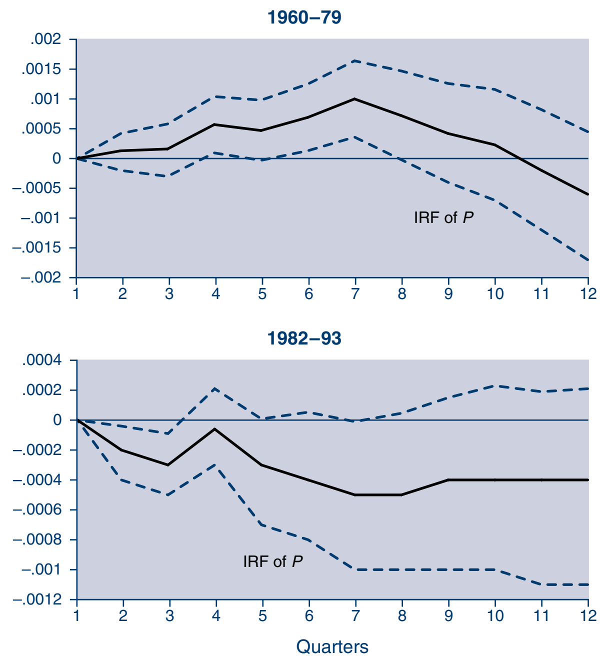


Figure 6
Impulse Response of *P*
To Federal Funds Rate Innovation
(Commodity Prices Included)



resolved when commodity prices are included in the basic VAR examined earlier. Commodity prices have been suggested to provide information about future inflation and could also be correlated with supply shocks. Thus, a priori, they are a good candidate for inclusion in the VAR. We replicate Christiano, Eichenbaum, and Evans' results when we include commodity prices in the basic vector autoregression.[9] Figure 5 plots the response of output and prices to a federal funds rate shock once commodity prices have been included in the VAR. As in Christiano, Eichenbaum, and Evans, a positive federal funds rate shock causes prices and output to fall. Furthermore, the reaction of the federal funds rate to a commodity price shock is consistent with the Sims and supply shock explanations of the price puzzle: a positive commodity price shock causes the federal funds rate to rise (not shown in figures). Indeed, the reaction of the federal funds rate to positive output and price shocks is consistent with a "lean against the wind" policy on the part of the Fed.

While including commodity prices in the VAR appears to solve the price puzzle for the full sample, when we examine the subsamples the evidence is less conclusive. Figure 6 presents the impulse response function of prices to a federal funds rate shock in the 1960–79 and 1982–93 subsamples.[10] While commodity prices succeed in eliminating the price puzzle in the latter period, they do not solve the price puzzle in the pre-1980 period. For the 1960–79 sample, prices are higher than their original level for nearly three years after a positive federal funds rate shock. Thus, even after includ-

[9] *In the vector autoregression, commodity prices come after output and aggregate prices but before the federal funds rate in the Choleski ordering.*

[10] *We tested whether the VAR parameter estimates were the same in the 1960–79 and post-1982 samples and could reject equality.*

Figure 7
Impulse Response Functions of *P*
To Federal Funds Rate
(Spread Included)

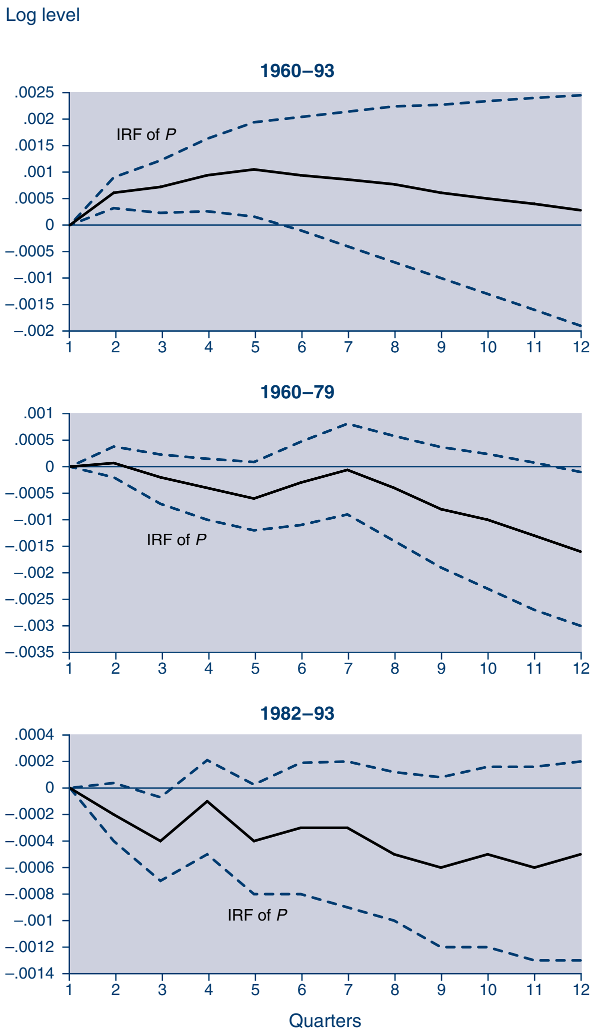
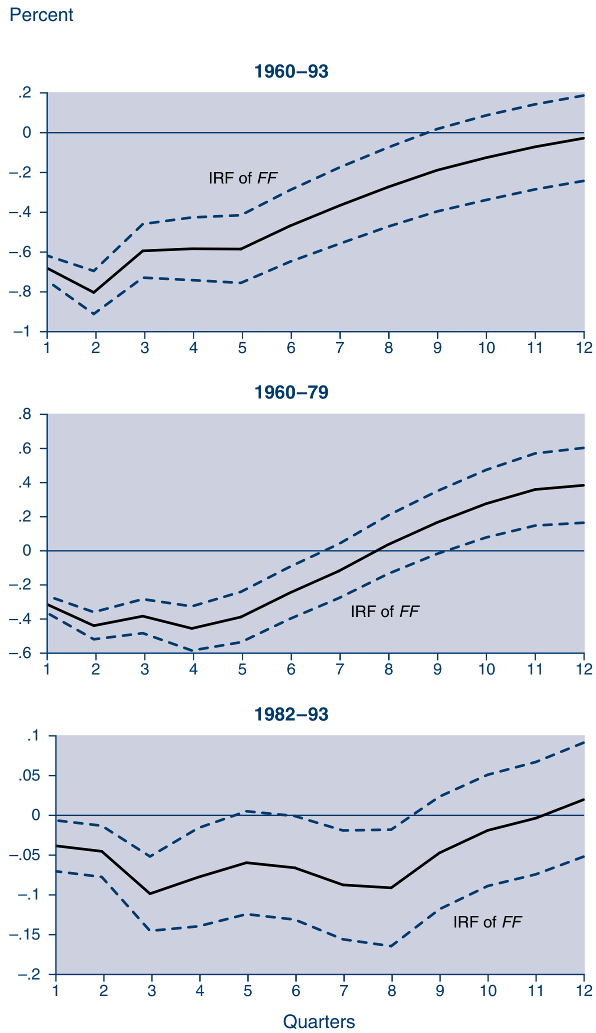


Figure 8
Impulse Response Functions
Of Federal Funds Rate (*FF*) to Spread
(Spread Included)



ing commodity prices in the VAR, the price puzzle is still unresolved for the 1960–79 period.

We examine whether other variables might be able to solve the price puzzle; these include the spread between long- and short-term interest rates, short- and long-term interest rates individually, oil prices, stock prices, unit labor costs, the index of leading economic indicators, and industrial capacity utilization.[11] Of these, only the spread between long- and short-term interest rates helps solve the price puzzle, but it does so only if the period 1979–82 is excluded.[12] Figure 7 shows the

[11] These were done one variable at a time to keep the dimension of the VAR relatively low. For the impulse responses, each of these variables is placed third in the Choleski ordering, behind output and prices but before the federal funds rate.

[12] The spread variable is defined to be the ten-year Treasury bond rate minus the three-month Treasury bill rate.

Figure 9
Impulse Response Functions
Of *Y* and *P* to Spread, 1960–93
(Spread Included)

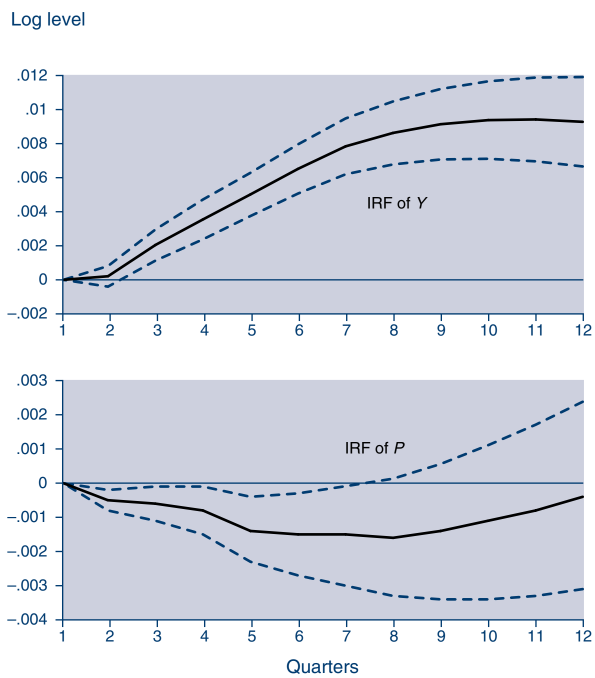
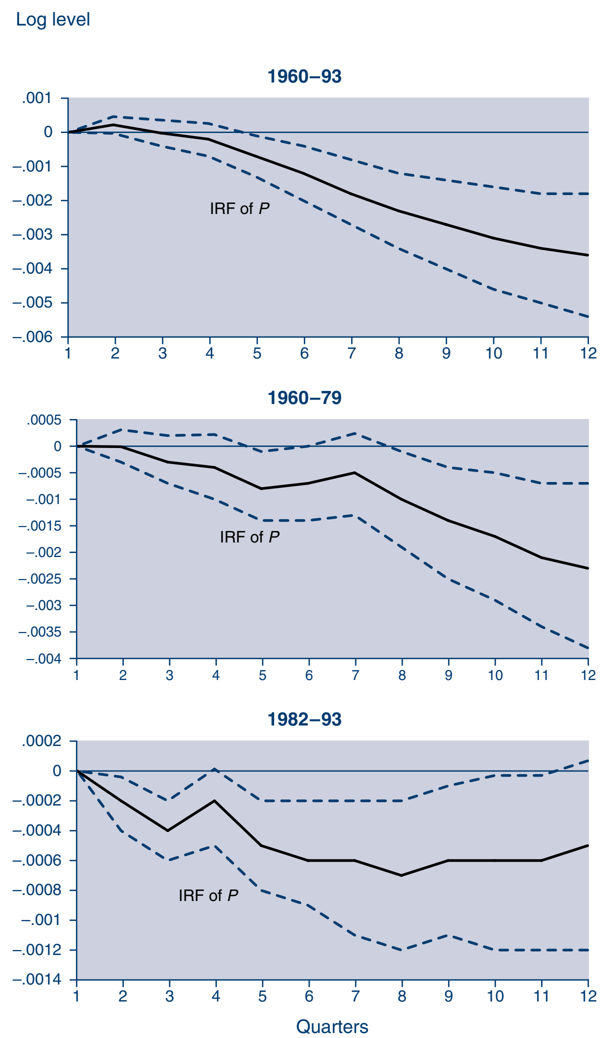


Figure 10
Impulse Response Functions of *P*
To Federal Funds Rate
(Commodity Prices and Spread Included)



response of prices to a federal funds shock for the full sample as well as the 1960–79 and post-1982 subsamples once the spread has been included in the VAR. In the full-sample VAR, including the spread does not solve the price puzzle; prices still rise in response to a positive federal funds shock. However, unlike commodity prices, including the spread eliminates the price puzzle during the 1960–79 period. The spread also eliminates the price puzzle in the post-1982 period. That the spread fails to work in the full sample may be the result of the extreme volatility exhibited by interest rates, including the federal funds rate, during the 1979–82 period.

Because the reaction function of the Fed is integral to explaining the price puzzle, we present the response of the federal funds rate to a shock in the interest rate spread (*Figure 8*). The impulse response functions suggest that the federal funds rate tends to move in the *opposite* direction of a spread shock; a positive spread shock (that is, long-term rates rising more than short-term rates) causes the federal funds rate to fall. The negative response of the federal funds rate to an increase in the spread does not entirely square with the pure inflation expectations explanation of the price puzzle. A rise in inflation expectations would more likely be reflected in an increase in the interest rate spread, as long-term rates respond more than short-term rates to expectations of future inflation. The negative response of the federal funds rate to the increase in the spread is not, therefore, consistent with the Fed tightening

Figure 11
Impulse Response Functions of Federal Funds Rate (Commodity Prices, *PSC*, and Spread, *SP*, Included)

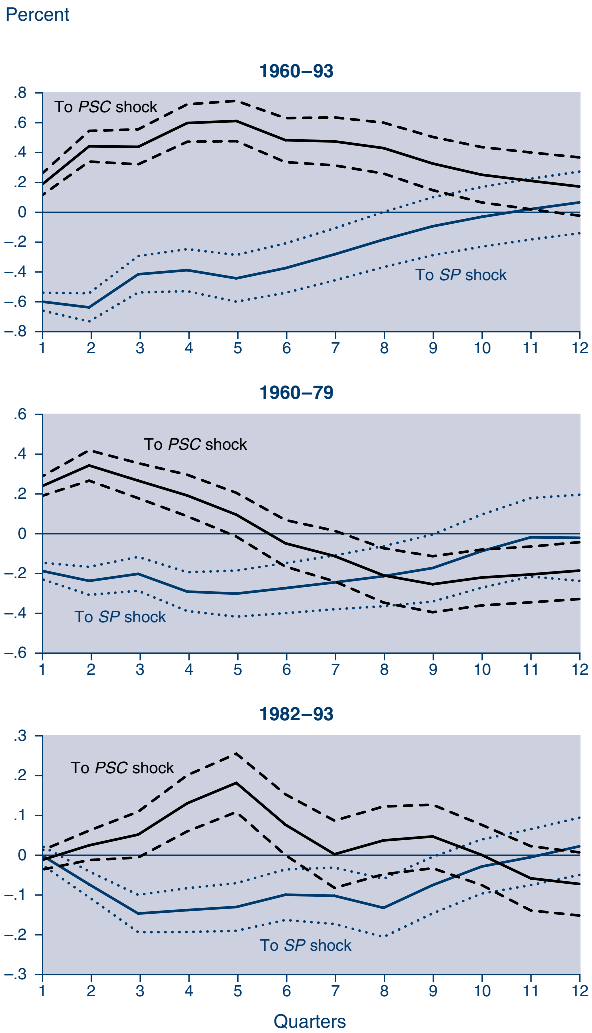
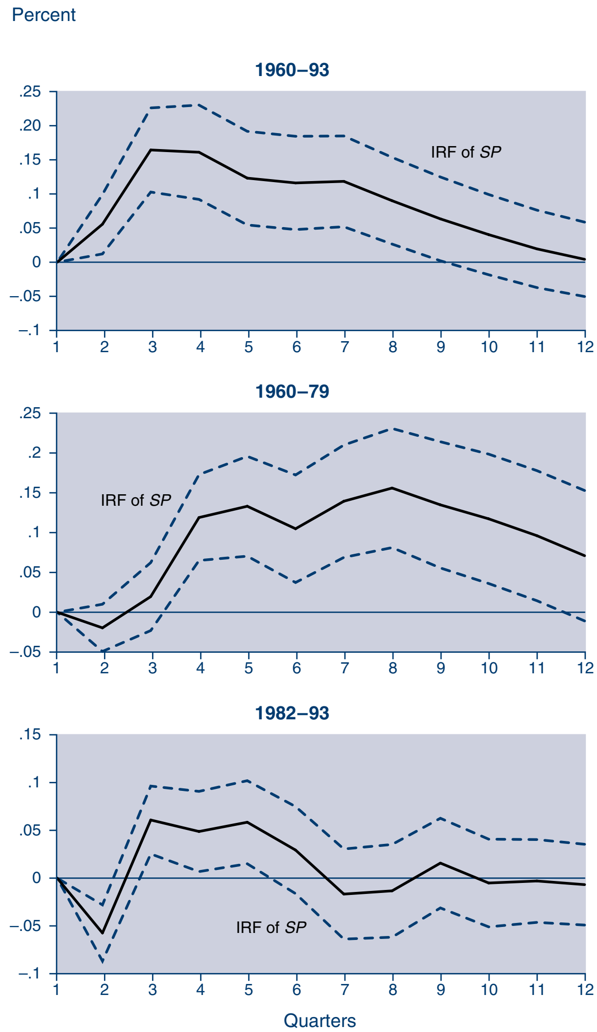


Figure 12
Impulse Response Functions Of Spread to Federal Funds Rate (Commodity Prices and Spread Included)



in anticipation of future inflation. However, it is consistent with the explanation that the Fed partially extinguishes the inflationary effects of a temporary supply shock. A temporary, negative supply shock, for example, would tend to cause the interest rate spread to fall as short-term real interest rates rise more than long-term real interest rates. At the same time the spread is falling, the Fed increases the federal funds rate to offset the inflationary effects of the supply shocks. This gives rise to the negative response of the federal funds rate to innovations in the interest rate spread.[13] The response to aggregate output and prices is also consistent with the supply shocks story as

[13] *Alternatively, the spread may decrease because of expectations that the Fed will tighten policy in response to a shock that increases prices and lowers output—namely, a supply shock.*

prices fall and output rises in response to a positive spread (supply) shock (*Figure 9*).

To determine the relative importance of the spread and commodity prices for explaining the price puzzle, especially in the later period, we estimate a VAR that includes both commodity prices and the interest rate spread.[14] Figure 10 displays the impulse response function of prices to shocks in the federal funds rate, and Figure 11 shows the response of the federal funds rate to shocks in commodity prices and the interest rate spread. For all three periods, the price puzzle no longer exists; prices respond negatively to a positive federal funds rate shock. Furthermore, the reaction of the federal funds rate to commodity price and interest rate spread shocks are consistent with the Fed tightening (at least initially) in response to an inflationary shock, in particular, supply shocks.[15] While we do not present the results here, the reaction of output and prices to commodity price shocks and interest rate shocks are consistent with those shocks reflecting supply-side shocks in the full and the 1960–79 samples. This interpretation is more tenuous in the post-1982 sample, as the response of GDP to commodity price and interest rate spread shocks is initially in the wrong direction for a supply shock.

Finally, an interesting result is that the interest rate spread rises in response to a federal funds rate innovation, particularly for the early sample period (*Figure 12*). This mirrors the response of the interest rate spread in spring 1994 as long-term interest rates rose more than short-term rates following an increase in the federal funds rate.

Conclusion

Using alternative approaches, we have documented a positive correlation between the federal funds rate and future inflation. Known as the "price puzzle," this positive correlation is surprising because increases in the federal funds rate, or tightenings in monetary policy, should theoretically lead to a lower, not a higher, price level. We have also documented that the price puzzle is much stronger during the pre-1980 period than since the early 1980s. In fact, the correlation between the funds rate changes and future price changes is close to zero in the later period.

We considered two explanations for the price puzzle. The first, from Sims (1992), is that the Federal Reserve systematically responds to expectations of higher future inflation by increasing the funds rate but not by enough to actually prevent inflation from rising. A second explanation is that the Federal Reserve systematically reacts to negative inflationary supply shocks by appropriately increasing the funds rate, but again, not by enough to extinguish the inflationary consequences of the shock.

We examine both of these explanations within a vector autoregression framework by including variables that may proxy for future inflation or supply shocks. While including commodity prices fully eliminates any price puzzle in the post-1980 period, the puzzle is still present in the pre-1980 period. We also find, however, that including the spread between ten-year and three-month Treasury rates eliminates the puzzle for both subsamples. Given the negative reaction of the funds rate to a shock in the spread, it appears that the spread is proxying for supply shocks rather than for increases in inflation expectations.

[14] *We conducted formal tests of structural stability and tests for whether commodity prices and/or the spread Granger caused the other variables in the system. Once again, structural stability was rejected when the 1960–79 and post-1982 subsamples were considered. Both commodity prices and the spread, individually and jointly, provide information about future values of the other variables for the full sample. The joint exclusion of commodity prices and the spread is rejected in the 1960–79 sample, but the spread variable can be excluded at the 10-percent significance level. For the post-1982 sample, one can exclude commodity prices and the spread jointly and can exclude the spread (at the 5-percent level) and commodity prices (at the 5-percent but not at the 10-percent level) individually. Note that a formal rejection of Granger causality for both commodity prices and the spread in the later sample may be due to the rather large number of parameters implied by the seven-variable VAR relative to the number of observations in the post-1982 sample.*

[15] *When the federal funds rate precedes the spread in the ordering, the response of the federal funds rate to a spread shock is still negative. Again, this is consistent with the Federal Reserve's tightening in response to a negative supply shock. However, for this ordering, the price puzzle remains in the 1960–79 sample.*

Thus, the implication is that the Federal Reserve responds to negative supply shocks by increasing the funds rate, but not by enough to fully offset the inflationary implications of the shock.

One interpretation of the muted price puzzle during the 1980s is that the Federal Reserve increased the weight it placed on price stability and reacted more strongly to inflationary shocks. Alternatively, it may be that the U.S. economy experienced fewer severe supply-side shocks during the 1980s, which allowed the Federal Reserve to avoid the difficult decisions associated with accepting either a run-up in prices or a larger short-run decline in output.

References

Balke, Nathan S., and Kenneth M. Emery (1994), "The Federal Funds Rate as an Indicator of Monetary Policy: Evidence from the 1980s," Federal Reserve Bank of Dallas *Economic Review*, First Quarter, 1–15.

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

Christiano, Lawrence J., Martin Eichenbaum, and Charles Evans (forthcoming), "Identification and the Effects of Monetary Policy Shocks," in M. Blejer, Z. Eckstein, Z. Hercowitz, and L. Leiderman, eds., *Financial Factors in Economic Stabilization and Growth* (Cambridge: Cambridge University Press).

——— (1994), "The Effects of Monetary Policy Shocks: Evidence from the Flow of Funds," Federal Reserve Bank of Chicago Working Paper Series, no. 94–2.

Cooley, Thomas F., and Stephen F. LeRoy (1985), "Atheoretical Macroeconometrics: A Critique," *Journal of Monetary Economics* 16 (November): 283–308.

Doan, Thomas A. (1992), *RATS User's Manual* (Evanston, Ill.: Estima).

Eichenbaum, Martin (1992), "Comments on 'Interpreting the Macroeconomic Time Series Facts: The Effects of Monetary Policy' by Christopher Sims," *European Economic Review* 36 (June): 1001–1011.

Goodfriend, Marvin (1992), "Interest Rate Policy and the Inflation Scare Problem: 1979–1992," Federal Reserve Bank of Richmond 1992 Annual Report, 7–19.

Hakkio, Craig S., and Charles S. Morris (1984), "Vector Autoregressions: A User's Guide," Federal Reserve Bank of Kansas City Research Working Paper no. 84–10.

Laurent, Robert D. (1988), "An Interest Rate-Based Indicator of Monetary Policy," Federal Reserve Bank of Chicago *Economic Perspectives*, January/February, 3–14.

McCallum, Bennett T. (1983), "A Reconsideration of Sims' Evidence Concerning Monetarism," *Economics Letters* 13 (2–3): 167–71.

Oliner, Stephen D., and Glenn D. Rudebusch (1992), "The Transmission of Monetary Policy to Small and Large Firms," manuscript, Board of Governors of the Federal Reserve System.

Runkle, David E. (1987), "Vector Autoregressions and Reality," *Journal of Business and Economic Statistics* 5 (October): 437–42.

Romer, Christina D., and David H. Romer (1989), "Does Monetary Policy Matter? A New Test in the Spirit of Friedman and Schwartz," in Olivier J. Blanchard and Stanley Fisher, eds., *NBER Macroeconomics Annual 1989* (Cambridge, Mass.: MIT Press), 121–70.

Sims, Christopher A. (1992), "Interpreting the Macroeconomic Time Series Facts: The Effects of Monetary Policy," *European Economic Review* 36 (June): 975–1000.

——— (1986), "Are Forecasting Models Usable for Policy Analysis?," Federal Reserve Bank of Minneapolis *Quarterly Review*, Winter, 2–15.

Todd, Richard M. (1990), "Vector Autoregression Evidence on Monetarism: Another Look at the Robustness Debate," Federal Reserve Bank of Minneapolis *Quarterly Review*, Spring, 19–37.

Zsolt Becsi
Economist
Federal Reserve Bank of Dallas*

Indicators of the General Price Level and Inflation

Price stability has emerged as a key long-term monetary policy goal. In a nutshell, a price stabilization policy seeks to minimize the disruptive effects of aggregate price movements and price uncertainty on economic decisions. This means that a long-run aggregate price level or some low inflation rate is targeted.[1] Consequently, the Federal Reserve's job is to monitor the value of currency. The problem for the Fed practitioner is how to monitor general price movements given that aggregate price data are noisy and imperfectly measured and that there are competing price measures. This article examines whether existing price indexes tell a consistent story about the general price level and its inflation rate.

To answer this question, the coverage of the most-watched price indexes is briefly reviewed.[2] It is shown how these indexes differ and that they may provide conflicting information. Does it then matter which price/inflation index is monitored? In response, alternative notions of the theoretically appropriate price index are discussed. Next, the time series properties of the price indexes are analyzed and compared. For instance, whether a time series is stationary is evaluated because this can determine if the effects of shocks on the series are temporary and will eventually die out. Cointegration tests reveal whether the price indexes have any stable, long-term relationships: cointegrated series have a common trend.

Popular price indexes such as the consumer price index (CPI) and the implicit price deflator for gross domestic product (PGDP) capture prices of currently produced final goods and services, while the producer price index (PPI) captures final goods prices at an early distribution stage. These indexes are found to have a stable long-run relationship (or, are cointegrated) with PGDP. For growth rates, the story is similar. Inflation rates tend to be nonstationary, although the evidence can be ambiguous, and except for the PPI, the inflation rates of the different series have stable long-run relationships with one another. Thus, monitoring any particular price series or inflation rate will capture long-run movements in the other series.

Some economists have argued that central banks should monitor a very general price level that would not only include final goods and services prices but also the prices of assets, intermediate goods, and services. How good are the above price indexes as indicators of a comprehensive aggregate price measure? Price indexes for intermediate goods, the producer price index for all commodities (PPIT), and asset prices, such as the Standard & Poor 500 stock index (S&P 500) and median housing price, are weakly cointegrated in levels with final-goods price indexes (such as the PGDP). In addition, the inflation rates of the

*I have benefited greatly from discussions with and helpful comments from Nathan Balke, Ken Emery, Evan Koenig, and Mark Wynne. I would also like to thank Chih-Ping Chang for superb research assistance. Of course, any errors are my own. *This article was completed while Zsolt Becsi was on the staff of the Federal Reserve Bank of Dallas. He is now a senior economist at the Federal Reserve Bank of Atlanta.*

[1] *See Ireland (1993) and Balke and Emery (1994) for an overview of the issues. Balke and Emery distinguish between long-run price level targeting, or strong price stability, and low (or zero) inflation rate targeting, or weak price stability.*

[2] *See the surveys of Brauer and Wu (1991), Davis (1991), Carlson (1989), Webb and Willemse (1989), and Wynne and Sigalla (1993) on the coverage of the producer price index of finished goods, the consumer price index, and the implicit price deflator for gross domestic product, as well as other indexes. The PPIT referred to below is the PPI for all commodities (seasonally adjusted), not the popular PPI, which only covers finished goods.*

PPIT and the median housing price (but perhaps not the S&P 500) are weakly cointegrated with PGDP inflation. Thus, monitoring final goods prices or inflation indexes may be adequate because shocks to the other series will be reflected in these indexes. However, to the extent the PGDP does not fully capture information about movements of the other price indexes, a more comprehensive price level index may be called for.

Coverage of major price indexes

In theory, the aggregate price level represents the average level of all prices in an economy at a point in time. However, existing price indexes measure the price level for a group of goods and services that is more or less broad. The CPI, PPI, and PGDP are the price indexes that receive the most attention. Of these, the CPI is the most widely watched measure of purchasing power. Not only is the CPI timely because monthly data exist, it is influential because it is used to index federal programs such as Social Security, income tax brackets, and wage contract negotiations.

The PGDP covers the prices of all goods and services included in GDP, so it tends to be the most comprehensive.[3] The CPI covers just the prices of consumption goods and services paid by urban consumers. The CPI includes imported consumption goods, while the GDP deflator covers only domestically produced goods. Thus, the PGDP is less sensitive to factors such as oil price shocks.

The PPIT measures the prices producers charge for goods used to produce other goods (crude materials, commodities, and semi-finished and finished goods). PPIT measures prices of goods at an earlier production and distribution stage than the CPI and PGDP; however, the PPIT does not cover services. The same can be said for the popular PPI, except that it only covers the wholesale prices of final (or finished) goods.

Which index comes closest to measuring the aggregate price level? Obviously, the PPI is too narrow by itself to reflect the general price level. The PPIT is broader and may contain useful information beyond that embodied in the final-goods price indexes. It seems that PGDP is closest to an aggregate measure of the general price level because it has the broadest coverage. However, PGDP has two disadvantages relative to the CPI. It is only measured quarterly and uses current quantity weights that make it an impure measure of price changes.[4]

What price index does theory suggest? Davis (1991) writes that "the CPI can reasonably be considered 'the' measure of inflation, since it is the only one specifically designed to measure the purchasing power of money for the average final consumer of goods and services." However, the price index for measuring the purchasing power of a unit of currency could easily be defined on a broader collection of goods and services than even the PGDP. For instance, the transactions approach of the quantity theory of money, as stated by Fisher (1920) or Friedman and Schwartz (1982), proposes an even broader price index that reflects all money-based market transactions within a time period.[5] The transactions approach suggests targeting a comprehensive price index with the broadest possible coverage of current final goods and services as well as assets and intermediate goods prices.[6] Fisher's quantity equation evokes the long-run link between monetary instruments and objectives: $MV = PT$. Here, M is the money stock, T is the total number of transactions, P is the aggregate price level, and V is the velocity of transactions. Thus, given the velocity and number of transactions, money influences the aggregate

[3] *Note that up to 80 percent of the PGDP is built up from components of the CPI and PPI as well as other indexes. See footnote 2 for references.*

[4] *Because the PPI and CPI are based on a fixed market basket of goods, they reflect price changes only. However, the fixed-base-year quantity weights come from surveys taken at ten-year intervals that become less and less relevant over time. Weights in the PGDP reflect the importance of the various items in the current market basket of goods. Thus, changes to the index reflect changes in the composition of GDP as well as prices. While a fixed-weight GDP price deflator exists, the series covers too short a time span to be useful.*

[5] *See Wynne and Sigalla (1993) and Santoni and Moehring (1994) for further discussion and references.*

[6] *This includes money transactions in the underground economy. For a survey of studies estimating the size of the underground economy, see Bendelac and Clair (1993).*

price level one-to-one in the long run.

As a practical matter, the quantity equation has been narrowed by substituting final goods output for total transactions, which are difficult to observe. In this income-based approach, the price index is the aggregate price level for final goods and services (or PGDP), and velocity is defined as the velocity of final goods and services transactions. It requires strong assumptions to presuppose that the price of final goods captures all movements of the aggregate price of money transactions. In essence, it must be true that within a period, total transactions are a constant multiple (or cointegrated in the long run) to the output of final goods; and similarly for the velocity of total transactions and output. If this is not the case, final-goods price indexes may imperfectly reflect sustained general price changes that are due to monetary policy and may make aggregate price targeting more difficult.

Definition of inflation

The rate of inflation is defined as the percentage rate of change in a price index from one period to the next. Policymakers are interested in sustained changes of the economy's aggregate price level. This is because the trend, or average rate, of money growth (relative to real potential output growth) is the main determinant of these changes. When central bankers speak of inflation, they are concerned with sustained aggregate price changes or price movements that are primarily determined by monetary policy.

However, many other factors can affect price statistics. For example, short-term price shocks to a small number of goods may cause one-time jumps in the price level that are not sustained. Also, and perhaps simultaneously, as sectors allocate resources, the relative price changes of some goods over time may be picked up as persistent effects on many price indexes. Because price data are very noisy, the public and policymakers may have trouble distinguishing all of the different sources of change. Thus, they may overreact to short-term movements in the published price indexes and make suboptimal economic choices.

To get a reliable measure of sustained aggregate price changes, policymakers try to sift out the noise from the aggregate price changes they can influence. The crudest attempts to factor out short-term variability were the core rates of inflation. These indexes subtract food and energy indexes from the PPI and CPI. Except for the volatility of the food and energy markets in the 1970s, there really is no basis for throwing away the information that may be contained in these series.[7]

Recently, more sophisticated attempts to filter out noise have been studies that estimate common factors in the subindexes of the major aggregate inflation indexes. These studies use the inflation rates of components of the price indexes as separate but noisy observations on common price changes. Bryan and Cecchetti (1993b) use subindexes of the CPI, while Dow (1993) uses components of both the PPI and CPI to estimate common factors. The time-varying common factor may be attributed to monetary policy.

Although this second approach comes closer to what policymakers want measured, it may not adequately capture sustained price movements. Capturing these movements requires focusing on long-term price and inflation series movements and good knowledge of the series' dynamic properties over long periods. A common trend (or common long-run factor) across different price series—if one exists—would capture the long-term price growth that is of interest to policymakers. The next section explores whether the PPI, CPI, and PGDP have common trends despite differences in coverage. It also explores the links between the PGDP and price indexes of assets and intermediate goods to determine whether final goods prices are satisfactory indicators of the general price level as suggested by the transactions version of the quantity equation. For instance, the information contained in the dynamics of final-goods price indexes may be insufficient to capture price movements of intermediate goods, financial assets (such as equities and bonds), and real assets

[7] *Golub (1993) cites studies in the early 1980s that find relative price variability due to food and energy prices caused inflation in the 1970s. However, there do not seem to be similar studies for the 1980s. Bryan and Cecchetti's (1993a) median estimator approach is a more subtle attempt to derive a core rate of inflation that does not exclude any particular sector as a source of temporary variability.*

Table 1a
Contemporaneous Correlation Among Quarterly Price Indexes

	PGDP	*CPI*	*PPI*	*PPIT*	*SP500*	*HMP*
PGDP		.998	.995	.991	.912	.995
CPI			.997	.993	.898	.993
PPI					.872	.989
PPIT					.856	.985
SP500						.877
HMP						

(such as the stock of residential and commercial real estate, land, art, and so on).

Empirical properties of existing price series

This section examines the time series properties of price statistics to see how well they capture sustained or money-induced price movements. As Table 1a shows, the quarterly price series for the *CPI*, *PPI*, and *PGDP* exhibit high correlations between 1947:2 and 1994:1; the same is true for inflation rate indexes between 1947:3 and 1994:1, as can be seen in Table 1b. This suggests that any series may be a reliable indicator of movements in the others. However, targeting any one of these series will not necessarily produce equivalent movements in those that remain. Simple correlations do not necessarily imply a stable long-run relationship that a central bank can exploit. In particular, such correlations do not distinguish between persistent or temporary movements and their sources.

Since the transactions approach of the quantity equation suggests that the general price level may differ from final goods prices, this section also looks at the prices of assets and intermediate goods to see whether final goods prices (as represented by the *PGDP*) capture general price level movements. Convenient asset price indexes are the median sales price of housing (available starting in 1963:1) and the S&P 500 index. Tables 1a and 1b include correlations of a broad final-goods price index, the *PGDP*, with *PPIT*, the *S&P 500* index, and the median home price (*HMP*). Despite the broader coverage of the *PPIT*, correlations of the *PPIT* in levels and differences are similar to that of the *PPI*. Both asset price series are also highly correlated with the final-goods price indexes in levels. However, they are only weakly correlated in differences; the *S&P 500* index shows a negative correlation and the housing price exhibits a positive correlation with *PGDP* inflation. Also, as Figures 1a and 1b show, the price series appear to trend together even though movements of these indexes can deviate over short horizons. Below, we investigate to what extent there exist common long-term trends (which are presumably due to common factors such as monetary policy).

Before looking at common long-term trends, this section looks first at the properties of the different time series in isolation. In particular, tests reveal whether a series is nonstationary or stationary. These tests estimate the persistence of prices and their growth rates. Stationarity implies that the effects of shocks are temporary and will eventually die out. Whether a series is stationary or not also gives a measure of the price uncertainty facing economic decisionmakers. Finally, this section examines whether pairs of different price series are cointegrated—that is, whether they have a stable long-run relationship. Cointegration tests tell us whether it matters which price index we monitor and if a stable long-run relationship exists between the various indexes that could be capitalized on by policy. Shared stochastic trends, or cointegration, reveal shared underlying pro-

Table 1b
Contemporaneous Correlation Among Quarterly Inflation Indexes

	ΔPGDP	*ΔCPI*	*ΔPPI*	*ΔPPIT*	*ΔSP500*	*ΔHMP*
PGDP		.81	.74	.704	–.081	.205
CPI			.862	.804	–.213	.017
PPI					–.236	.117
PPIT					–.232	.181
SP500						–.07
HMP						

cesses and imply that permanent shocks to the trend of one series will be transmitted to the trend of the other series.

Stationarity tests determine the importance of a deterministic trend relative to a stochastic trend within a variable's long-run dynamics. Loosely, a (trend) stationary time series contains a deterministic trend but not a stochastic trend.[8] This means that shocks to the series cause temporary fluctuations. Because the series always reverts to its trend, there is no long-run uncertainty about the series. Therefore, one can forecast the series' long-run component with complete certainty or zero forecast variance. In contrast, a nonstationary series contains a stochastic trend. Instead of trend-reverting fluctuations, shocks cause permanent changes to the series. As a result, the series never completely returns to its original trend. For such a series, there is a great deal of uncertainty about its long-run behavior that increases with time. This uncertainty is reflected by an increasing forecast variance. Strong versions of price stability attempt to eliminate this uncertainty by making price levels and inflation rates stationary.

To determine whether a series is stationary or not requires a battery of tests.[9] The first of these is the augmented Dickey–Fuller (ADF), which tests the null hypothesis that the variable is nonstationary (or difference stationary). Additional information is available from the KPSS test by Kwiatkowski et al. (1992), which directly tests the null hypothesis that the series is stationary.[10] The ADF and KPSS tests may have difficulty in establishing whether a series is stationary or not. This is because it takes a very large sample to distinguish between stationarity and nonstationarity, and the postwar quarterly sample used in this article may be too small. Finally, the variance-ratio test of Cochrane (1988) compares the size of the permanent (or stochastic trend) component relative with the temporary (trend-reverting) component in a series by calculating the ratio of the components' variances. The variance ratio tends toward zero (or one) the smaller (or larger) the stochastic

8 For a trend stationary process, the deterministic trend is linear and can be written as (a + bt). A mean stationary process is a trend stationary process with b = 0. The important issue of nonlinear deterministic trends and structural breaks will not be explored here. For more on this, see, for instance, Balke (1991) and Hamilton (1994).

9 The tests are further described in the summary tables. In the spirit of Nelson and Plosser (1982), the ADF and KPSS (nonstationarity) tests define difference stationary by focusing on unit-root processes. In the spirit of Beveridge and Nelson (1981), McCallum (1993) argues for allowing more general nonstationary (or difference stationary) processes and that time series combine trend and difference stationary components with one or the other dominating. The variance ratio below agrees with this view.

10 Difference stationary processes can be either trend or mean stationary.

Figure 1a
Movements of *PGDP*, *PPI*, and *CPI*

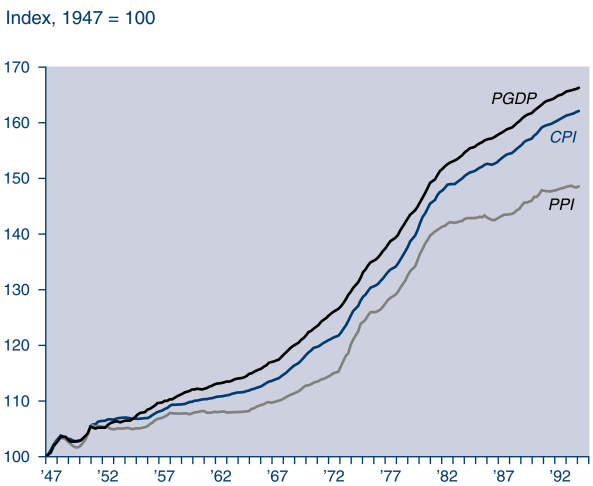
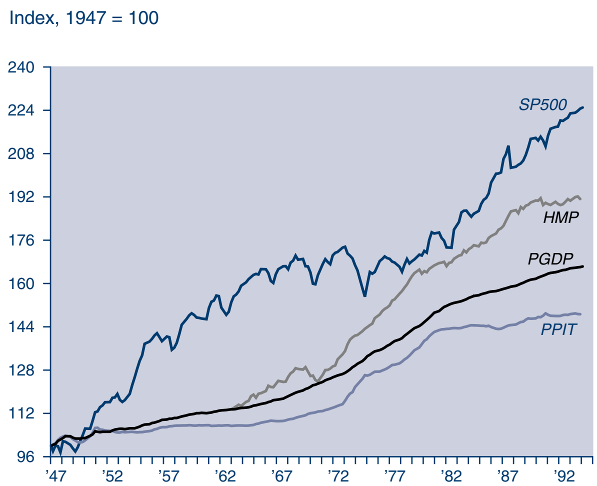


Figure 1b
Movements of *PGDP*, *PPIT*, *SP500*, and *HMP*



trend component in the series. However, if there is positive (or negative) serial correlation, the variance ratio will have an upward (or downward) bias from one. Differences in this ratio across price series indicate possible heterogeneous trends resulting from the different coverage of the indexes.

Results of the ADF and KPSS stationarity tests are summarized in Tables 2 and 3. The tests agree that *PGDP*, *CPI*, and *PPI* are nonstationary in levels. However, the tests on the stationarity of the series' inflation rates disagree and are sensitive to the lag specification.[11] The ADF test fails to reject nonstationarity for the growth rates of all three series. The results for the KPSS test tend to be less crisp. For the *PGDP* and *CPI* inflation rates, stationarity is (weakly) rejected, while stationarity fails to be rejected for the *PPI*. Thus, the ADF and KPSS give weak or conflicting results for the inflation series (especially for the *PPI*) but indicate that the levels of the price series are nonstationary.

The variance ratios of the price level series in Table 4 are large and growing, which suggests a large permanent component with positive serial correlation. The results agree with the above finding of nonstationary price level series. Thus, unexpected shocks to the price level cause the series to diverge from its initial path. By contrast, the variance ratios for the inflation rate series are small and falling. This is evidence that the inflation rates are stationary or that the permanent (or stochastic trend) component is dominated by the temporary (or trend-reverting) component over longer horizons. A large temporary component may distort the finite sample critical values of the ADF and KPSS tests and may explain why the test results are ambiguous (Schwert 1987). Furthermore, as the variance ratios differ from one index to another, the size of the stochastic trend component relative to the temporary component differs too. At long horizons, the *CPI* and *PPI* inflation rates have similar variance ratios, while *PGDP* tends to have the lowest. This suggests that shocks to the *CPI* and *PPI* inflation rates are more persistent than shocks to the growth rates of *PGDP*.

What about the price indexes for intermediate goods and assets? Despite the broader coverage of the *PPIT*, ADF and KPSS tests reveal that the *PPIT* in levels behaves similar to the *PPI* in terms of nonstationarity. However, in contrast to the *PPI*, these tests suggest that the *PPIT* inflation rate is stationary. This conclusion is supported by the variance ratios (which in levels and differences

[11] *The discussion of the ADF and KPSS tests considers the optimal lag length that comes closest to eliminating serial correlation for the ADF test statistic. The results for other lag lengths are in Tables 2 and 3.*

Table 2
Augmented Dickey–Fuller Unit-Root Test Statistics[1]

Variable	Sample	With trend			Without trend		
		Lags = 4	Lags = 8	Lags = 12	Lags = 4	Lags = 8	Lags = 12
PGDP	1947:1–94:1	–2.365 (.072)	–1.941 (.414)	–2.153 (.048)	1.114 (.088)	–.276 (.472)	–.701 (.105)
CPI	1947:1–94:1	–2.197 (.001)	–1.448 (.151)	–1.875 (.003)	1.337 (.002)	.564 (.228)	–.193 (.009)
PPI	1947:2–94:1	–1.965 (.04)	–1.5 (.101)	–1.877 (.010)	.528 (.044)	–.031 (.144)	–.189 (.023)
PPIT	1948:2–93:4	–1.904 (.005)	–1.424 (.019)	–1.756 (.0003)	.376 (.008)	–.10 (.032)	–.365 (.001)
SP500	1947:1–94:1	–2.153 (.088)	–1.716 (.002)	–1.705 (.007)	–.899 (.075)	–.88 (.004)	–.805 (.012)
HMP	1963:1–93:4	–1.21 (.169)	–1.628 (.243)	–1.398 (.147)	–1.109 (.226)	–.822 (.435)	–0.959 (.194)
Δ*PGDP*	1947:2–94:1	**–4.122*** (.08)**	–1.453 (.2)	–1.229 (.03)	**–3.729*** (.1)**	–1.86 (.29)	–1.755 (.04)
Δ*CPI*	1947:2–94:1	**–3.908** (.003)**	–2.411 (.19)	–2.166 (.14)	**–3.494*** (.005)**	2.471 (.23)	–1.865 (.02)
Δ*PPI*	1947:3–94:1	**–4.134*** (.02)**	–2.705 (.07)	–2.166 (.14)	**–3.975*** (.3)**	**–2.798* (.099)**	–2.054 (.18)
Δ*PPIT*	1948:3–93:4	**–4.388*** (.002)**	**–2.80*** (.004)**	**–2.034*** (.8)**	**–4.297*** (.004)**	**–2.893*** (.007)**	**–2.072*** (.8)**
Δ*SP500*	1947:2–94:1	**–6.73*** (.075)**	**–5.23*** (.004)**	**–4.32*** (.013)**	**–6.749*** (.11)**	**–5.223*** (.007)**	**–4.33*** (.022)**
Δ*HMP*	1963:2–93:4	**–3.926** (.09)**	–2.823 (.52)	–2.522 (.22)	**–3.865*** (.13)**	**–2.72*** (.59)**	–2.472 (.32)

* Significant at the .10 level.
** Significant at the .05 level.
*** Significant at the .01 level.

Significance denotes that the null hypothesis of nonstationarity is rejected. Numbers in parentheses are the significance level determined by the Ljung-Box Q statistic for whether serial correlation is eliminated for a given lag length.

[1] The ADF test is determined by the regression: $y_t = \alpha + \beta_t + \rho y_{t-1} + \sum_{j=1}^{n} \gamma_j \Delta y_{t-j} + e_t$,

where y_t is the variable in period t, $\Delta y_{t-j} = y_{t-j} - y_{t-j-1}$, and n is the lag length. The null hypothesis that y_t is nonstationary is rejected when $\hat{\rho}$ differs significantly from one. The critical values are found in Fuller (1976) and Hamilton (1994).

Table 3
KPSS Unit-Root Test Statistics[1]

Variable	Sample	With trend Lags = 4	With trend Lags = 8	With trend Lags = 12	Without trend Lags = 4	Without trend Lags = 8	Without trend Lags = 12
PGDP	1947:1–94:1	.47***	.258***	.183**	2.128***	1.159***	.811***
CPI	1947:1–94:1	.485***	.266***	.188**	2.092***	1.14***	.8***
PPI	1947:2–94:1	.42***	.231***	.165**	2.05***	1.114***	.780***
PPIT	1948:2–93:4	.40***	.221***	.157**	2.03***	1.104***	.722***
SP500	1947:1–94:1	.254***	.147**	**.108**	1.959***	1.09***	.781***
HMP	1963:1–93:4	.14*	**.089**	**.077**	1.447***	.8***	.571**
Δ*PGDP*	1947:2–94:1	.244***	.174**	.125*	.574**	.383*	**.272**
Δ*CPI*	1947:2–94:1	.181**	.141*	**.116**	.527***	.377*	**.297**
Δ*PPI*	1947:3–94:1	.172***	.138*	**.112**	**.3**	**.233**	**.186**
Δ*PPIT*	1948:3–93:4	.163**	.143*	**.115**	**.242**	**.206**	**.165**
Δ*SP500*	1947:2–94:1	**.098**	.411***	.236***	**.098**	.411*	**.237**
Δ*HMP*	1963:2–93:4	.127*	**.108**	**.107**	**.166**	**.140**	**.137**

* Significant at the .10 level.
** Significant at the .05 level.
*** Significant at the .01 level.

Significance means that the null hypothesis of stationarity is rejected.

[1] The test statistics are derived by computing the test statistic $\left(1/T^2\right)\sum_{t=1}^{T}\left(S_t^2/\sigma^2(l)\right)$, where T is the sample size, $S_t = \sum_{i=1}^{T} e_t$ and e_i is the residual from a regression of the variable in question, y_t, on an intercept and a time trend. Also, $\sigma^2(l)$ is a consistent estimator of the long-run variance of y_t and is constructed as in Kwiatkowski et al. (1992). Critical values for the above test statistics can be found in Kwiatkowski et al. (1992).

appear to be very similar to those of the *PPI*). ADF and KPSS tests reveal that the *S&P 500* is nonstationary in levels. The evidence is mixed for the growth rates. ADF and KPSS tests seem to imply stationarity, and variance ratios show that the temporary component tends to dominate in differences, which is evidence for stationarity. Finally, the ADF and KPSS tests indicate that median home prices are nonstationary in levels and possibly stationary in differences.

Finally, cointegration tests of various price series are depicted in Tables 5a and 5b. Essentially, the cointegration test is a test for common trends and indicates whether the same processes underlie the different price indexes, even if the indexes cover different goods and services. The method proposed by Johansen (1988, 1991) and Johansen and Juselius (1990) was chosen to determine whether various price indexes share a common stochastic trend.[12] The evidence is that the *PGDP* is cointegrated with the *PPI* and weakly cointegrated with the *CPI*, but the *PPI* and *CPI* are not cointegrated. Because coverage of the *PGDP* comprises components of both *CPI* and *PPI*, it is not surprising that *PGDP* shares a trend with these indexes. Surprisingly, *PPI* and *CPI* do not share a trend. Transitivity would imply that

[12] *See Campbell and Perron (1991) and Gonzalo (1994) for comparisons of methods.*

Table 4
Cochrane's Variance-Ratio Statistics*

Variable	Sample	k = 4	k = 8	k = 12	k = 24	k = 36	k = 48
PGDP	1947:1–94:1	2.76	4.35	6.06	12.114	18.38	24.43
		(.046)	(1.03)	(1.76)	(4.985)	(9.26)	(14.21)
CPI	1947:1–94:1	3.18	5.13	6.69	11.73	17	21.96
		(.53)	(1.22)	(1.95)	(4.83)	(8.57)	(12.78)
PPI	1947:2–94:1	2.82	4.25	5.29	8.02	10.66	12.524
		(.48)	(1.01)	(1.54)	(3.31)	(5.39)	(7.31)
PPIT	1948:2–93:4	2.72	3.81	4.53	6.54	8.56	9.82
		(.46)	(.91)	(1.32)	(2.698)	(4.32)	(5.73)
SP500	1947:1–94:1	1.4	1.16	.93	1.16	1.29	1.27
		(.23)	(.28)	(.27)	(.48)	(.65)	(.74)
HMP	1963:1–93:4	1.3	1.64	1.84	2.07	2.88	4.01
		(.27)	(.48)	(.66)	(1.05)	(1.79)	(2.88)
Δ*PGDP*	1947:2–94:1	.393	.215	.154	.094	.075	.068
		(.066)	(.051)	(.045)	(.039)	(.038)	(.04)
Δ*CPI*	1947:2–94:1	.55	.398	.258	.134	.111	.098
		(.093)	(.095)	(.075)	(.055)	(.056)	(.057)
Δ*PPI*	1947:3–94:1	.658	.43	.276	.134	.108	.095
		(.11)	(.103)	(.081)	(.055)	(.055)	(.056)
Δ*PPIT*	1948:3–93:4	.583	.377	.223	.108	.087	.078
		(.098)	(.09)	(.065)	(.044)	(.044)	(.045)
Δ*SP500*	1947:2–94:1	.365	.187	.113	.068	.037	.024
		(.062)	(.045)	(.033)	(.028)	(.019)	(.014)
Δ*HMP*	1963:2–93:4	.269	.146	.086	.048	.026	.019
		(.056)	(.043)	(.031)	(.024)	(.017)	(.014)

* Cochrane's (1988) variance-ratio statistics for the difference horizon k are estimated as the following ratio of variances: $\frac{Var(y_{t+k}-y_t)}{kVar(y_{t+1}-y_t)}$. Bartlett standard errors are given in the parentheses and are computed as $(4k/3T)^{1/2}$, where T is the sample size.

the *CPI* is cointegrated with *PPI*. This may, in fact, be the case; however, cointegration tests may have insufficient strength to yield consistency. An alternative explanation is that *PGDP* and *CPI* are not cointegrated. In addition, the inflation rates of the *PGDP* and the *CPI* are cointegrated. However, the cointegration tests with the *PPI* inflation indicate that there are as many cointegrating relationships as there are series included in the test regression. This result implies that the inflation series are stationary. Thus, the tests with *PPI* inflation are inconclusive and contradict the findings from ADF and KPSS tests (although they are consistent with the variance ratios).

What about the price indexes for intermediate goods and assets? The *PPIT* is cointegrated with *PGDP* in levels. Despite the earlier finding of stationarity, *PPIT* inflation is cointegrated with *PGDP* inflation. Thus, monitoring any of the final-goods price or inflation indexes (as represented by the *PGDP*) is sufficient because those price or inflation indexes for total intermediate goods will not

Table 5a
Cointegration Test Statistics for Price Levels

		With trend				Without trend			
Variables	Sample size	Eigen-values	Null hypothesis	λ-max test	Trace test	Eigen-values	Null hypothesis	λ-max test	Trace test
CPI, *PPI*	1947:2–94:1	.027	*r* = 0:	4.97	5.85	**.064**	***r* = 0:**	**11.95**	**14.68**
	T = 188	.005	*r* = 1:	.89	.89	**.015**	***r* = 1:**	**2.73**	**2.73**
						0			
PGDP, *CPI*	1947:1–94:1	**.071**	***r* = 0:**	**13.42****	**14.05****	.112	*r* = 0:	21.51***	24.06***
	T = 189	**.003**	***r* = 1:**	**.63**	**.63**	.014	*r* = 1:	2.56	2.56
						0			
PGDP, *PPI*	1947:2–94:1	.032	*r* = 0:	5.84	8.57	**.101**	***r* = 0:**	**19.35*****	**23.39*****
	T = 188	.015	*r* = 1:	2.73	2.73	**.022**	***r* = 1:**	**4.05**	**4.05**
						0			
CPI, *PPIT*	1948:2–93:4	.03	*r* = 0:	5.45	5.95	**.067**	***r* = 0:**	**12.42**	**14.51**
	T = 183	.003	*r* = 1:	.496	.496	**.012**	***r* = 1:**	**2.09**	**2.09**
						0			
PGDP, *PPIT*	1948:2–93:4	.02	*r* = 0:	3.69	6.45	**.077**	***r* = 0:**	**14.52****	**18.01****
	T = 183	.015	*r* = 1:	2.76	2.76	**.019**	***r* = 1:**	**3.49**	**3.49**
						0			
PGDP, *HMP*	1963:1–93:4	.107	*r* = 0:	13.52**	15.62***	**.109**	***r* = 0:**	**13.85****	**19.98****
	T = 124	.015	*r* = 1:	2.16	2.16	**.05**	***r* = 1:**	**6.13**	**6.13**
						0			
PGDP, *SP500*	1947:1–94:1	.049	*r* = 0:	9.31	11.38	**.141**	***r* = 0:**	**28.03*****	**33.22*****
	T = 189	.011	*r* = 1:	2.07	2.07	**.028**	***r* = 1:**	**5.19**	**5.19**
						0			

** Significant at the .10 level.
*** Significant at the .05 level.

Critical values are from Johansen and Juselius (1990) Table A1 for the model estimated with a trend and Table A3 for the model without. After looking at the graph of the price level series, it was determined that the vector error-correction model should be estimated with a trend. To test whether the null hypothesis of a trend or the alternative of no trend fit the data better, a likelihood-ratio test was performed. The test statistic, $-T\sum_{i=r+1}^{p} \ln \frac{1-\lambda_i^{trend}}{1-\lambda_i^{no\ trend}}$, is distributed $x^2(p-r)$ where $p = 2$ is the number of variables, r is the number of cointegrating vectors, and the eigenvalues, λ_i, are arranged in descending order (or, $\lambda_1 > \lambda_2$). Note that this test statistic is conditioned on the r found to be significant in the model with a trend. Finally, the bold-faced statistics in the table indicate which model passes the likelihood-ratio test.

move independently in the long run. This cointegration result suggests a stable long-run link between final-goods price indexes and the general price level. Also, the *S&P 500* and *PGDP* are cointegrated in levels. However, the relationship is inconclusive for the growth rates of the *S&P 500* and *PGDP*, which is in line with the univariate evidence that the *S&P 500* is stationary in differences. The median home price series is also cointegrated with *PGDP* in levels and weakly cointegrated in growth rates (which contradicts the univariate evidence of stationarity). Thus, there is evidence that asset prices share common trends with final goods prices. Since the general price level may comprise intermediate goods and asset prices, and *PGDP* tends to share common trends

Table 5b

Cointegration Test Statistics for Inflation Rates

Variables	Sample size	With trend: Eigen-values	Null hypothesis	λ-max test	Trace test	Without trend: Eigen-values	Null hypothesis	λ-max test	Trace test
ΔCPI, ΔPPI	1947:3–94:1	.109	r = 0:	20.99***	31.41***	**.110**	**r = 0:**	**21.03*** **	**31.52*** **
	T = 187	.056	r = 1:	10.43***	10.43***	**.056**	**r = 1:**	**10.49*** **	**10.49*** **
						0			
$\Delta PGDP$, ΔCPI	1947:2–94:1	.165	r = 0:	32.54***	39.43***	**.165**	**r = 0:**	**32.64*** **	**39.53*** **
	T = 188	.037	r = 1:	6.89***	6.89***	**.037**	**r = 1:**	**6.89**	**6.89**
						0			
$\Delta PGDP$, ΔPPI	1947:3–94:1	.142	r = 0:	27.72***	37.2***	**.142**	**r = 0:**	**27.73*** **	**37.74*** **
	T = 187	.054	r = 1:	10***	10***	**.054**	**r = 1:**	**10*** **	**10*** **
						0			
ΔCPI, $\Delta PPIT$	1948:3–93:4	.102	r = 0:	19.46***	26.1***	**.103**	**r = 0:**	**19.54*** **	**26.2*** **
	T = 182	.036	r = 1:	6.65***	6.45***	**.036**	**r = 1:**	**6.65**	**6.65**
						0			
$\Delta PGDP$, $\Delta PPIT$	1948:3–93:4	.152	r = 0:	29.69***	37.04***	**.152**	**r = 0:**	**29.71*** **	**37.08*** **
	T = 182	.04	r = 1:	7.36***	7.36***	**.04**	**r = 1:**	**7.38**	**7.38**
						0			
$\Delta PGDP$, ΔHMP	1963:2–93:4	.182	r = 0:	23.93***	27.51***	**.182**	**r = 0:**	**23.93*** **	**27.51*** **
	T = 123	.03	r = 1:	3.58**	3.58**	**.03**	**r = 1:**	**3.58**	**3.58**
						0			
$\Delta PGDP$, $\Delta SP500$	1947:2–94:1	.218	r = 0:	45.19***	58.6***	**.218**	**r = 0:**	**45.2*** **	**58.71*** **
	T = 188	.07	r = 1:	13.4***	13.4***	**.07**	**r = 1:**	**13.5*** **	**13.5*** **
						0			

** Significant at the .10 level.
*** Significant at the .05 level.

Critical values are from Johansen and Juselius (1990) Table A1 for the model with a deterministic trend and Table A3 for the model without a trend. After looking at the graph of the inflation rate series, it was determined that the vector error correction model should be estimated without a linear trend. To test whether the null hypothesis of no trend or the trend alternative fit the data better, a likelihood-ratio test was performed. The test statistic, $-T\sum_{i=r+1}^{p} ln \frac{1-\lambda_i^{no\ trend}}{1-\lambda_i^{trend}}$, is distributed $x^2(p-r)$, where $p = 2$ is the number of variables, r is the number of cointegrating vectors, and the eigenvalues, λ_i, are arranged in descending order (or, $\lambda_1 > \lambda_2$). Note that this test statistic is conditioned on the r found to be significant in the model without a trend. Finally, the bold-faced statistics in the table indicate which model passes the likelihood-ratio test.

with some indexes of assets and intermediate goods prices, one can conclude that final goods prices are valid indicators for the general price level. However, this conclusion does not necessarily hold true for the inflation rates of intermediate goods and assets, which tend to exhibit weak or mixed evidence of stationarity and so render the cointegration tests inconclusive.

Conclusion

Past measures of sustained movements in the general price level were based on popular price indexes such as the CPI, PPI, and PGDP. This article extends the search for a general price level measure and money-induced (or sustained) price movements beyond the final goods and

List of Variables

PGDP	Implicit Price Deflator for GDP (seasonally adjusted and in logarithms)
CPI	Consumer Price Index (seasonally adjusted and in logarithms)
PPI	Producer Price Index for Finished Goods (seasonally adjusted and in logarithms)
PPIT	Producer Price Index for All Intermediate Goods (seasonally adjusted and in logarithms)
SP500	Standard & Poor's 500 Index (in logarithms)
HMP	Median Home Price (seasonally adjusted and in logarithms)
Δ	Difference operator

services prices covered by these popular price indexes. Theory suggests expanding the coverage of the popular indexes by adding information contained in asset prices and intermediate goods prices. According to some, such an expansion is necessary for a theoretically satisfactory measure of the aggregate price level and aggregate inflation.

To determine whether price indexes for final goods and services and price indexes for intermediate goods and assets provide similar information, this article investigates the time series characteristics of the above-mentioned popular price indexes, PPIT, and asset price series such as the S&P 500 and the median housing sales index. Because monetary authorities are particularly interested in sustained price changes, I focused on the long-run characteristics of these series. Examining whether the price series and their growth rates are stationary gives an idea of how close we are to price stability. More important is whether the different price series or their growth rates are cointegrated. If so, shocks to the trend of one series will be transmitted to the trends of the others. In other words, cointegration implies a stable long-run relationship between the series—a relationship that simplifies monitoring of the general price level and that a policymaker may exploit. Given cointegration, a specific price index or inflation rate can serve as an effective indicator for the other series. Otherwise, auxiliary information or a more general price measure will be necessary.

Evidence suggests that the different price level series are nonstationary, but evidence is weak or conflicting on whether their growth rates are also nonstationary. Since nonstationarity implies forecast uncertainty and potentially inefficient decision-making, achieving price level stability appears to be a more distant goal than inflation rate stability. While data from the 1980s may reveal that our price stabilizing performance has improved, more work has to be done before we can tell for sure. Tests reveal that the PGDP tends to be cointegrated with a wide range of price indexes. Thus, the PGDP is an effective indicator of the general price level. However, because there is mixed evidence that some inflation series are stationary, evidence is weak that other inflation rates are cointegrated with PGDP inflation. Thus, it is unclear whether final-goods price inflation is a useful indicator for monetary policy decisions. Future work may investigate the information contained in asset price inflation that is not contained in the PGDP and other indexes of final goods prices.

References

Balke, N. S. (1991), "Modeling Trends in Macroeconomic Time Series," Federal Reserve Bank of Dallas *Economic Review*, May, 19–33.

Balke, N. S., and K. M. Emery (1994), "The Algebra of Price Stability," *Journal of Macroeconomics* 16 (Winter): 77–97.

Bendelac, J., and P.-M. Clair (1993), "Macroeconomic Measures of the Hidden Economy" (Paper presented at the 49th Congress of the I.I.P.F., Berlin, Germany).

Beveridge, S., and C. R. Nelson (1981), "A New Approach to Decomposition of the Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of Business Cycles," *Journal of Monetary Economics* 7 (January): 151–74.

Brauer, D., and L. Wu (1991), "In Brief: An Overview of Inflation Measurement," Federal Reserve Bank of New York *Quarterly Review*, Summer, 25–29.

Bryan, M. F., and S. G. Cecchetti (1993a), "Measuring Core Inflation," NBER Working Paper Series, no. 4303 (Cambridge, Mass.: National Bureau of Economic Research).

———, and ——— (1993b), "The Consumer Price Index as a Measure of Inflation," Federal Reserve Bank of Cleveland *Economic Review* 29(4): 15–24.

Campbell, J. Y., and P. Perron (1991), "Pitfalls and Opportunities: What Macroeconomists Should Know About Unit Roots," in *NBER Macroeconomics Annual* (Cambridge, Mass.: MIT Press for National Bureau of Economic Research).

Carlson, K. M. (1989), "Do Price Indexes Tell Us About Inflation? A Review of the Issues," Federal Reserve Bank of St. Louis *Review*, November/December, 12–30.

Cochrane, J. H. (1988), "How Big Is the Random Walk in GDP?" *Journal of Political Economy* 96(5): 893–920.

Davis, R. G. (1991), "Inflation: Measurement and Policy Issues," Federal Reserve Bank of New York *Quarterly Review*, Summer, 13–24.

Dow, J. P. (1993), "Measuring Inflation Using Multiple Price Indexes," Department of Economics, University of California–Riverside, March, mimeo.

Fisher, I. (1920), *The Purchasing Power of Money: Its Determination and Relation to Credit, Interest and Crises*, 2nd edition (New York: Macmillan).

Friedman, M., and A. Schwartz, (1982), *Monetary Trends in the United States and the United Kingdom: Their Relation to Income, Prices and Interest Rates, 1867–1975* (Chicago: University of Chicago Press).

Fuller, W. A. (1976), *Introduction to Economic Time Series* (New York: Wiley).

Golub, J. E. (1993), "Inflation, Inflation Uncertainty, and Relative Price Variability: A Survey," Federal Reserve Bank of Kansas City, Working Paper no. 93–15 (Kansas City, Mo., August).

Gonzalo, J. (1994), "Five Alternative Methods of Estimating Long-Run Equilibrium Relationships," *Journal of Econometrics* 60 (January/February): 203–33.

Hamilton, J. D. (1994), *Time Series Analysis* (Princeton, N.J.: Princeton University Press).

Ireland, P. N. (1993), "Price Stability Under Long-Run Monetary Targeting," Federal Reserve Bank of Richmond *Economic Quarterly*, Winter, 25–45.

Johansen, S. (1991), "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models," *Econometrica* 59: 1551–1580.

——— (1988), "Statistical Analysis of Cointegration Vectors," *Journal of Economic Dynamics and Control* 12 (June/September): 231–54.

———, and K. Juselius (1990), "Maximum Likelihood Estimation and Inference on Cointegration—With Applications to the Demand for Money," *Oxford Bulletin of Economics and Statistics* 52 (May): 169–210.

Kwiatkowski, D., P. C. B. Phillips, P. Schmidt, and Y. Shin (1992), "Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root," *Journal of Econometrics* 54 (October–December): 159–78.

McCallum, B. T. (1993), "Unit Roots in Macroeconomic Time Series: Some Critical Issues," Federal Reserve Bank of Richmond *Economic Quarterly*, Spring, 13–43.

Nelson, C. R., and C. I. Plosser (1982), "Trends Versus Random Walks in Macroeconomic Time Series: Some Evidence and Implications," *Journal of Monetary Economics* 10 (September): 139–62.

Santoni, G. J., and H. B. Moehring (1994), "Asset Returns and Measured Inflation," *Journal of Money, Credit, and Banking* 26 (May): 232–48.

Schwert , G. W. (1987), "Effects of Model Specification on Tests for Unit Roots in Macroeconomic Data," *Journal of Monetary Economics* 20 (July): 73–103.

Webb, R. E., and R. Willemse (1989), "Macroeconomic Price Indexes," Federal Reserve Bank of Richmond *Economic Review*, July/August, 22–32.

Wynne, M. A., and F. Sigalla (1993), "A Survey of Measurement Bias in Price Indexes," Federal Reserve Bank of Dallas, Research Paper no. 9340 (Dallas, October).