# The "Price Puzzle" Reconsidered \*

Michael S. Hanson<sup>\*</sup>

Department of Economics, Wesleyan University, Middletown, CT 06459

#### Abstract

A large literature has employed structural vector autoregressive (SVAR) models to investigate the empirical effects of U.S. monetary policy. Many of these models regularly produce a "price puzzle" — a rise in the aggregate price level in response to a contractionary innovation to monetary policy — unless commodity prices are included. Conventional wisdom maintains that commodity prices resolve the price puzzle because they contain information that helps the Federal Reserve forecast inflation. I examine a number of plausible alternative indicator variables and find little correlation between an ability to forecast inflation and an ability to resolve the price puzzle. Additionally, a sub-sample investigation reveals that evidence of a price puzzle is associated primarily with the 1959 – 1979 sample period, and that most indicators — including commodity prices — cannot resolve the puzzle over this period.

Key words: Structural VAR models, monetary policy reaction function, inflation forecasting, commodity prices. JEL Classification: E52, E31, C32.

<sup>\*</sup> Thanks to Robert Barsky, Allan Brunner, Kathryn Dominguez, Martin Eichenbaum, Charles Evans, Jon Faust, Richard Grossman, Phil Howrey, Erik Hurst, Michael Kiley, Lutz Kilian, Eric Leeper, Gil Skillman, Matthew Shapiro, and an anonymous referee for helpful comments and suggestions. Additional thanks to seminar participants at the University of Michigan, the Federal Reserve Board of Governors, the NBER Monetary Economics Meeting, Wesleyan University, and the Midwest Economics Association meetings. Ady Wijaya provided able research assistance. Any remaining errors are the author's sole responsibility.

<sup>\*</sup> Corresponding author.: Department of Economics, Wesleyan University, 238 Church Street, Middletown, CT 06459–0007; Telephone: + 1-860-685-4634; Fax: + 1-860-685-2301.

Email address: mshanson@wesleyan.edu (Michael S. Hanson).

# 1 Introduction

Over the past decade, a large number of researchers have estimated structural vector autoregressive (SVAR) models to establish "stylized facts" about U.S. monetary policy.<sup>1</sup> Contrary to intuition and most commonly-accepted macroeconomic theories, many studies find a protracted rise in the price level following an exogenous contractionary innovation to monetary policy. Sims (1992) first commented on this empirical anomaly, dubbed the "price puzzle" by Eichenbaum (1992).

Despite subsequent advances in SVAR modeling, the price puzzle has generally remained a problem for empirical researchers. Some authors have argued that the presence of a price puzzle should serve as an informal specification test of a VAR model: if such an anomalous result is observed, then what one has labeled as "monetary policy" probably has not been correctly identified. Proponents of this view include Zha (1997), Sims (1998), and Christiano et al. (1999). Viewed this way, understanding the price puzzle is a prerequisite for measuring the effects of monetary policy.

Sims (1992) first demonstrated that the price puzzle largely disappeared if commodity prices were included in his VAR. He proposed that commodity prices served as an "information variable," i.e. as an indicator of nascent inflation, in the Federal Reserve's policy reaction function. Failure to include a variable that signaled future inflation thus would constitute a misspecification of the VAR model.

Most subsequent research has adopted commodity prices as a necessary variable in monetary VAR models.<sup>2</sup> As this tactic has since evolved into a "conventional wisdom," many authors now make only a passing reference to the problem commodity prices are intended to resolve. Nor is an *a priori* rationale regularly provided for including commodity prices in an otherwise parsimonious VAR model. And while a VAR system often is meant to correspond with a theoretical business cycle model, most theories do not accord an explicit macroeconomic role for commodity prices.

This essay examines the empirical consistency of this "solution" to the price puzzle. While offering some intuitive appeal, how does this approach hold up

 $<sup>\</sup>overline{}^{1}$  Some authors refer to these models as "identified VARs." Whichever nomenclature is used, the concept is the same: restrictions placed upon the model allow the researcher to imbue the estimated disturbances (or some subset thereof) with a particular economic interpretation.

 $<sup>^2</sup>$  For similar reasons, commodity prices also have found their way into empirical monetary models that use other estimation techniques. See, e.g., Sims (1999) and Clarida et al. (2000).

under more extensive econometric investigation? In an earlier study, Balke and Emery (1994) reported some support for both the above "information variable" view and for commodity prices serving as proxies for "supply shocks." <sup>3</sup> My results cast some doubt on both of these proposed explanations, suggesting the need for additional research into the role of commodity prices (and other inflation indicators) in these models. Given the sensitivity of the effects of policy to this modeling decision — particularly conclusions about the effects of monetary policy upon inflation — my analysis augurs that a reconsideration of certain popular identifying assumptions may be warranted.

Specifically, in section 3 I examine the ability of various potential indicators of inflationary pressure to resolve the price puzzle. Across a fairly broad set of indicators I find at best a limited relationship between forecasting power and mitigation of the price puzzle. In section 4, I investigate the robustness of the existence and resolution of the price puzzle across sample periods. Two conclusions emerge in this section. First, the price puzzle is more pronounced in a sample period ending in October 1979. Second, the consequences of excluding any "information variable" — including commodity prices — vary over the sample periods studied.

Each of these results contrasts with the implications of a simple model of information variables in a monetary policy rule, which is developed in the next section. Further, they may lead one to ask whether commodity prices play a unique role in resolving the price puzzle, and whether they belong in a SVAR (merely) because of their forecasting ability.

#### 2 Empirical Policy Measurement and the Price Puzzle

A structural (or identified) VAR model can be written as

$$\Phi(L) X_t = \varepsilon_t , \ E\left[\varepsilon_t \varepsilon_t'\right] = I .$$
(1)

 $X_t$  is an *n*-dimensional vector of observed endogenous variables, and  $\varepsilon_t$  a vector of unobserved structural disturbances. The structural shocks are assumed to be both mutually and serially uncorrelated, with their variances normalized to unity.

One element of  $X_t$  is the policy instrument of the monetary authority, denoted as  $m_t$ . The policy instrument can be decomposed into two components: a systematic or endogenous portion — the "reaction function" — and an un-

 $<sup>^3\,</sup>$  Christiano et al. (1996b) also suggest that commodity prices may belong in the VARs to account for oil price "supply shocks."

forecastable or exogenous policy shock. That is,

$$m_t = f(X^t) + \mu_t \,, \tag{2}$$

where  $X^t$  is the history of the observed data through date t. Under certain assumptions, a VAR model allows for the identification of the monetary policy shock,  $\mu_t$ , as an element of  $\varepsilon_t$ . This separation is important for two reasons. First, as explained below, the conventional wisdom posits that misspecification of the systematic part of policy produces the price puzzle anomaly. Second, the literature typically focuses upon responses to exogenous monetary policy shocks. Of course these two concepts are inexorably linked: given a time path for the policy instrument, specifying the endogenous component implies a particular set of exogenous policy shocks and vice versa.

A variety of approaches to identification have been investigated in the VAR literature; I focus upon a common technique that uses covariance restrictions in a block recursive specification. Examples of this approach can be found in Strongin (1995), Christiano et al. (1996a,b), and Bagliano and Favero (1998), to name a few. Bernanke and Mihov (1998) compare several such models within an encompassing framework. Additional surveys of this approach include Cochrane (1994), Leeper et al. (1996), and Christiano et al. (1999). With the exception of Strongin (1995), all of these papers — and the many derived from them — include commodity prices in their estimation.<sup>4</sup>

To illustrate the price puzzle, I first consider a baseline model without commodity prices or other "indicator" variables. Models augmented by such variables will be taken up in section 3. The variables in  $X_t$  are ordered as  $[Y_t : M_t]'$ :  $Y_t$  contains real output and the aggregate price level, while  $M_t$  contains monetary instruments and/or target variables, as discussed below. Then the structural VAR in equation (1) can be expressed as

$$\begin{bmatrix} \phi_{YY0} & \phi_{YM0} \\ \phi_{MY0} & \phi_{MM0} \end{bmatrix} \begin{bmatrix} Y_t \\ M_t \end{bmatrix} = \begin{bmatrix} \Phi_{YY}(L) & \Phi_{YM}(L) \\ \Phi_{MY}(L) & \Phi_{MM}(L) \end{bmatrix} \begin{bmatrix} Y_{t-1} \\ M_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_t^Y \\ \varepsilon_t^M \end{bmatrix}. \quad (3)$$

Identification is achieved in part by assuming that monetary policy only affects the activity variables in  $Y_t$  with a one-period lag:  $\phi_{YM0} = 0$ . As these

<sup>&</sup>lt;sup>4</sup> Two alternative classes of identifying assumptions can be found in the monetary VAR literature: non-recursive models of contemporaneous restrictions, and models of long-run restrictions. The former utilize more extensive and intricate identifying assumptions, and still require commodity prices to address the price puzzle. According to Christiano et al. (1999), such models return qualitatively similar results as those studied herein. The latter tend not to generate price puzzles but their identifying assumptions have been subjected to intense scrutiny; see, e.g., Faust and Leeper (1997) and Pagan and Robertson (1998).

restrictions are only approximate descriptions of actual behavior, they are more plausible with higher frequency data.<sup>5</sup> Below I focus upon estimates from monthly data.

Additional restrictions on the policy block,  $\phi_{MM0}$ , follow from the assumption that the Federal Reserve (perfectly) targets the funds rate, making the effective supply of reserves (perfectly) elastic: any change in demand for reserves would be accommodated by the Fed to return the funds rate to its targeted level. As in Bernanke and Blinder (1992) and much of the subsequent monetary VAR research, the policy instrument  $m_t$  is the Federal funds rate. The monetary policy shock is identified by placing the funds rate first in the policy block, followed by nonborrowed reserves and total reserves.<sup>6</sup>  $\phi_{YY0}$  and  $\phi_{MY0}$ , the blocks that capture the effects of non-policy shocks in equation (3), are left unrestricted.<sup>7</sup> In this sense the specification is semi-structural in nature: the only identified innovation is the monetary policy shock.

#### 2.1 Estimated Results for Baseline Models

The vector moving average (VMA) form of the structural model in equation (1) gives the dynamic responses of the macroeconomic variables to an exogenous monetary policy innovation. For  $\Theta(L) \equiv (\Theta_0 + \Theta_1 L + \Theta_2 L^2 + ...) = \Phi(L)^{-1}$ , the VMA can be written as

$$X_t = \Theta(L) \varepsilon_t \,. \tag{4}$$

If  $\mu_t$  is the  $j^{\text{th}}$  element of  $\varepsilon_t$  then  $\theta_{ijh}$  measures the *ceteris paribus* response of  $X_{i,t+h}$  to a  $\mu_t$  shock, a one-time exogenous increase in the federal funds rate. Figure 1 presents the estimated impulse response functions for a baseline model estimated without any indicator variables. Log levels of all variables (except the funds rate) are used in estimation.<sup>8</sup> Estimation is monthly from

 $<sup>^5</sup>$  Feedback from policy into production and pricing decisions becomes more likely as the length of a period (or the magnitude of a policy shock) increases. Notice that while these restrictions have a time-dependent interpretation, their economic justifications — often based on costs of adjustment — tend to be state-dependent in nature.

<sup>&</sup>lt;sup>6</sup> A closely related set of identifying assumptions, which associates the policy instrument  $m_t$  with nonborrowed reserves, can be found in Strongin (1995). Over the sample examined herein, Strongin's model yields a slightly less protracted price puzzle. See Leeper et al. (1996) for a critical analysis of this identification.

<sup>&</sup>lt;sup>7</sup> Keating (1996) demonstrates that the ordering of variables in the non-policy block does not affect the interpretation of responses to a policy shock as identified above.
<sup>8</sup> While unit root tests on these series tend not to reject a null of non-stationarity, I follow previous authors and estimate the VAR with log levels. Bernanke and Mihov (1998) report few differences between estimation in log differences and in log levels.

1959:01 to 1998:12, and 12 lags are included in each equation.

For each impulse response plotted, the solid line reports the point estimate for the response of the variable listed at the top of a given column to the monetary policy shock,  $\mu_t$ . The responses are reported as percentage points except for the own response of the funds rate, which is in basis points. The 68% bootstrapped confidence interval is indicated in dark grey, while the light grey region represents a 95% interval.<sup>9</sup>

Each row of figure 1 represents a different specification of the activity variables,  $Y_t$ ; the policy block always is specified as above. The first three rows use industrial production as the output measure.<sup>10</sup> The U.S. consumer price index (CPI) measures the aggregate price level in the first row. The second uses the CPI excluding shelter (CPIXH).<sup>11</sup> In the third row the price level is measured by the personal consumption expenditure deflator (PCE). These three price measures imply fairly similar dynamics for all variables shown in figure 1. Over the sample period in question the Fed followed the consumer price index most closely, only formally switching to the PCE deflator in February 2000.<sup>12</sup> The final row uses quarterly values for gross domestic product as the output measure and the GDP deflator (PGDP) as the price measure.<sup>13</sup>

All four measures of the activity variables yield a statistically significant price puzzle that persists for over one-half year, according to the 95% confidence intervals. Comparing my results with those of Sims (1992) and Christiano et al. (1996b), who report 68% error bands only, reveals a price response that lies above this interval for at least a year and a half in all cases. The point estimates finally turn negative during the third year following a monetary policy contraction.

Although the price puzzle is most pronounced with the CPI and least pronounced with the PCE deflator, a broad consistency exists across the various price measures. The pattern of the output responses is nearly identical in each case as well: exogenous policy shocks have a lagged effect on output; a statis-

According to Sims et al. (1990), estimation in (log) levels will be consistent.

 $<sup>^9</sup>$  See Kilian (1998) for a discussion of these bias-corrected bootstrapped confidence intervals. 1000 bootstrap replications were performed for each response, drawn from the sample distribution of residuals.

<sup>&</sup>lt;sup>10</sup> Similar results were obtained with the unemployment rate as the output measure. <sup>11</sup> Prior to 1983, the method used to impute the cost of owner-occupied housing mismeasured actual CPI inflation.

 $<sup>^{12}</sup>$  In a comparison of the CPI and the PCE deflator, Clark (1999) concludes that CPI measures are more appropriate for the basis of monetary policy decisions.

<sup>&</sup>lt;sup>13</sup> Quarterly estimation of the specifications corresponding to the first three rows of figure 1 tended to yield less pronounced but positive price responses and wider confidence bands.

tically discernible reduction occurs roughly six to nine months following the shock, which peaks at about 0.8 percent of a reduction two years after the initial impulse.<sup>14</sup> The own responses of the funds rate also are quite similar for the monthly estimates; the estimated quarterly monetary policy shocks in figure 1 have a greater variance and therefore a larger impact effect (roughly 125 basis points versus 50).

#### 2.2 Modeling the Conventional View of the Price Puzzle

Sims (1992) argued that the price puzzle was a result of erroneously identifying the exogenous part of monetary policy. Suppose what had been labeled an "exogenous shock" in fact contained some portion of the endogenous response of the Fed to higher expected inflation. Then the impulse response to a contractionary policy shock would appear to lead to an *increase* in prices: higher interest rates are followed by higher inflation. But notice that causality is the reverse: the realized increase in expected inflation has caused the prior (endogenous) increase in the funds rate.<sup>15</sup> The implication is that an empirical researcher could more accurately identify the truly exogenous component of monetary policy by including in the VAR variables that indicate future inflation.

To see the consequences of this now common view, start by formalizing the monetary policy reaction function in equation (2) as

$$r_t^f = \beta[\pi_t^e - \bar{\pi}] + g(X^t) + \mu_t \,, \tag{5}$$

where  $r_t^f$  is the Fed funds rate,  $\pi_t^e$  is the expected future rate of inflation based on time t information,  $\bar{\pi}$  is the Fed's target inflation rate, and  $g(\cdot)$  represents other possible arguments of the reaction function (for example, the output gap or lags of the policy instrument).

The central question, according to this perspective, is the determination of inflationary expectations. Consider two sets of variables that are correlated with expected inflation:  $\Omega_t$ , included in the estimated model by the researcher, and  $Z_t$ , initially excluded from the estimation.<sup>16</sup> Expected inflation then can

 $<sup>^{14}</sup>$  Each row of figure 1 also exhibits a mild "output puzzle" that disappears within a quarter or so.

<sup>&</sup>lt;sup>15</sup> This effect would be more pronounced the greater the degree of accommodation of inflation by the Fed, or the lower expectations relative to the actual realized future inflation.

<sup>&</sup>lt;sup>16</sup> A purely statistical explanation of the price puzzle does not require the information in  $Z_t$  to enter the Fed's reaction function; it need only improve the forecast of prices within the VAR model. The results presented in section 3 are consistent with such an approach, as well as with the economic model developed here.

be separated into two parts:

$$\pi_t^e = \pi_t^m(\Omega_t) + \pi_t^z(Z_t|\Omega_t) \,,$$

where  $\pi_t^m(\Omega_t)$  represents the "measured" inflation expectations based only upon the information contained in  $\Omega_t$ , and  $\pi_t^z(Z_t|\Omega_t)$  captures the *incremental* contribution of including  $Z_t$  in the information set for expected inflation.

Provided  $Z_t$  possesses some forecasting power for future inflation, a policy rule that is estimated without  $Z_t$  will result in a misspecified model,

$$r_t^f = \beta[\pi_t^m - \bar{\pi}] + g(X^t) + \eta_t \,,$$

in which the mismeasured policy shock is contaminated by a portion of the endogenous response of policy to the information about expected future inflation contained in  $Z_t$ :

$$\eta_t = \beta \, \pi_t^z (Z_t | \Omega_t) + \mu_t \,. \tag{6}$$

One immediate consequence is omitted variable bias. More significantly for the VAR approach, the estimated monetary policy shock mismeasures the actual exogenous component of policy:  $\eta_t \neq \mu_t$ .

Conventional macroeconomic theory suggests that a contractionary monetary policy shock should (eventually) reduce the price level, but never raise it:  $\frac{\partial P_{t+h}}{\partial \mu_t} \leq 0$  for all h > 0. Yet as illustrated in figure 1, the opposite appears in the estimated baseline VAR models. Notice that

$$\frac{\partial P_{t+h}}{\partial \eta_t} = \frac{\partial P_{t+h}}{\partial \mu_t} + \frac{\partial P_{t+h}}{\partial [\beta \pi_t^z]},\tag{7}$$

as  $\pi_t^z$  and  $\mu_t$  are orthogonal by construction. For relatively short horizons (i.e. small values of h), the estimated impulse responses  $\left(\frac{\partial P_{t+h}}{\partial \eta_t}\right)$  are positive while those implied by theory  $\left(\frac{\partial P_{t+h}}{\partial \mu_t}\right)$  are zero or negative. Therefore, by equation (7) the conventional view of the price puzzle requires

$$\frac{\partial P_{t+h}}{\partial [\beta \pi_t^z]} > -\left(\frac{\partial P_{t+h}}{\partial \mu_t}\right) \text{ for some } h > 0.$$
(8)

The above framework isolates the components needed to support the conventional view of the price puzzle. First, the excluded information contained in  $Z_t$  must offer *incremental* forecasting power for future inflation beyond that already in  $\Omega_t$ . Should  $Z_t$  indicate higher future inflation  $(\pi_t^z > 0)$  — and provided that the Fed raises interest rates in response to higher expected inflation  $(\beta > 0)$  — then equation (6) shows that the estimated policy shock could be positive even when the "true" exogenous innovation to policy were zero or negative. In other words, by excluding  $Z_t$  an empirical researcher would incorrectly infer that a contractionary policy shock, rather than an endogenous response to greater expected inflation, had raised interest rates. Second, the dynamic response of the price level to the excluded inflation forecast information must be sufficiently large and positive to offset the negative impact of a "true" policy shock upon prices, as in equation (8).

Direct tests of these hypotheses are precluded as neither the "true" policy shock nor inflationary expectations can be observed. However, some implications of this model can be studied empirically. First, indicators with greater incremental forecasting power should exhibit greater reductions in the price puzzle. Comparisons across indicators and forecast horizons are examined in the next section. Second, the severity of the perverse price response should increase with  $\beta$ , the degree to which the Fed reacts to inflationary pressures. A number of researchers recently have concluded that this parameter has varied significantly over time in the U.S. Cross-sample evidence on the price puzzle is investigated in section 4.

# 3 Indicator Variables in SVAR Models

Following Sims (1992) and Christiano et al. (1996b), the information variable most commonly added in the VAR literature is commodity prices. Yet other variables might plausibly contain information about future inflation.<sup>17</sup> In this section I investigate the forecasting power of a large set of candidate indicators over multiple horizons, and ask whether resolution of the price puzzle is related to forecasting power. Tables 1 and 2 list the particular series I examine. This list of potential indicators is meant only to be representative, not exhaustive.<sup>18</sup>

There are two broad classifications for the role played by potential indicator variables. One is a "chain of production" or "pass through" channel: an increase in the costs of intermediate inputs could lead to an increase in the prices of final goods directly. For example, certain indicators might impact marginal costs, a key component of inflation dynamics in the emerging "New Neoclassical Synthesis." In a similar vein, some variables could measure "supply shocks." Examples for this channel may include producer price indices,

<sup>&</sup>lt;sup>17</sup> Bernanke and Mihov (1998) explore including the index of leading indicators as an information variable for future output. They ultimately discard this variable from their preferred specifications.

<sup>&</sup>lt;sup>18</sup> Cecchetti (1995), Stock and Watson (1999), and Cecchetti et al. (2000) provide complementary investigations of inflation indicators. Christiano et al. (1996a) remark in a footnote that, unlike commodity prices, oil prices did not resolve the price puzzle in their model. Balke and Emery (1994) investigate some of these measures and conclude that commodity prices and the long-short interest rate spread, under certain circumstances, can resolve the price puzzle. I analyze each of these indicators in greater detail below.

the price of oil, labor costs and capacity measures. Import prices or exchange rates may also fit this story.

A second channel exploits differing degrees of price flexibility. If the aggregate price level adjusts sluggishly to shocks, more flexible prices could signal a future increase in aggregate prices without necessarily feeding into them in a direct, causal manner. An "informational" story is the basis of the exchange rate overshooting model of Dornbusch (1976). Other asset prices may perform in a similar fashion, including interest rates (i.e. the prices of bonds). Notice that commodity prices plausibly could represent either channel, and that these channels are not mutually exclusive.

### 3.1 Forecasting Power of Candidate Indicators

Tables 1 and 2 report the root mean squared error (RMSE) for a VAR-based forecast of the price level at various horizons. From equation (4), the forecast error for the price level h-periods ahead, conditional on information available at time t, can be written as:

$$p_{t+h} - \hat{p}_{t+h|t} = \theta_{\cdot 20} \varepsilon_{t+h} + \theta_{\cdot 21} \varepsilon_{t+h-1} + \theta_{\cdot 22} \varepsilon_{t+h-2} + \dots + \theta_{\cdot 2h-1} \varepsilon_{t+1},$$

where  $\hat{p}_{t+h|t}$  represents the forecasted value of  $p_{t+h}$  based upon data observed in period t or earlier.  $\theta_{2h}$ , a  $1 \times n$  vector, is the second row of  $\Theta_h$ .<sup>19</sup> In table 1 the price level is measured by the CPI excluding shelter; in table 2 it is the personal consumption expenditure (PCE) deflator.

The mean squared error (MSE) of the price level forecast, h periods ahead, is

$$MSE(\hat{p}_{t+h|t}) = E\left[(p_{t+h} - \hat{p}_{t+h|t})(p_{t+h} - \hat{p}_{t+h|t})'\right] \\ = \theta_{\cdot 20} E[\varepsilon_t \varepsilon'_t] \theta'_{\cdot 20} + \theta_{\cdot 21} E[\varepsilon_t \varepsilon'_t] \theta'_{\cdot 21} + \dots + \theta_{\cdot 2h-1} E[\varepsilon_t \varepsilon'_t] \theta'_{\cdot 2h-1} \\ = \theta_{\cdot 20} \theta'_{\cdot 20} + \theta_{\cdot 21} \theta'_{\cdot 21} + \dots + \theta_{\cdot 2h-1} \theta'_{\cdot 2h-1}$$

since  $E[\varepsilon_t \varepsilon'_t] = I$ . The RMSE is the square root of  $MSE(\hat{p}_{t+h|t})$ , measured as percentage points.

The first line of tables 1 and 2 gives the RMSE of the baseline VAR. Subsequent rows are for augmented 6-variable VARs with the indicator in the first column included as the final variable of  $Y_t$ , the non-policy block. These rows display the percentage reduction in the RMSE of the price level forecast from including the indicator variable listed at left, relative to the baseline model (i.e. the baseline-model RMSE less the augmented-model RMSE). Thus a larger

<sup>&</sup>lt;sup>19</sup> The price level always is the second variable in  $X_t$  for the models considered here.

value implies more incremental for ecasting power for the particular indicator at the specified for ecast horizon.  $^{20}\,$ 

Before discussing the results, note that the way in which the indicator variables enter the VAR is not innocuous. By placing them among the activity variables, the block-recursive structure imposes that these indicators cannot respond to contemporaneous monetary policy innovations. This restriction may be undesirable for variables operating within an "informational" channel as outlined above. While this issue may seem especially acute for interest rates, commodity prices — which are set in forward-looking asset markets — also are likely to suffer this problem. For consistency with the broader literature, I treat all candidate indicator variables symmetrically, and enter them as commodity prices commonly appear in other VAR studies. This potential misspecification seems unlikely to account for the "success" of commodity prices, nor bias the subsequently reported results in favor of any potential indicator over another. To the extent the question of interest regards adding information to the monetary policy rule rather than modeling the complete system, the import of this issue may be lessened.<sup>21</sup>

The first two tables report RMSE for several horizons likely to be relevant for monetary policy; in light of "long and variable" outside lags of policy, forecastability at longer horizons may be important for monetary policy makers. Below I focus on the 6- to 12-month horizons, as the puzzle almost always peaks within this forecast interval when estimated over the full sample period. Recall one implication of the model in section 2.2: if the price puzzle is due to an excluded inflation forecast measure, then at horizons for which the puzzle is larger, indicators that resolve the puzzle should have greater incremental forecasting power.

Given that intuition, several broad results from tables 1 and 2 warrant attention. First, the results are qualitatively similar between the two tables. Second, the incremental forecasting power of many indicators increases monotonically with the length of the forecast horizon, at least over the first year. However, relative to the total RMSE of the baseline model, the proportion of the unexplained variation in prices that can be accounted for by any given indicator tends to fall as the horizon increases in length. These proportions also are relatively small: adding any one of these indicators does not improve

 $<sup>^{20}</sup>$  I have also tested whether each indicator Granger causes the price level in the context of the augmented VAR models. Although not reported here, the results parallel those in tables 1 and 2.

<sup>&</sup>lt;sup>21</sup> Some authors, e.g. Leeper et al. (1996) and Sims and Zha (1998), have posited non-recursive identifying assumptions to address this issue. These assumptions can be more difficult to justify and therefore are viewed by some as controversial. More significant for this study, they vary by particular indicator variable, making it difficult to separate the role of the indicator from that of the identification scheme.

dramatically the forecast of future prices at the horizons examined.<sup>22</sup> Finally, no single indicator (or class of indicators) produces the lowest RMSE at all horizons shown.

The conventional view of commodity prices as indicators of future inflation is evident in both tables, at least at shorter horizons. Almost all the commodity price measures exhibit a fall in their incremental forecasting power (in absolute value and as a proportion of the baseline) at horizons of one year or greater. While commonly-used broad commodity price indices — the CRB spot index and the IMF overall index — perform well, both tables indicate that the more narrowly-defined "raw materials" indices — the CRB raw industrial materials index, the price of sensitive materials index, and the IMF agricultural raw materials index — generally perform even better.<sup>23</sup>

To distinguish the two channels defined above, first consider measures of the "pass-through" or "marginal cost" channel. Neither the oil price level nor the net increase in oil prices over the year, as defined in Hamilton (1996), have much incremental forecasting power as compared with commodity prices. The two PPI measures, however, show improvements in the RMSE comparable to those for commodity prices. <sup>24</sup> A nearly similar improvement can be attributed to average hourly earnings, especially as the forecast horizon is lengthened. If moving through the chain of production takes a while, a "pass-through" effect may be most significant at longer horizons. Measures of constraints on production, such as the unemployment rate and capacity utilization, exhibit minimal (but positive) incremental forecasting power.

As for the "information" channel, the expectations hypothesis of the term structure implies that a long-short interest rate spread should signal expected future inflation. However, neither the long-short nor the "quality" spread measure listed in tables 1 and 2 differs substantially from the baseline model at any horizon. Kozicki (1997) reports better inflation forecasting power from interest rate levels instead of spreads, particularly the long bond rate. But in tables 1 and 2, levels perform similarly to spreads and do not appreciably improve upon the baseline case.

Monetary aggregates provide somewhat greater information about future prices by this metric, particularly M2 at longer horizons — but still less than com-

 $<sup>^{\</sup>overline{22}}$  Computational considerations precluded testing multiple indicators simultaneously.

<sup>&</sup>lt;sup>23</sup> The IMF overall index is particularly interesting, as several of its components are reported as separate indices. The sub-indices for foodstuff and metals have noticeably lower incremental forecasting power than the agricultural raw materials index.

 $<sup>^{24}\,{\</sup>rm Sims}$  and Zha (1998) use the crude materials PPI in place of a broad commodity price index.

modity prices. Friedman (1997), among others, has noted that the relationship between money and inflation has broken down during the last decade or so, and today few economists would propose tracking a broad monetary aggregate for the purpose of forecasting inflation over the horizons considered here.

Finally, the nominal trade-weighted exchange rate exhibits noticeable forecasting power at all horizons, with relative strength at horizons of one year or more. Indeed, at a two-year horizon it has the single greatest incremental forecasting power of all the indicators considered. Note that the exchange rate could reflect either channel discussed above.

# 3.2 Alternative Indicators and the Price Puzzle

Forecasting power is one side of the conventional wisdom regarding the role of inflation indicators in VAR models; a "successful" indicator variable also must eliminate the price puzzle. In this section I estimate a series of models, each augmented with one indicator from table 1. To match the approach most common in the literature, the indicator variable is included in the non-policy block,  $Y_t$ , after output and the aggregate price level. As mentioned above, this practice could introduce a separate source of misspecification, but one that *a priori* is unlikely to favor any particular indicator.

With that caveat in mind, figures 2 through 5 report the responses of the CPI price level (excluding housing), industrial production, and the federal funds rate to a contractionary policy shock in several augmented models.<sup>25</sup> The first row in each of these figures reproduces the 5-variable baseline case as in figure 1. The remaining rows plot the responses to the policy shock when the estimation is augmented with the indicator variable listed at the left of the row. The final column gives the response of that indicator variable to the policy shock.

The baseline case was presented in section 2.1: the price level response is significantly positive for at least 9 months, while output falls after a little more than half a year and the funds rate remains above its initial value for just over a year. This pattern of responses was replicated across all price and output measures displayed in figure 1.

The augmented cases tend to resemble the baseline, although there are some important differences. First, including any of the broader commodity price indexes in the VAR does reduce the length of the positive price response

 $<sup>^{25}</sup>$  Replacing the CPI excluding housing with the PCE deflator yields similar impulse responses across the indicators shown. In the interest of space, those figures are not reproduced here.

but does not completely eliminate it. This finding of a residual price puzzle is replicated in several recent monthly studies, including Leeper et al. (1996), and Christiano et al. (1999), and contrasts with previous quarterly results (see, e.g., Christiano et al., 1996b). Commodity prices also appear to marginally reduce the length of the funds rate response. The responses of these indicators to the policy shock are plausible as well: commodity prices do not exhibit a price puzzle themselves, although the magnitudes of their point estimates are quite a bit larger than the aggregate price level responses.

Figure 3 illustrates the impulse responses for various measures of "costs of production," i.e. measures associated directly with the "pass-through" channel. The two producer price indices both reduce — but again, do not eliminate — the duration and magnitude of the counter-intuitive price response. Notice that their ability to reverse the puzzle is not as strong as the commodity prices in figure 2. Further, these measures themselves exhibit a mild positive response to exogenous policy shocks (albeit statistically insignificant). In contrast to the commodity price measures, which are generally set in flexible auction markets, the more sluggish nature of both PPI measures — as with the aggregate (final goods) prices shown in figure 1 — appears to be associated with a perverse response to contractionary policy shocks. This finding suggests that the "information" channel may be the more important one for understanding the role of commodity prices.

On the other hand, the oil price — measured in figure 3 by Hamilton's (1996) net oil price increase — has no discernible effect upon any of the impulse responses, including the aggregate price level. Moreover, this oil measure itself exhibits a statistically significant price puzzle.<sup>26</sup> To the extent that rapid increases in oil prices commonly are thought to represent "adverse supply shocks" or "cost shocks," the evidence does not suggest a significant role for these variables for the resolution of the price puzzle. Arguably, the "success" of commodity prices then is not attributable to being a proxy for the adverse inflationary episodes often associated with the oil price shocks.

Selected interest rates and spreads are considered in figure 4. The long-short spread (defined as the 10-year bond less the 3-month T-bill rate) mitigates the price puzzle only by a few months at most, and arguably worsens the perverse — but short-lived — positive output response to contractionary policy evidenced in monthly VARs. Contractionary monetary policy initially flattens the yield curve according to these estimates, with all the action coming from the short end of the market, as revealed in the third and fourth rows of figure 4. Indeed, the reduction of the price puzzle appears to occur primarily through the short-term interest rate (which also happens to eliminate any "output puzzle") — perhaps because it reduces the measured policy shock by nearly

<sup>&</sup>lt;sup>26</sup> The results for unfiltered oil price data are qualitatively similar on both counts.

one-half. Again, the lack of feedback from the various interest rate measures to policy, as codified in the block recursive identification, may represent a misspecified model. This same critique could be levied against most standard VAR models that include commodity prices, however, as the above evidence suggests the "information" channel is most likely the reason commodity prices play any macroeconomic role in a monetary VAR.<sup>27</sup>

In his 1992 paper, Sims added both a commodity price and the exchange rate to his baseline VAR models. The impulse responses in figure 5 indicate why exchange rates no longer are routinely included as inflationary indicators in VAR models for the U.S.: the estimated price response is virtually identical. Policy also appears to have no statistically discernible effect on the exchange rate itself — although the point estimates imply an appreciation in response to a contractionary shock, as expected. Like commodity prices, exchange rates could help forecast inflation through either an "information" channel (as in a Dornbusch-style overshooting framework) or a "pass-through" channel (as a component of the cost of imported intermediate goods). However, unlike commodity prices they have little consequence for the impulse responses.

Figure 5 also includes two other candidate indicators: average hourly earnings and the capacity utilization rate. The former has no noticeable effect upon any of the impulse responses, while the latter does reveal a reduction in the duration of the price puzzle — at the cost of a larger output puzzle. Interestingly, capacity utilization itself exhibits the same small and very short, yet perverse, positive response to a contractionary policy shock. Unexpected contractionary policy reduces average hourly earnings briefly; in combination with the apparent rise in prices a monetary contraction would appear to reduce real wages for a sustained period of time.

# 3.3 Are Forecasting Power and the Price Response Linked?

The logic behind Sims's (1992) original justification for including commodity prices in a monetary VAR model, to which a sizable number of subsequent authors have given their assent, is straight-forward: commodity prices resolve (or at least greatly mitigate) the price puzzle because they purge the estimated policy shock of endogenous responses of the Fed to expected future price changes. This notion was formalized in section 2.2 with a model that linked the inflation forecasting power of a candidate indicator with the magnitude of the positive price level response.

 $<sup>^{27}</sup>$  Work by Leeper et al. (1996) and Sims and Zha (1998) that use non-recursive assumptions to identify an "information sector" — including variables such as commodity prices — still tend to produce residual price puzzles.

The incremental forecasting power of individual indicators is plotted against the average reduction in the price response to a contractionary policy shock for the consumer price index (CPI), the CPI less shelter (CPIxH), the personal consumption expenditure deflator (PCE) and the GDP implicit price deflator (PGDP, measured quarterly) in figures 6 to 9, respectively. The horizontal axis measures the percentage reduction in the RMSE of the price level forecast from including the indicator variable shown, as in tables 1 and 2. The vertical axis shows the reduction in the level of the impulse response function averaged over the length of the specified forecast horizon (i.e. the average size of the point estimate of the price level response in the baseline model, less the average size of the point estimate response of the appropriate augmented model). Notice that the maximum value of both axes tends to increase with the forecast horizon plotted.

Variables with stronger forecasting power are positioned towards the right side of each graph. Those that reduce the positive price level response the most lie near the top of the graph. If greater forecasting power implies a less pronounced puzzle, one might initially expect the points plotted to generally lie along a line from the upper-right to the lower-left corners. With a few exceptions there is no conclusive evidence of such a relationship, and several outliers emerge. At shorter horizons, the CPIxH (figure 7) and PCE (figure 8) measures of prices at first appear to satisfy the hypothesized relationship between forecast power and price response. However, in each of these cases most of the indicators lie in a cloud in the lower-left corner, with a few commodity price measures in the upper-right region. As these commodity price indexes often share components (and thus are highly correlated by construction), they should not necessarily be treated as independent observations on the hypothesis in question.

A weaker implication of the model in section 2.2 — which accords with an intuitive justification for including commodity prices — is that indicators that exhibit greater incremental forecasting power should coincide with indicators that more substantially mitigate the price puzzle. However, rankings by indicator of forecasting power and puzzle mitigation do not coincide for the price measures and horizons plotted in figures 6 through 9. Additionally, several indicators contradict this implicit ordered relationship among the list of candidates. For example, at horizons of less than a year, the 3-month Treasury bill rate (TBILL) and the long-short spread (SPRD1) exhibit a comparable ability to resolve the price puzzle as the various commodity price measures when the CPI is the measure of prices (figure 6). Yet neither candidate indicator has much incremental forecasting power, suggesting that forecasting power is not necessary for resolving the price puzzle. A similar result is found for capacity utilization (CU) in the VARs with the GDP deflator as the price measure (figure 9).

Conversely, for each price measure at horizons when the price puzzle is the largest (generally centered around a year) there often exist at least one or two indicators that have substantial forecasting power yet fail to appreciably change the path of the price response. For example, at the 12-month horizon for a VAR with the PCE deflator as the aggregate price measure (figure 8), an approximately vertical line can be drawn through no fewer than eight candidate indicators: all have almost identical incremental forecasting power, yet vary widely in their impact upon the price response, with some (e.g. average hourly earnings (LAHE) and M2) actually exacerbating the price puzzle relative to the baseline case.<sup>28</sup> For CPI excluding shelter (figure 7), the exchange rate (TWXR) exhibits similar behavior vis-à-vis various commodity price measures at the 9-month horizon and higher. Together, these results suggest that forecasting power may not be sufficient to resolve the price puzzle.

Perhaps the most interesting finding among the commodity price measures is the dominant forecasting power of the IMF's agricultural raw materials price index (IMFA). For each of the four aggregate price measures considered in figures 6 to 9, this index always raises the RMSE of the price level forecast the most — although it exhibits the largest reduction only for the GDP deflator. As the raw materials in this index comprise a small fraction of the total inputs into production in the U.S., a "pass through" or "marginal cost" channel is even less likely to be the reason commodity prices appear to resolve the price puzzle. Rather, this finding lends further credence to an "information" channel, in which flexible commodity price movements anticipate subsequent changes in more sluggish consumer prices.

Also surprising is the non-impact of interest rate spreads upon this hypothesized relationship: the spreads (or the interest rates themselves) do not appear to have much incremental forecasting power, despite a well-developed theory relating interest rate spreads (the term structure) and inflationary expectations. In contrast, the theoretical link between commodity prices and aggregate consumer prices is tenuous at best. And with the exception of the CPI index at certain horizons, the various interest rate measures do not significantly reduce the price level response to contractionary policy shocks. In light of the identifying assumptions, these results must be interpreted with care. But why a possibly misspecified model should reveal one class of information variable (i.e. commodity prices) to satisfy a hypothesis linking forecasting power and the price response, and not another (i.e. interest rates), remains unclear under the conventional approach.

Recall that the peak of the point estimate for the positive price level response is nearly a year and a half after the initial policy shock. To the extent that hori-

 $<sup>^{28}</sup>$  Similar results, albeit not quite as striking, can be found at the 9-month horizon as well.

zons beyond one year are relevant for policymakers (perhaps due to "long and variable lags" of monetary policy), the results at 18- and 24-months reverse the relationship: for a majority of cases, the relationship between forecasting power and resolution of the price puzzle (for those indicators that do either) appears to be negative. In particular, commodity prices exhibit little incremental forecasting power beyond the baseline estimates yet greatly reduce the positive price response, whereas several indicators — such as the exchange rate, average hourly earnings, and the S&P 500 stock index — exhibit significant forecasting power while scarcely affecting the impulse response for the aggregate price level.

In summary, the link between incremental forecasting power for prices and the degree of resolution of the price puzzle is not nearly as strong as one might suspect from only considering commodity prices: among the indicators examined here, the former attribute is neither necessary nor sufficient for the latter. These results offer a challenge to the intuition that forms the basis of the conventional wisdom for including commodity prices in a monetary VAR model, as commodity prices appear to be more the exception than the rule.<sup>29</sup> Perhaps the true "price puzzle" is the apparently unique role commodity prices appear to play in monetary VAR models. This restatement becomes more salient in light of the lack of a theoretical justification for their inclusion and concerns about their inclusion under some common identifying schemes used in the VAR literature. Thus, the continued inclusion of commodity prices in estimated monetary models may warrant a re-examination by empirical practitioners.

# 4 Analysis of Sub-sample Estimates

In October 1979, Fed Chairman Paul Volcker announced a shift from effectively targeting the federal funds rate to explicitly targeting nonborrowed reserves. This policy change potentially impacts the models estimated above in two ways. First, identifying the policy instrument with the funds rate would be inappropriate for the 1979 to 1982 period.<sup>30</sup> Second, this shift in instrument might have been accompanied by a more general change in the policy reaction function. Taylor (1999) has argued that the Fed accommodated inflation to

 $<sup>^{29}</sup>$  With regard to the general hypothesis linking forecasting power and the price response, a particularly stark interpretation is that the commodity price results simply represent type I error. I thank Chris Foote for this observation.

 $<sup>^{30}</sup>$  Bernanke and Mihov (1998) report that a Fed funds targeting model does not fit this period very well. They also demonstrate that a nonborrowed reserves model performs better over the 1979 – 1982 period than any other period they examine. Strongin (1995) discusses Federal Reserve operating procedures since 1959.

a much greater extent prior to 1979 than after 1982. Clarida et al. (2000) confirm these results with a generalized version of a "Taylor rule" that incorporates interest rate smoothing and forward-looking behavior. In the terms of the model of section 2.2, both studies conclude that  $\beta$ , the coefficient on the difference between (forecasted) inflation and its target value, was significantly smaller for the pre-Volcker Fed. But it is also possible that the components of the information set — or the forecasting power of various indicators — have changed over time. For these reasons, estimation over a sample that contains pre- and post-1979 (or 1982) observations may be inappropriate to examine the causes and cures of the price puzzle. This section examines whether variation within the longer sample exists, and whether such variation supports the conventional explanation of the price puzzle.

Figure 10 replicates the baseline model for the full sample and two sub-sample periods: 1959:01 – 1979:10 and 1982:11 – 1998:12. These break dates are chosen to correspond with the period of experimentation with nonborrowed reserves targeting and thus match those used elsewhere in the literature. Monthly CPI excluding housing is the measure of prices reported here; qualitatively similar results are obtained with other price level variables, and with quarterly data. A comparison down the first column suggests that the price puzzle is a regime-specific phenomenon: the price level response is not statistically significant in the post-1982 sub-sample, while the pre-1979 puzzle is quite protracted — nearly twice as long as in the full sample — and relatively large in magnitude.

The output response, shown in the second column, also differs substantially across regimes. Industrial production shows a deeper, albeit slightly less protracted, decline in response to the policy shock for the earlier period relative to the full sample results. By contrast, output falls much less in the later sub-sample and this decline is less evident statistically.

Consistent with the broader monetary literature, the impulse responses for the funds rate also reveal apparent sharp differences in the nature of monetary policy before and after the Volcker disinflation. First, the magnitude of the initial impulse to the funds rate is almost twice as great during the 1959:01 – 1979:09 sample as during the 1982:11 – 1998:12 one. That is, the average contractionary innovation was substantially larger in the pre-Volcker period. Second, the dynamic path of the funds rate to its own policy innovation reveals a pronounced reversal of policy — more than a 25 basis point *reduction* in the funds rate — a little over two years following the initial increase in rates. This response function is consistent with a view of "stop-go" monetary policy during that period, as well as the claim that the Fed responded more to output than inflation. Notice that 18 months following a contractionary shock, when the funds rate response moves into negative territory, output has fallen nearly a full percent while prices have not yet fallen at all — indeed, the point estimate still shows an increase in prices.

These results are not inconsistent with  $\beta < 1$  in the earlier sample period, as reported by Taylor (1999), Clarida et al. (2000), and others. Moreover, there is relatively little evidence of a significant price puzzle in the later, supposedly more stable, policy regime of Volcker and Greenspan. The point estimate of the price response is still positive but with a generally smaller magnitude than in the other sample periods, and it is statistically indistinguishable from zero for most of the period shown. Recall the second implication mentioned at the end of section 2: the larger the value of  $\beta$ , the greater should be the magnitude of the puzzle — if the true reason for the puzzle is an excluded indicator variable — and the stronger should be the link between the forecasting power of such indicators and their ability to reduce the positive price response. The first part of this implication does not correspond with the impulse responses shown in figure 10: a smaller  $\beta$  commonly is attributed to the nature of monetary policy during 1959:01 – 1979:09 period, yet this period exhibits a more pronounced price puzzle.

An alternative explanation for the sub-sample variation shown in figure 10 is that the dynamics of the economy differ for reasons unrelated to the specification of the monetary policy reaction function. A growing recent literature has investigated whether the U.S. macroeconomy, and output in particular, has become more less volatile since the early 1980s. While this question is yet unresolved, it also is consistent with some of the results illustrated in figure 10. For example, the greater magnitude of policy shocks in the earlier sample period may themselves be a source of greater instability, or they may reflect the need for larger movements in an exogenous forcing variable to control a less responsive and more volatile system. Further research into this issue is warranted, and may provide additional clues into the nature of the price dynamics illustrated here.

Despite the above results, a more narrow view of the conventional approach to the price puzzle might posit that the estimated sub-sample policy innovations still could be mis-measured if an appropriate inflation forecasting variable were excluded from these models. Thus, figures 11 and 12 illustrate the relationship between incremental forecasting power and the CPI (excluding shelter) price response for the 1959:01 – 1979:09 and 1982:11 – 1998:12 sub-samples, respectively.

For the earlier sub-sample period (figure 11), the IMF raw agricultural price index (IMFA) is again the most preferred indicator among the commodity price measures — at least at short horizons. Other commodity price indices perform similarly to a variety of other indicators, and have no reason to be preferred to them. Most interesting may be the long-short spread (SPRD1) and the exchange rate (TWXR) indicators: at the 3-month forecast horizon both show approximately the same forecasting power, yet the spread largely eliminates the price puzzle while the exchange rate generates less than one-fourth as large a reduction in the positive price response. As the horizon increases, the forecasting power of the exchange rate rises without nearly as great a relative improvement in its ability to reduce the price response. In contrast, the long-short spread always exhibits the strongest or second-strongest reduction in the price response, yet its incremental forecasting power actually *falls* as the forecast horizon increases. These results do not appear to offer support for the conventional view of the cause and resolution of the price puzzle.

For the latter sub-sample, figure 12 yields further results that are at odds with the conventional view of the price puzzle: at each horizon shown, M2 and the exchange rate (TWXR) always are among the indicators with the largest incremental forecasting power, yet for most horizons less than two years in length, including these measures in the VAR actually raises the price response to contractionary policy shocks. As these indicators worsen the price puzzle, the original logic behind the inclusion of an inflation indicator is turned on its head. Notice from figure 12 that the commodity price indicators have, on average, nearly zero incremental forecasting power and lead to only moderate reductions in the price puzzle.

#### 5 Conclusion

Within the empirical monetary policy literature, a puzzle commonly arises in vector autoregressive (VAR) models: a significant, protracted increase in aggregate prices following what the researcher has labeled a contractionary monetary policy shock. The conventional approach, following a suggestion by Sims (1992), is to include commodity prices in the VAR as an indicator of future inflation. Lacking a theoretical foundation, previous justification for the inclusion of any "information variable" — such as commodity prices within a monetary VAR has been fairly *ad hoc*. A simple model indicative of the conventional wisdom for including such measures is developed in section 2. Once a broader set of indicators is considered, several implications of this model are not strongly supported in the data. In particular, the results of section 3 do not appear to be consistent with a general, systematic relationship between the ability of a given indicator to forecast prices (as measured by the RMSE of the price level forecast from the VAR) and its ability to prevent a price puzzle in standard VAR models.<sup>31</sup> Further, a sub-sample investigation in section 4 reveals results at odds with both the conventional wisdom and

<sup>&</sup>lt;sup>31</sup> Barth and Ramey (2001) have suggested that a "working capital" or "cost channel" can explain why prices increases follow interest rates increases in the short run. Christiano et al. (2003) develop a model with such a channel that (in conjunction with other assumptions) accounts quantitatively for a positive initial price response to a monetary shock.

the implications of the model.

One interpretation of these findings is to call into question the hypothesized link between indicator variables and the resolution of the price puzzle, and by extension whether monetary policy has been identified correctly in such models. In light of the regular inclusion of commodity prices in empirical monetary models, more research into the macroeconomic role of commodity prices seems warranted. As commodity prices have begun to appear within a broader class of empirical monetary models — ostensibly as inflation indicators for the purpose of monetary policy — these results may serve as a warning for practitioners outside the VAR literature as well.

To the extent that commodity prices do succeed in mitigating the price puzzle, the analysis in section 3 indicates this may be due to an "information" channel — commodity prices respond more quickly than aggregate goods prices to future inflationary pressures — rather than serving as a proxy for marginal costs or otherwise measuring costs of production. Thus a possible interpretation of the findings of this paper is that the traditional identifying assumptions are inappropriate for such an information variable. <sup>32</sup> Determining the appropriate specification of monetary VAR models is necessary for accurately separating policy into its endogenous and exogenous components, and thereby correctly measuring the contribution of monetary policy to economic fluctuations. Under that interpretation, these results might suggest a reconsideration of some of the now-common modeling strategies in the extensive monetary VAR literature.

 $<sup>^{32}</sup>$  Giordani (2001) and Leeper and Roush (2002) study the price puzzle under non-recursive identifying assumptions for monetary VARs.

#### References

- Bagliano, F. C., Favero, C. A., 1998. Measuring monetary policy with VAR models: An evaluation. European Economic Review 42, 1069–1112.
- Balke, N. S., Emery, K. M., 1994. Understanding the price puzzle. Federal Reserve Bank of Dallas Economic Review 4th quarter, 15–26.
- Barth, III, M. J., Ramey, V. A., 2001. The cost channel of monetary transmission. In: Bernanke, B. S., Rogoff, K. (Eds.), NBER Macroeconomics Annual 2001. MIT Press, Cambridge, MA, pp. 199–240.
- Bernanke, B. S., Blinder, A. S., 1992. The federal funds rate and the channels of monetary policy. American Economic Review 82 (4), 901–921.
- Bernanke, B. S., Mihov, I., 1998. Measuring monetary policy. Quarterly Journal of Economics 113 (3), 869–902.
- Cecchetti, S. G., 1995. Inflation indicators and inflation policy. In: Bernanke, B., Rotemberg, J. (Eds.), NBER Macroeconomics Annual 1995. MIT Press, Cambridge, MA, pp. 189–219.
- Cecchetti, S. G., Chu, R. S., Steindel, C., 2000. The unreliability of inflation indicators. Federal Reserve Bank of New York, Current Issues in Economics and Finance 6 (4), 1–6.
- Christiano, L. J., Eichenbaum, M., Evans, C., 1996a. The effects of monetary policy shocks: Evidence from the flow of funds. Review of Economics and Statistics 78 (1), 16–34.
- Christiano, L. J., Eichenbaum, M., Evans, C., 1996b. Identification and the effects of monetary policy shocks. In: Blejer, M. I., Eckstein, Z., Hercowitz, Z., Leiderman, L. (Eds.), Financial Factors in Economic Stabilization and Growth. Cambridge University Press, New York, pp. 36–74.
- Christiano, L. J., Eichenbaum, M., Evans, C., 1999. Monetary policy shocks: What have we learned and to what end? In: Woodford, M., Taylor, J. (Eds.), Handbook of Macroeconomics, volume 1A. Elsevier Science, North-Holland, New York, pp. 65–148.
- Christiano, L. J., Eichenbaum, M., Evans, C., 2003. Nominal rigidities and the dynamic effects of a shock to monetary policy, forthcoming, Journal of Political Economy.
- Clarida, R., Galí J., Gertler, M., 2000. Monetary policy rules and macroeconomic stability: Evidence and some theory. Quarterly Journal of Economics 115 (1), 147–180.
- Clark, T. E., 1999. A comparison of the CPI and the PCE price index. Federal Reserve Bank of Kansas City Economic Review Third Quarter, 15–29.
- Cochrane, J. H., 1994. Shocks. Carnegie-Rochester Conference Series on Public Policy 41, 295–364.
- Dornbusch, R., 1976. Expectations and exchange rate dynamics. Journal of Political Economy 84 (6), 1161–1176.
- Eichenbaum, M., 1992. Comment on 'Interpreting the macroeconomic time series facts: The effects of monetary policy'. European Economic Review 36 (5), 1001–1011.

- Faust, J., Leeper, E., 1997. When do long-run identifying restrictions give reliable results? Journal of Business and Economic Statistics 15 (3), 345–353.
- Friedman, B. M., 1997. The rise and fall of money growth targets as guidelines for U.S. monetary policy. In: Kuroda, I. (Ed.), Towards More Effective Monetary Policy. St. Martin's Press, New York, pp. 137–164.
- Giordani, P., 2001. An alternative explanation of the price puzzle, Stockholm School of Economics.
- Hamilton, J. D., 1996. This is what happened to the oil price-macroeconomy relationship. Journal of Monetary Economics 38 (2), 215–220.
- Keating, J. W., 1996. Structural information in recursive VAR orderings. Journal of Economic Dynamics and Control 20 (9–10), 1557–1580.
- Kilian, L., 1998. Small sample confidence intervals for impulse response estimates. Review of Economics and Statistics 80 (2), 218–230.
- Kozicki, S., 4th quarter 1997. Predicting real growth and inflation with the yield spread. Federal Reserve Bank of Kansas City Economic Review, 39–57.
- Leeper, E. M., Roush, J., 2002. Putting 'M' back in monetary policy, Indiana University.
- Leeper, E. M., Sims, C. A., Zha, T., 1996. What does monetary policy do? In: Brainard, W. C., Perry, G. L. (Eds.), Brookings Papers on Economic Activity. Vol. 2. The Brookings Institute, Washington, DC, pp. 1–63.
- Pagan, A. R., Robertson, J. C., 1998. Structural models of the liquidity effect. Review of Economics and Statistics 80 (2), 202–217.
- Sims, C. A., 1992. Interpreting the macroeconomic time series facts: The effects of monetary policy. European Economic Review 36 (5), 975–1000.
- Sims, C. A., November 1998. Comment on Glenn Rudebusch's 'Do measures of monetary policy in a VAR make sense?'. International Economic Review 39 (4), 933–941.
- Sims, C. A., July 1999. Drifts and breaks in monetary policy, Yale University.
- Sims, C. A., Stock, J. H., Watson, M. W., 1990. Inference in linear time series models with some unit roots. Econometrica 58 (1), 113–144.
- Sims, C. A., Zha, T., July 1998. Does monetary policy generate recessions? Federal Reserve Bank of Atlanta Working Paper 98-12.
- Stock, J. H., Watson, M. W., 1999. Forecasting inflation. Journal of Monetary Economics 44 (2), 293–335.
- Strongin, S., 1995. The identification of monetary policy disturbances: Explaining the liquidity puzzle. Journal of Monetary Economics 35 (3), 463– 497.
- Taylor, J. B., 1999. A historical analysis of monetary policy rules. In: Taylor, J. B. (Ed.), Monetary Policy Rules. University of Chicago Press, Chicago, pp. 319–341.
- Zha, T., 1997. Identifying monetary policy: A primer. Federal Reserve Bank of Atlanta Economic Review 82 (2), 26–43.

	Forecast Horizon								
Indicator	3 mo.	6 mo.	9 mo.	1 yr.	2 yr.				
Baseline	0.5775	0.8665	1.1918	1.5539	2.9888				
Percent improvement over Baseline:									
CRB Spot Index	0.0717	0.1380	0.1811	0.2177	0.0647				
CRB Raw Industrials	0.0768	0.1437	0.1988	0.2304	0.0757				
Price of Sensitive Mat'ls	0.0837	0.1521	0.2120	0.2504	0.1289				
Gold Price	0.0423	0.1027	0.1864	0.2674	0.4400				
IMF Overall Index	0.0632	0.1382	0.2096	0.2680	0.2008				
IMF Foodstuffs	0.0379	0.0792	0.1096	0.1174	-0.0197				
IMF Agr. Raw Mat'ls	0.0957	0.2092	0.2871	0.3482	0.2834				
IMF Metals	0.0221	0.0471	0.0615	0.0595	-0.0605				
IMF Oil Index	0.0240	0.0150	0.0284	0.0391	0.0665				
Crude Oil Price	0.0166	0.0254	0.0392	0.0528	0.1220				
Hamilton's Measure	-0.0000	0.0012	0.0078	0.0121	0.0088				
Crude Materials PPI	0.0432	0.0839	0.1365	0.1960	0.2513				
Intermediate Mat'ls PPI	0.0383	0.0778	0.1335	0.2128	0.5291				
3 mo. Treasury Bill	-0.0025	-0.0071	-0.0105	-0.0108	-0.0124				
3 mo. Financial Paper	0.0010	-0.0003	-0.0057	-0.0073	-0.0186				
10 year Gov't Bond	-0.0029	-0.0051	-0.0115	-0.0179	-0.0413				
Bond – T-bill Spread	-0.0015	-0.0026	-0.0081	-0.0189	-0.0378				
FP – T-Bill Spread	0.0045	0.0179	0.0226	0.0204	0.0346				
Monetary Base	0.0141	0.0194	0.0265	0.0385	0.1013				
M1 (level)	0.0056	0.0196	0.0282	0.0484	0.1154				
M2 (level)	0.0155	0.0439	0.0742	0.1186	0.3226				
Exchange Rate	0.0363	0.0950	0.1725	0.2665	0.6117				
Ave. Hourly Earnings	0.0199	0.0574	0.1139	0.1953	0.6755				
Unemployment Rate	0.0050	0.0066	0.0134	0.0271	0.0632				
Capacity Utilization	0.0220	0.0419	0.0527	-0.0625	-0.0284				
S&P 500 Index	0.0123	0.0399	0.0833	0.1449	0.4621				

# Table 1Root Mean Squared Error for CPI Forecasts, 1959:01 – 1998:12

**Notes:** Table reports difference between RMSE of the baseline model (no indicator) and RMSE of augmented model (model includes indicator listed the left-hand column). Negative numbers indicate those indicators that worsen the forecast relative to the baseline. CPI is measured by the urban consumer price index excluding shelter.

	Forecast Horizon								
Indicator	3 mo.	6 mo.	9 mo.	1 yr.	2 yr.				
Baseline	0.4017	0.6459	0.9142	1.2032	2.4198				
Percent improvement over Baseline:									
CRB Spot Index	0.0416	0.0837	0.1183	0.1442	0.0558				
CRB Raw Industrials	0.0502	0.0974	0.1389	0.1589	0.0568				
Price of Sensitive Mat'ls	0.0558	0.1029	0.1421	0.1611	0.0714				
Gold Price	0.0250	0.0572	0.1028	0.1467	0.2509				
IMF Overall Index	0.0472	0.1055	0.1553	0.1869	0.1122				
IMF Foodstuffs	0.0201	0.0464	0.0666	0.0712	-0.0173				
IMF Agr. Raw Mat'ls	0.0752	0.1680	0.2328	0.2789	0.2790				
IMF Metals	0.0189	0.0384	0.0456	0.0378	-0.0699				
IMF Oil Index	0.0088	0.0039	0.0065	0.0082	0.0133				
Crude Oil Price	0.0021	0.0048	0.0082	0.0129	0.0319				
Hamilton's Measure	-0.0020	0.0009	0.0022	0.0064	0.0159				
Crude Materials PPI	0.0282	0.0648	0.1076	0.1518	0.2182				
Intermediate Mat'ls PPI	0.0169	0.0417	0.0803	0.1339	0.3926				
3 mo. Treasury Bill	-0.0004	-0.0018	-0.0039	-0.0054	-0.0032				
3 mo. Financial Paper	0.0039	0.0057	0.0003	-0.0050	0.0009				
10 year Gov't Bond	-0.0005	-0.0003	-0.0057	-0.0110	-0.0198				
Bond – T-bill Spread	0.0036	0.0081	0.0097	0.0095	0.0216				
FP – T-Bill Spread	0.0094	0.0228	0.0299	0.0353	0.0595				
Monetary Base	0.0273	0.0484	0.0701	0.0972	0.1991				
M1 (level)	0.0055	0.0162	0.0250	0.0445	0.0949				
M2 (level)	0.0193	0.0507	0.0873	0.1391	0.3940				
Exchange Rate	0.0104	0.0329	0.0666	0.1109	0.3049				
Ave. Hourly Earnings	0.0147	0.0434	0.0871	0.1511	0.5501				
Unemployment Rate	0.0031	0.0053	0.0089	0.0154	0.0875				
Capacity Utilization	0.0258	0.0486	0.0607	0.0548	0.0224				
S&P 500 Index	0.0153	0.0426	0.0843	0.1444	0.4327				

Table 2Root Mean Squared Error for PCE Forecasts, 1959:01 – 1998:12

**Notes:** Table reports difference between RMSE of the baseline model (no indicator) and RMSE of augmented model (model includes indicator listed the left-hand column). Negative numbers indicate those indicators that worsen the forecast relative to the baseline. PCE is measured by the monthly deflator for personal consumption expenditures.









































0.6

0.4

0.2

0

-0.05 |--0.2

0.6

0 0.2 0.4 Percent Improvement in RMSE

-0.1 -0.2

0.6

0 0.2 0.4 Percent Improvement in RMSE

-0.08

-0.06 -

SP500

ZZ M

Percent Improvement in RMSE

TWXR

R1∀

M1 MMERUR

SPRD

0

TWXR

**TBILL** 

0

PRB20K1Y

-0.05

TWXR

SPRD & 1)

-0.04

SRB91 TB PPIIM R10Y

SP500