Submitted for the November 1999 NBER Monetary Economics Conference
Preliminary and Incomplete Draft

# On the Identification of Monetary Policy: the "Price Puzzle" Reconsidered

by

Michael S. Hanson
*Wesleyan University*

Current version: September 20, 1999

**Abstract.** A large literature has used 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 the "exogenous" component of 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. A sub–sample investigation reveals that evidence of a price puzzle is associated primarily with the 1959 – 1979 sample period, and that none of the indicators studied — including commodity prices — resolve the puzzle over this period. Variation in the monetary policy reaction function across sub–samples may contribute towards this finding.

**Address.** Department of Economics, Wesleyan University, Middletown, CT, 06459-0007

# On the Identification of Monetary Policy: the "Price Puzzle" Reconsidered

Michael S. Hanson

## 1. Introduction

Over the past several years, a large number of researchers have estimated structural vector autoregressive (SVAR) models to establish "stylized facts" about U.S. economic fluctuations.[1] The results of these studies often are used to assess a given model of the business cycle or to discriminate among competing models. Many of these studies have focused on quantifying the effects of monetary policy. 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). Sims showed that this result disappears if commodity prices are included in the SVAR.[2] He proposed including commodity prices to serve as an "information variable" for the Federal Reserve, i.e. as an indicator of nascent inflation for the policy reaction function.

Despite numerous advances in SVAR modeling since then, the price puzzle has generally remained a problem for empirical researchers. Almost all subsequent research has adopted commodity prices as a necessary variable in monetary VAR models. As this tactic has evolved into a "conventional wisdom," many authors make only a passing reference to the problem commodity prices are intended to resolve. Nor is any *a priori* rationale given for including commodity prices in an otherwise parsimonious VAR model. And while the VAR system often is meant to correspond with a theoretical business cycle model, rarely do such theoretical models incorporate any role for commodity prices.

Other 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" has probably not been correctly identified.

Thanks to Robert Barsky, Allan Brunner, Kathryn Dominguez, Jon Faust, Phil Howrey, Erik Hurst, Michael Kiley, Lutz Kilian and Matthew Shapiro for helpful comments and suggestions. Additional thanks to seminar participants at the University of Michigan, the Federal Reserve Board of Governors, and the 1998 Meeting of the Midwest Economics Association (presented under the title "Identifying Monetary Policy: The Role of Commodity Prices"). All errors remain the author's sole responsibility.

[1] Some authors refer to these models as "identified VARs" rather than "structural VARs." Whichever nomenclature is used, the idea is the same: restrictions placed upon the model allow the researcher to imbue (some subset of) the estimated disturbances with a particular economic interpretation.

[2] Sims (1992) investigated both exchange rates and commodity prices as potential inflationary indicators for the monetary authority. Christiano, Eichenbaum and Evans (1996b) demonstrated that commodity prices alone appear to resolve the price puzzle across a number of SVAR modeling assumptions.

Proponents of this view include Sims (1996), Zha (1997), and Christiano, Eichenbaum and Evans (1998).[3] Viewed this way, the price puzzle goes to the very heart of empirical measurement of the effects of monetary policy. Clearly one should be cautious about drawing inferences about "stylized facts" from a misspecified model. The exclusion of a measure of future inflation is the particular misspecification cited.

This paper seeks to examine the empirical consistency of these explanations. While each offers some intuitive appeal, it is desirable to know whether these stories are supported by a more complete econometric investigation. Balke and Emery (1994) undertook an early study of some of these issues. They reported some support for both the above "information variable" view and for commodity prices serving as proxies for "supply shocks."[4] My results cast significant doubt on both of these proposed explanations. Note that if the "solution" to the price puzzle is not obtained for commonly accepted reasons, by implication the problem itself may be poorly understood. The results of this paper suggest the need for additional research into the role of commodity prices (and other inflation indicators) in the macroeconomy. Given the sensitivity of the effects of policy to this poorly understood modeling decision — particularly conclusions about the effects of monetary policy upon inflation — some of the "stylized facts" of the literature may need to be reconsidered.

I divide my investigation of these issues into two parts. First, I examine the ability of various potential indicators of inflationary pressure to resolve the price puzzle. This investigation allows me to determine 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. To foreshadow my findings, across a fairly broad set of indicators I find little relation between forecasting power and an ability to resolve the price puzzle.

Second, I examine the robustness of the existence and resolution of the price puzzle across sample periods. As will be discussed in detail below, the presence of a price puzzle may be a consequence not of excluding relevant variables from the model, but of misspecifying an unchanged monetary policy reaction function over a long sample period. Most observers agree that the Federal Reserve changed its operating procedures in 1979, and again by early 1983. Sub–sample estimation reveals that, relative to the post–1982 period, the pre–1979 period exhibited a siignificantly more pronounced price puzzle, greater variance in the monetary policy shocks, and less persistence of the federal funds rate. Indeed, the impulse responses for the early sample period suggest that on average, the estimated reaction function *reversed* the contractionary policy shocks within two and one–half years of their enactment. I also provide evidence that the persistence properties of inflation

[3] See also Sims and Zha (1998). More formal specification tests (though not specific to the price puzzle) have been proposed by Uhlig (1997) and Faust (1998).

[4] Christiano, Eichenbaum and Evans (1996b) propose that commodity prices may belong in the VARs to account for oil price "supply" shocks as well .

differ significantly between the two sample periods. Together these results point in the direction of an economically important change in the monetary policy reaction function. While beyond the scope of the current paper, such an implication would be consistent with mounting evidence that the reaction function before 1979 differs from that since.[5] Whether or not a change in policy is ultimately responsible for these findings, my results have two important implications for applied work in this literature. First, SVAR models estimated over more recent sample periods (in particular since 1983) generally do not exhibit statistically significant price puzzles. Second, the exclusion of "information variables" — including commodity prices — has no effect on the estimated responses to exogenous monetary policy shocks for this recent sample.

The next section of this paper discusses the monetary policy reaction function as an organizing framework. The recent literature is also reviewed. Identification is taken up in section 3, where estimated results from common models in the literature are presented. In section 4 I then investigate the inclusion of various candidate indicators of expected inflation in an SVAR model. Section 5 presents evidence that the price puzzle is a sample–specific phenomenon. Section 6 concludes.

## 2. The Identification of Monetary Policy

While I focus upon a particular empirical puzzle for much of this paper, the central issue is the empirical measurement of the macroeconomic effects of monetary policy. Monetary policy consists of two distinct yet interrelated components: a systematic or endogenous part (a "feedback rule") which summarizes how the monetary authorities respond to the state of the economy (past, present or expected future), and an exogenous part which captures shifts in policy objectives (or other features of policy that cannot be captured by the estimated rule).[6]

This distinction is important for two reasons. First, as is explained below, the conventional wisdom asserts that misspecification of the systematic part of policy is responsible for the price puzzle anomaly. Second, the literature typically focuses upon responses to exogenous monetary policy shocks only. In this paper I wish both to evaluate these "stylized facts" as well as to ask whether important variation in the endogenous component has occurred in the data.

[5] In Hanson (1999) I explore the evidence of change in the monetary policy reaction function in the context of an SVAR model. For other studies of a change in the reaction function, see Taylor (1997) and Clarida, Galí and Gertler (1998). De Long (1997) and Sargent (1998) investigate the pre–1979 policy situation in more depth.

[6] For the purposes of this investigation, the rule need only be interpreted as the best forecast of the actions of the monetary authorities conditional on the state of the economy, rather than as a behavioral relationship. For further discussion of this issue, as well as what policy "shocks" may actually measure, see Christiano, Eichenbaum and Evans (1998).

*The Monetary Policy Reaction Function*

In order to understand the conditions under which a price puzzle will be generated — or resolved — one first must consider how the effects of monetary policy are identified empirically. Much of the SVAR literature identifies one equation which can be interpreted as the monetary policy reaction function. In general terms, the reaction function of the monetary authority can be written as:

$$m_t = f(\mathcal{I}_t) + \mu_t \tag{1}$$

where $m_t$ represents the policy instrument(s) controlled by the monetary authority. $\mathcal{I}_t$ denotes the information set upon which the monetary authority can base its policy decisions. It includes the past history of all the variables in the system, as well as those variables assumed to be contemporaneously observed by the monetary authority. Thus, the $f(\cdot)$ function represents the systematic or endogenous component of monetary policy. The final term, $\mu_t$, represents the component of policy which is not a function of any current or past information contained in $\mathcal{I}_t$, and is in that sense "exogenous."[7] The SVAR approach allows a researcher to specify both the policy instrument and the form of the reaction function, and thereby obtain estimates of the exogenous shocks to monetary policy. These structural innovations cannot be observed directly in the data, but can be estimated once sufficient restrictions are imposed on the model.

The true information set $\mathcal{I}_t$ presumably contains a rich array of variables — many more than can be accommodated within a standard econometric model. An econometrician attempting to estimate equation (1) typically only observes some subset of the information utilized by the monetary authority: $\mathcal{H}_t \subset \mathcal{I}_t$.[8] Thus, the econometrician estimates

$$m_t = g(\mathcal{H}_t) + \eta_t \tag{2}$$

where $g(\cdot)$ is a (linear) function of the observed information. As some information will be excluded from the econometrician's estimated policy reaction function, the goal is to specify $\mathcal{H}_t$ so that $g(\mathcal{H}_t)$ approximates the "true" reaction function $f(\mathcal{I}_t)$ well.

The common view of the price puzzle posits that the exclusion of certain key variables — namely those which should contain information about future inflation — could result in

[7] Hall (1997) refers to this component as the "spontaneous random element" of policy. Unfortunately such a characterization has lead to some confusion about SVAR analysis: that it implies truly "random" — and by implication, irrational or sub–optimal — behavior by the monetary policy makers. Rather, the $\mu_t$ component signifies changes in policy that are unforecastable to agents who observe the information set $\mathcal{I}_t$; no normative implications for such policy actions follow from this specification.

[8] Parsimonious estimation generally will require a much smaller information set than $\mathcal{I}_t$. An example of data which would be difficult to incorporate into an econometric model is anecdotal evidence on the state of the economy brought to the FOMC meetings by regional Fed Bank presidents. On the other hand, the econometrician may have access to data revisions which preclude $\mathcal{H}_t$ from being a subset of $\mathcal{I}_t$.

a poor estimate of $g(\cdot)$. Suppose the excluded inflationary indicator variable signaled a rise in future inflation, to which the Fed endogenously reacted with contractionary policy. The econometrician, not observing the indicator variable, would (falsely) conclude that the Fed had undertaken an exogenous shift in policy. If the incipient inflation was realized before the effects of the policy change took place (due to an "outside lag" of policy), or if the Fed partially accommodated the inflation, then the econometrician would observe an increase in inflation following what he (falsely) concluded was an exogenous contractionary policy shock. This is the explanation of the price puzzle offered by Sims (1992): the confounding of endogenous and exogenous changes in monetary policy. The conventional wisdom states that an unforecastable (i.e. surprise) contractionary monetary policy shock — that is, a true $\mu_t$ shock — should not lead to an increase in inflation or the price level.[9]

In the above example, the change in the policy instrument — a contractionary response to indicators of higher future inflation — was codified into the reaction function, which itself was assumed constant. In the analogy of Taylor (1997), this would be a "movement along" as opposed to a "shift in" the reaction function. Yet the monetary authority certainly has the ability to change its reaction function. Failure to properly take into account such "regime changes" also could result in incorrectly measures endogenous and exogenous components of policy. A number of recent papers have suggested that the objectives of Federal Reserve policy have changed over the sample period commonly analyzed in the literature. Further discussion of this topic can be found in section 5.[10]

As an aside, note that the source of misspecification for both of these proposed explanations lies with the monetary policy reaction function. Alternatively the price equation may have been misspecified. For example, commodity prices may serve as a proxy for marginal cost considerations which are otherwise absent from the VAR. This interesting hypothesis is worthy of further consideration but is beyond the scope of this paper.

*Literature Review*

Through the 1980's and early 1990's, much of the debate in the VAR literature (and the money–income causality literature more generally) had centered around finding a single variable which measured exogenous changes in money supply, and therefore monetary policy. Numerous variations on the broad money aggregates commonly were used. Bernanke and Blinder (1992) countered by presenting evidence in favor of the federal funds rate

[9] But see Barth and Ramey (1999).

[10] I do not model here the decision to change the reaction function. Certainly the form of the reaction function is a choice variable of the monetary authorities, and it would be interesting to consider (a) what determines the menu of possible reaction functions available to the monetary authorities, and (b) what would induce the monetary authority to change the reaction function (in particular, whether such changes can be treated as exogenous).

as an appropriate measure of monetary policy. Moreover, they explicitly recognized that the federal funds rate is in large part determined endogenously. Separately, Sims (1992) demonstrated that shocks to various monetary aggregates resembled shocks to money demand rather than money supply, whereas shocks to short–term interest rates produced impulse responses that broadly matched the hypothesized effects of monetary policy, *provided the VAR model was augmented by certain "indicator variables"*. (He used commodity prices and the exchange rate.) Eichenbaum (1992) countered Sims's claims by proposing shocks to nonborrowed reserves as a suitable measure of monetary policy.

The federal funds rate and nonborrowed reserves are jointly determined in the market for bank reserves. Both Bernanke and Blinder (1992) and Eichenbaum (1992) recognized that institutional details of the reserves market could help to disentangle shocks to monetary policy from other shocks. The class of models that evolved from that investigation includes those developed by Strongin (1995) and Christiano, Eichenbaum, and Evans (1996a, 1996b). Bernanke and Mihov (1998) compare several models in this class within an encompassing framework. Further investigation has been undertaken by Leeper, Sims and Zha (1996), Christiano, Eichenbaum, and Evans (1998), and Bagliano and Favero (1998), to name but a few. [11] With the notable exception of Strongin (1995), all of the papers cited here include commodity prices in their estimation.

A central feature of all these models is the block recursive structure imposed to identify the monetary policy shocks and the reaction function.[12] All the models within this class also are "semi–structural," in that they identify the policy shocks while leaving the non–policy side of the model unrestricted. Christiano, Eichenbaum and Evans (1996b, 1998) confirm that a price puzzle arises in models in this class (when policy shocks are associated either with innovations to nonborrowed reserves or to the federal funds rate), and that the inclusion of commodity prices tends to eliminate the puzzle.[13]

Non–recursive semi–structural models also have been estimated in the literature. Recent examples include Gordon and Leeper (1994), Leeper (1995), Leeper, Sims and Zha (1996), and Sims and Zha (1998). Unlike the reserve market models above, these models tend to focus on broader monetary aggregates and achieve identification in part by positing an aggregate money demand function; variables which do not affect money demand are used to identify money supply (that is, monetary policy). Tellingly, all of these papers include commodity prices in their estimation.

[11] Similar models, though not encompassed by the Bernanke and Mihov specification, have been estimated by Cochrane (1994), Kim (1996) and Bernanke, Gertler and Watson (1997).

[12] Here "block recursive" is used in the sense of Keating (1996), who presents conditions under which such recursive estimates can be given a structural interpretation (examined in more detail below). Some authors, e.g. Cushman and Zha (1997), use the term "block recursive" for a different construct.

[13] See also Cochrane (1994).

In this paper I focus on models whose supply and demand shocks originate in the reserves market. These models of the economy have several features which recommend them for my purposes. First, restrictions in the reserves market are more closely related to actual Federal Reserve policy making and the institutional relationship between the Fed and the banking sector, and do not rely upon a posited aggregate money demand relationship. Second, these models constitute a popular class of models in the literature; in a chapter prepared for the forthcoming *Handbook of Macroeconomics*, Christiano, Eichenbaum and Evans (1998) make a case for their adoption.[14] Moreover, Bernanke and Mihov (1998) report that certain models in this class have strong empirical support. Most importantly, much research regarding empirical anomalies such as the price puzzle has been conducted within this framework. Thus it seems natural to start the analysis with these models.[15]

## 3. Empirical Measurement of Policy and the Price Puzzle

This section contains a more detailed discussion of how monetary policy is identified in models of the reserves market, and an examination of the empirical conditions under which the price puzzle will arise in these models. As the specification and estimation of structural VAR models is by now fairly standard, a brief presentation of the general model and notation follows. The discussion then turns to the identification schemes used in the remainder of this paper. This section concludes with the estimated impulse response functions for several "baseline" models: models which do not include any of the candidate inflation indicators.

*Overview of SVAR Estimation*

The SVAR problem described here is fairly standard. Consider the following representation of the *structural model* of the economy:

$$X_t = \Theta(L)\,\varepsilon_t\,. \tag{3}$$

$X_t$ is an $n$-dimensional vector of observed endogenous variables included in the model, and $\varepsilon_t$ a vector of unobserved structural disturbances. These are white noise primitives whose

[14] One reason for the popularity of these models undoubtedly is the ease of estimating recursive models in many statistical packages. However, Christiano, Eichenbaum and Evans (1998) make the stronger claim that the non–recursive models mentioned previously lead to similar inference about the effects of monetary policy as the recursive models examined below.

[15] It is worth reiterating that these models reflect current thinking on empirical modeling in macroeconomics and are not "straw men." It would be informative to conduct the analysis here on the models of Gordon and Leeper (1994) and Sims and Zha (1998); I plan to do so in subsequent work. An alternative class of models relies upon long–run neutrality properties to identify policy shocks and does not establish a policy reaction function *per se*. Thus they differ fundamentally from the models discussed above. See, e.g., Galí (1992), Johansen and Juselius (1994), Lastrapes and Selgin (1995) and Gerlach and Smets (1995). Faust and Leeper (1997) and Pagan and Robertson (1998) offer critiques of this class of models.

history accounts for the fluctuations of the endogenous variables. The monetary policy shock, $\mu_t$, is an element of $\varepsilon_t$. $\Theta(L)$ is a one–sided (infinite–order) matrix polynomial in the lag operator $L$.[16] Thus equation (3) is the vector moving average (VMA) representation of $X_t$ in the structural disturbances $\varepsilon_t$. The estimated $\Theta(L)$ will give the impulse response functions of interest. The covariance matrix for the structural disturbances is given by

$$E\left[\varepsilon_t \varepsilon_t'\right] = \Omega\,. \tag{4}$$

The structural shocks are assumed to be both mutually and serially uncorrelated, implying $\Omega$ is diagonal. It will be convenient to normalize the system so that $\Omega = I$.

Under standard regularity conditions, the model in equation (3) also has a structural vector autoregressive (SVAR) form:

$$\Phi(L)\, X_t = \varepsilon_t\,, \tag{5}$$

where $\Phi(L) = \Theta(L)^{-1}$ is also a polynomial in the lag operator. As in the usual simultaneous equations problem, the structural model of equation (5) cannot be directly estimated (except under very special conditions). Instead one must estimate the reduced–form system and then impose restrictions to identify the structural parameters. For this problem, one can estimate the following *reduced–form* VAR:

$$A(L)\, X_t = v_t\,, \tag{6}$$

where $v_t$ is a vector of $n$ reduced–form residuals. $A(L)$ is a lag polynomial in $L$; both $\Phi(L)$ and $A(L)$ are of order $q$, chosen to be large enough to capture the dynamics of the data. Thus, $A(L)$ can be written as

$$A(L) = I - A_1 L - \ldots - A_q L^q\,,$$

where $A_k$ is the $n \times n$ matrix of coefficients on $X_{t-k}$ — the $k^{th}$ lag of the variables in the system — and $I$ the $n \times n$ identity matrix. Simple rearrangement yields the standard practice of writing equation $i$ with $x_{it}$ as the dependent variable and with lags 1 through $q$ of each of the $n$ variables as (predetermined) regressors. In this case the simultaneity of the system is summarized by the covariance matrix,

$$E\left[v_t v_t'\right] = \Sigma\,, \tag{7}$$

in which some $\sigma_{ij}$ $(i, j = 1, \ldots, n;\ i \neq j)$ are not zero.

[16] The lag operator is defined as $Lx_t = x_{t-1}$.

Given the definitions and the normalization discussed above, the reduced–form VAR system (6) can be consistently estimated equation–by–equation via OLS, which yields estimates of the coefficient matrices, $\widehat{A_1}$ through $\widehat{A_q}$, and the covariance matrix, $\widehat{\Sigma}$. The estimated residuals from the VAR are forecast errors from the least–squares projections upon all the variables in the VAR (the information set). Since $v_t$ and $\varepsilon_t$ span the same linear space, identification consists of finding the $n \times n$ matrix $S$ such that

$$v_t = S\,\varepsilon_t\,. \tag{8}$$

Recall that the elements of $\varepsilon_t$ are orthogonal by construction.

Equation (8) implies the following relationship between the structural and reduced–form covariance matrices:

$$\begin{aligned}\Sigma = E\left[v_t v_t'\right] &= E\left[S\varepsilon_t\varepsilon_t' S'\right] \\ &= S\,\Omega\,S' \\ &= S\,S'\,.\end{aligned}$$

This follows from equations (4) and (7), and the previous normalization of $\Omega = I$. Recall that $\Sigma$ is an $n \times n$ symmetric matrix; it has $\frac{n(n+1)}{2}$ unique elements. Since there are $n^2$ parameters of $S$, the researcher must impose $\frac{n(n-1)}{2}$ restrictions to identify the model.

Equation (8) also allows the recovery of the structural parameters of $\Phi(L)$:

$$A(L)\,X_t = v_t = S\,\varepsilon_t = S\,\Phi(L)\,X_t\,.$$

It follows that

$$\Phi_0 = S^{-1}$$

and

$$\Phi_k = S^{-1}\,A_k$$

for $k = 1, \ldots, q$. I turn now to a more specific discussion of the identification of monetary policy within the class of models mentioned above.

*Identification of Monetary Policy in a SVAR*

The models estimated herein have a block recursive structure, in which the set of endogenous variables can be divided in non–policy and policy blocks. Let the vector of non–policy (or "activity") variables be denoted by $Y_t$. In the baseline specifications, $Y_t$ contains a measure of real output and an aggregate price index. Later this non–policy block will be augmented with one of a number of candidate inflation indicators.[17] Let the block of monetary policy variables be denoted by $M_t$. Each specification includes the

[17] The position of the indicator variable in the SVAR is not an innocuous choice, as discussed below.

same three variables determined within the reserves market: the federal funds rate, total reserves and nonborrowed reserves. (Identifying assumptions for the reserves market will be discussed below.) The SVAR of equation (5) can now be divided into policy and non–policy blocks and rewritten as

$$\begin{bmatrix} \phi_{YY0} & \phi_{YM0} \\ \phi_{MY0} & \phi_{MM0} \end{bmatrix} \begin{bmatrix} Y_t \\ M_t \end{bmatrix} = \begin{bmatrix} \Phi_{YY}(L) & \Phi_{YM}(L) \\ \Phi_{MY}(L) & \Phi_{MM}(L) \end{bmatrix} \begin{bmatrix} Y_t \\ M_t \end{bmatrix} + \begin{bmatrix} \varepsilon_t^Y \\ \varepsilon_t^M \end{bmatrix}. \tag{9}$$

Since

$$S = \begin{bmatrix} S_{YY} & S_{YM} \\ S_{MY} & S_{MM} \end{bmatrix} = \Phi_0^{-1},$$

the corresponding reduced form VAR model, corresponding to equation (6), then can be expressed as

$$\begin{bmatrix} Y_t \\ M_t \end{bmatrix} = S \begin{bmatrix} \Phi_{YY}(L) & \Phi_{YM}(L) \\ \Phi_{MY}(L) & \Phi_{MM}(L) \end{bmatrix} \begin{bmatrix} Y_t \\ M_t \end{bmatrix} + S \begin{bmatrix} \varepsilon_t^Y \\ \varepsilon_t^M \end{bmatrix}.$$

An essential identifying assumption is the predetermination of the non–policy block within a period. That is, policy is assumed to react to current innovations to the non–policy variables, while changes in policy only affect the aggregate activity variables — i.e. output and the price level (and later, an inflation indicator) — with a one–period lag. In the block recursive model, these restrictions imply $S_{YM} = 0$. Such timing restrictions are common in the literature. As they are only approximate descriptions of actual behavior, these restrictions are more plausible with higher frequency data.[18]

These restrictions are not sufficient to identify the policy shock $\mu_t$. Institutional features of the reserves market will motivate additional restrictions on the policy block, $S_{MM}$. Following the literature, I consider two candidate models: a federal funds rate targeting model ($m_t = \text{FF}$), and a nonborrowed reserves targeting model ($m_t = \text{NBR}$). The blocks which capture the effects of non–policy shocks, $S_{YY}$ and $S_{MY}$, are left unrestricted.[19] In this sense the specifications are semi–structural in nature: only the monetary policy shock is identified as a structural innovation.

In a federal funds rate targeting model, as in Bernanke and Blinder (1992) and Bernanke and Mihov (1998),[20] the Federal Reserve (perfectly) targets the funds rate, making the

[18] Feedback from policy into production and pricing decisions becomes more likely as the length of a period increases, and with "large" policy shocks. Notice that while the restrictions have a time–dependent interpretation, the economic justifications for these restrictions — often based on costs of adjustment — tend to be state–dependent in nature.

[19] Keating (1996) demonstrates that the ordering of variables in the non–policy block does not affect the responses to a policy shock.

[20] Bernanke and Mihov (1998) reinterpret the Bernanke and Blinder (1992) model to fit within their broader framework, which uses slightly different data. However, both impose restrictions that non–policy variables are predetermined within the period with respect to policy variables.

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. Changes in the funds rate would then be identified as policy shocks, whereas changes in the quantity of reserves would reflect shocks to both supply and demand. Many observers believe this solution to the identification problem to be a close approximation to actual Fed policy over most of the sample period examined below.

To make this model operational, the policy shock is identified by placing the funds rate first in the block of policy variables and imposing a lower triangular structure upon $S_{MM}$. Doing so permits the residual of the funds rate equation to be interpreted as the monetary policy shock. Subsequently I refer to this model as model "F".[21]

Alternatively, the Fed could target the quantity of reserves in the market by supplying reserves (perfectly) inelastically; in this case changes in reserves would be interpreted as policy changes (i.e. money supply shocks), whereas changes in the funds rate would confound both supply and demand shocks. Christiano, Eichenbaum and Evans (1996a) argue in favor of nonborrowed reserves as being under tighter control by the Fed than any of the broader measures of money. Below I refer to this specification as model "N". One might expect this model to perform particularly well in the 1979 to 1982 period, when the Federal Reserve eschewed interest rate smoothing to experiment with explicit targeting of nonborrowed reserves, but less well over other samples. The results reported by Bernanke and Mihov (1998) broadly confirm this intuition.

A variation on model "N" has been proposed by Strongin (1995), who identifies as the policy instrument that portion of nonborrowed reserves which is orthogonal to innovations in total reserves. That is, a contractionary policy shock shifts the "mix" away from nonborrowed reserves for a given quantity of total reserves.[22] He views innovations to total reserves as being determined exclusively by demand considerations which the Fed has no choice but to accommodate in the short run. While this assumption may be attractive for data measured bi–weekly (the length of the reserve settlement periods), it holds less weight at monthly (let alone quarterly) frequencies. On the other hand, Strongin's model does have an attractive feature: as the degree of demand accommodation goes to zero, Strongin's model becomes one of inelastic supply of nonborrowed reserves. In this sense, the estimated policy shock for model "N" is a special case of the shock from Strongin's

[21] A variant of this model has also been estimated by Christiano, Eichenbaum and Evans (1996a,b). They order the policy variables as {FF, NBR, TR}, whereas Bernanke and Mihov (1998) suggest ordering the policy variables as {FF, TR, NBR}. Notice that the order of the two reserves variables does not affect the estimation of the policy shock in a recursive model.

[22] Strongin uses the ratio of nonborrowed reserves and of total reserves to the previous period's level of total reserves, rather than the log level of each variable as in the other models. Following Strongin's naming convention, I refer to these transformed variables as NBRX and TRX, respectively.

model.[23] I refer to this specification as model "S".

This model also employs a recursive identification scheme for $S_{MM}$, with total reserves ordered first (to capture demand shocks to reserves), nonborrowed reserves second, and the federal funds rate last. This ordering implies that model "F" is *not* a special case of Strongin's model, even when the degree of accommodation approaches unity. In fact, the interpretation of the orthogonal shock to the funds rate itself is unclear in this model. Model "S" is predicated upon claim that the orthogonal shocks to nonborrowed reserves fully and uniquely measure the monetary policy shocks.

*Estimated Results for Baseline Models*

The first empirical results are for the baseline cases — estimated without any indicator variables — of the models discussed above. Figure 1 presents the estimated impulse response functions for model "F". The model is estimated monthly from 1959:01 to 1997:06, and 12 lags are included in each equation. Output is measured as industrial production and the price level by the consumer price index.[24]

For all the impulse response figures, the solid lines report the point estimates of the response of the variable listed at the top of a given column to the shock listed to the left of that row. The structural shock to monetary policy is a one–time increase in the federal funds rate. The responses are in percentage points except for the own response of the funds rate, which is in basis points. The dotted lines report 68% bootstrapped confidence bands, while the dashed lines represent 95% bootstrapped confidence bands.[25] While 95% error bands are becoming more common in the literature, I also report the 68% bands for comparison with earlier papers.[26]

[figure 1 here]

The third row of figure 1 shows the responses of each of the five variables to the monetary policy shock, which in this model is the shock to the federal funds rate. (As

[23] Technically, Strongin's model does not nest model "N" as both are just identified.

[24] Log levels of all variables (except the funds rate) are used in estimation. All of these series exhibit a pronounced upward trend over the sample period. While unit root tests on these series tend not to reject the null of non–stationarity, I follow the specification choices of the original authors and include these series in log levels instead of (log) differences. Bernanke and Mihov (1998) claim that there is little effect on their results if the various models are estimated in log differences instead of log levels. Sims, Stock and Watson (1990) suggest that estimation in (log) levels will be consistent (but not necessarily efficient).

[25] See Kilian (1998) for a discussion of these bias–corrected bootstrapped confidence intervals. 500 bootstrap replications where performed for each response.

[26] The bootstrap procedure used herein draws from the sample residuals and does not impose any distributional assumptions. The 68% and 95% bands would correspond to one– and two–standard error bands, respectively, under Gaussian assumptions.

discussed above, the other shocks in this model do not have a structural interpretation.) In general, a contractionary monetary policy shock depresses output, raises the funds rate and (mildly) depresses total reserves. These models implicitly assume a constant discount rate, so a contractionary policy shock which raises the funds rate also will encourage banks to shift from nonborrowed reserves into borrowed reserves, all else equal. This intuition is confirmed by the immediate (albeit temporary) fall in nonborrowed reserves.

The second column shows the response of the price level to a contractionary policy shock. The point estimate (the solid line) is *positive* for roughly three years. Moreover, this response is statistically greater than zero at the 68% confidence level for almost two years, a result which corresponds closely with prior findings (e.g. Sims, 1992; Christiano, Eichenbaum and Evans, 1996b). The qualitative nature of the responses is preserved at the 95% level: the price level (measured by the GDP deflator) rises significantly for about one year, and is indistinguishable from zero for the remaining years plotted.

Thus, while most the responses coincide with commonly–accepted views of how policy affects the variables of the model, the price response is a glaring anomaly. Most economists do not believe the price level should rise for a year or more in response to unanticipated contractionary shocks to monetary policy, and many would be surprised to find that policy shocks have no statistically discernible affect on the aggregate price level four years after the shock. This model fails the informal specification test discussed in the introduction.[27]

An alternative assumption to targeting the federal funds rate is targeting nonborrowed reserves. Figure 2 reports the impulse responses for model "S" (described above).[28] Since the restrictions are more plausible at higher frequencies, I estimate the model only with monthly data (consistent with Strongin, 1995). The fourth row, which shows the responses to orthogonalized innovations in nonborrowed reserves, represents the responses to a monetary policy shock in this model. Recall that innovations to total reserves (the third row) should reflect shocks to reserve demand, which the Fed accommodates. The policy shock is a one–time *decrease* in nonborrowed reserves, to reflect contractionary policy.[29]

[figure 2 here]

Similar to the effects of the policy shock in model "F", a contractionary policy shock

[27] Other researchers have used non–farm employment or interpolated GDP as the output measure. These have little effect upon the impulse responses. Alternative measures of the price level (such as the personal consumption expenditure deflator and the CPI excluding shelter) exhibit qualitative similar — albeit less dramatic — positive responses. Quarterly estimates using real GDP and the GDP deflator are virtually indistinguishable from the results in figure 1.

[28] Recall the transformation of the nonborrowed and total reserves variables, which prevents comparison of the responses of reserves across models. Strongin (1995), who proposes this transformation, reports similar responses of output and prices in a version of his model with logarithmic levels of NBR and TR.

[29] By contrast, figure 1 records the effect of a positive shock to nonborrowed reserves in the fourth row.

in model "S" significantly reduces output for several years (with about a six–month lag), reduces nonborrowed reserves and increases the federal funds rate (a liquidity effect). However, unlike the price response in model "F", the price level shows little evidence of an increase following a policy shock. In other words, identifying monetary policy shocks with orthogonalized innovations to nonborrowed reserves appears to eliminate the price puzzle. If a price puzzle serves as a specification test, these findings might suggest a preference for model "S" over model "F". Indeed, one possible interpretation is that the true source of misspecification is not excluding certain information variables but choosing the wrong variable as the policy instrument.[30]

Before accepting that conclusion it is worthwhile to examine the role of innovations to the federal funds rate in model "S". The fifth row of figure 2 reports responses to the "structural" federal funds rate shock: that portion of the innovation to the funds rate which is orthogonal to both the demand shock (equal to the innovation in total reserves, conditional on current activity variables) and the policy shock (the innovation in nonborrowed reserves which is orthogonal to the demand shock). Strongin (1995) does not offer any economic interpretation of this final shock. However, the impulse responses to a FF shock in model "S" mirror those of model "F". In particular, one observes a price puzzle nearly a year in length. Indeed, the reduced form innovations to the federal funds rate and nonborrowed reserves appear to be largely independent of each other, and so either orthogonalization implies similar "structural" shocks.

Not only does the orthogonalized funds rate shock in model "S" resemble a monetary policy shock, but there are reasons to believe it may have a policy interpretation, either instead of or in addition to the NBRX shock. For example, if the Federal Reserve actively tried to smooth interest rates, then it is likely that some portion of the FF shock in model "S" is in fact attributable to monetary policy. Moreover, as emphasized by Leeper, Sims and Zha (1996), models which target a reserve variable potentially imply dramatic swings in the federal funds rate in response to unexpected changes in discount window borrowing, which seems at odds with actual practice by the monetary authorities.[31]

The choice of which variable (and which set of identifying restrictions) best reflects policy is obviously important for assessing the role of monetary policy in the economy.

[30] This result has precedence in the literature. Eichenbaum (1992) proposed nonborrowed reserves as the policy instrument in response to Sims (1992), and found the price puzzle was greatly diminished. This is confirmed by Christiano, Eichenbaum and Evans (1996b, 1998). Strongin reports that his model does not exhibit a price puzzle, consistent with our findings. On the other hand, Leeper, Sims and Zha (1996) conclude that Strongin's model generates a significant price puzzle. Their version of Strongin's model differs from my model "S" in several respects, most importantly their use of a Bayesian estimation approach. They also use an interpolated measure of GDP instead of industrial production. Additionally, they include 6 rather than 12 lags in the VAR, and they use a slightly different sample: 1960:01 – 1996:03.

[31] Note that these criticisms apply with equal force to model "N." Leeper, Sims and Zha (1996) offer additional criticism of the reserves instrument approach.

Using the existence of a price puzzle as an informal specification test might suggest models which identify the policy instrument with the federal funds rate, rather than nonborrowed reserves, are misspecified. But at the same time, nonborrowed reserves shocks may not completely capture exogenous shifts in monetary policy. The recent SVAR literature has generally emphasized models which associate federal funds rate shocks with policy shocks. Another reason for examining a funds rate targeting policy is the recent attention given to interest rate rules, at the expense of (broad or narrow) money rules, in the theoretical literature. For all these reasons, plus the fact that the phenomenon I wish to examine is more pronounced in model "F", I focus on that model in the following section.

## 4. Inflation Indicators in SVAR Models

Although the models presented above offer a more rigorous foundation for the identification of monetary policy shocks relative to earlier (S)VAR models — via restrictions motivated by the market for bank reserves — they nonetheless tend to generate puzzling responses for prices. Sims (1992) suggested such puzzles were a consequence of incorrectly specifying the information set for the monetary authority's reaction function, and he proposed including additional information variables (namely commodity prices) in the SVAR as a way to resolve these puzzles. In this section I investigate the forecasting power of a large set of candidate indicators, and ask whether resolution of the price puzzle is related to forecasting power. Addressing this issue is one way to answer the question: Which variables belong in the (estimated) reaction function? Determining the answer is important for accurately separating policy into its endogenous and exogenous components.

Following Sims (1992) and Christiano, Eichenbaum and Evans (1996b), the information variable most commonly added in the the literature is commodity prices. Yet other variables might plausibly contain information about future inflation.[32] It is of course impossible to examine every conceivable indicator, but the results below suggest which broad classes of candidates are more likely to exhibit forecasting power than others. However, consideration of the channels through which a particular variable might be expected to help forecast inflation suggests a number of candidate indicators.

One prominent channel might be via a "chain of production" or "pass–though" effect: an increase in the costs of inputs or intermediate goods may lead to an increase in the prices of final goods. In this case, indicators which directly measure such costs play a causal role toward the aggregate price level. Alternatively, they could measure "supply shocks." Examples of such variables include producer price indices, including the price of oil and (perhaps) general commodity price indices, measures of labor costs and capacity

[32] 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.

measures. Import prices or exchange rates may also fit this story. Of course the shocks in question could be purely monetary, in which case a monetary aggregate might be an appropriate indicator.

A second channel exploits the distinction between flexible and temporarily fixed prices. If the aggregate price level adjusts sluggishly to various shocks, more flexible prices might respond initially to such shocks, followed only later by aggregate prices. In this case, the candidate indicator would signal a future increase in aggregate prices without necessarily feeding into them in a causal manner. This "informational" story is the basis of the classic exchange rate overshooting model of Dornbusch (1976). Other asset prices may perform in similar fashion, including interest rates (i.e. the prices of bonds). As many commodity prices are set in auction markets, they too may exhibit this behavior.

Notice that commodity prices plausibly could represent either channel (and these channels are not mutually exclusive). The results cited below suggest that the "pass–through" channel is probably not the most important. In particular, I find little evidence to support the "supply shock" story proposed by Balke and Emery (1994) and others. Table 1 lists the particular series I examine. This list of potential indicators is meant only to be representative, not exhaustive.[33]

*Forecasting Power of Candidate Indicators*

Table 1 reports the root mean squared error (RMSE) for a VAR–based forecast of the price level at various horizons. From equation (3), the forecast error for the price level $h$–periods ahead, conditional on information available at time $t$, can be written as:

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

where $\hat{p}_{t+h|t}$ represents the forecasted value of $p_{t+h}$ based upon data observed in period $t$ or earlier. $\theta_{\cdot 2k}$, a $1 \times n$ vector, is the second row of $\Theta_k$. (The price level always will be the second variable of $X_t$ in the models considered here.)

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

$$\begin{aligned}
\mathrm{MSE}(\hat{p}_{t+h|t}) &= E\left[(p_{t+h} - \hat{p}_{t+h|t})(p_{t+h} - \hat{p}_{t+h|t})'\right] \\
&= \theta_{\cdot 20}\,E[\varepsilon_t\varepsilon_t']\,\theta_{\cdot 20}' + \theta_{\cdot 21}\,E[\varepsilon_t\varepsilon_t']\,\theta_{\cdot 21}' + \ldots + \theta_{\cdot 2\,h-1}\,E[\varepsilon_t\varepsilon_t']\,\theta_{\cdot 2\,h-1}' \\
&= \theta_{\cdot 20}\,\theta_{\cdot 20}' + \theta_{\cdot 21}\,\theta_{\cdot 21}' + \ldots + \theta_{\cdot 2\,h-1}\,\theta_{\cdot 2\,h-1}'
\end{aligned}$$

[33] Cecchetti (1995) provides a complementary investigation of indicators of inflation. Christiano, Eichenbaum and Evans (1996a) remark in a footnote that oil prices did not resolve the price puzzle in their model while several commodity price indexes did. 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. Below I analyze those claims in greater detail.

since $E[\varepsilon_t \varepsilon_t'] = I$. The RMSE is simply the (positive) square root of MSE($\hat{p}_{t+h|t}$). The RMSE is more readily interpretable than the MSE, as the units of the former are in percentage terms (a consequence of estimating the VAR in log levels).

The first line of table 1 gives the RMSE of the baseline 5–variable VAR model, without any indicator variables. The remaining rows are for augmented 6–variable VARs with the indicator in the first column included as the third variable of $X_t$ (that is, the final variable of $Y_t$, the non–policy block). Variables with a lower RMSE are better indicators, in the sense that they reduce the forecast error of the price level relative to the baseline. The table reports RMSE for several horizons likely to be relevant for monetary policy; forecastability at horizons of one year or greater may be more important for policy makers, in light of the outside lag of monetary policy.[34]

Several interesting results appear in table 1. First, while all the reported indicators appear to reduce the RMSE relative to the baseline at a 3–month horizon, progressively fewer indicators do better than the baseline as the horizon increases in length. Thus, the indicators capture high frequency movement in the price level, but add little information beyond that contained in the own lags of prices as the horizon increases. Second, only a few indicators significantly improve upon the baseline case — producing RMSEs, say, ten percent below the baseline result. In other words, adding any one of these particular indicators does not dramatically improve the forecast of future prices at the horizons examined.[35] Finally, no single indicator (or class of indicators) produces the lowest RMSE at all horizons between three months and three years.

[Table 1 here]

Some additional results can be generalized across similar indicators. Broad commodity price measures are represented by three indexes: the Commodity Research Bureau spot index (CRB), the IMF overall commodity price index (IMF), and the price of sensitive materials (PSM). These series appear to outperform most other indicators tested at short horizons (a year or less) and still do well at a 2–year horizon. By 3 years, however, these broad commodity indexes actually worsen the forecast relative to the baseline case.[36]

Narrower commodity price indexes, such as those that focus on agricultural goods or metals, follow a similar pattern of doing well at short horizons but falling off as the horizon

[34] I have also tested whether each indicator Granger causes the price level in the context of the augmented VAR models. Although not reported here (in the interest of space), the results parallel those in table 1.

[35] Computational considerations precluded testing multiple indicators simultaneously.

[36] Prices of financial assets are probably even more flexible than commodity prices (which are set both on spot markets and by longer–term contracts), and therefore may be more likely to respond to information regarding future inflation. Consistent with this view, the S&P 500 stock market index performs well over 2– and 3–year horizons.

lengthens. The broad measures of commodity prices might be expected to outperform more narrowly defined measures, as the former tend to average out idiosyncratic shocks which do not impact the price level. There is not strong evidence for this hypothesis until individual commodities are examined. Consider oil prices, generally thought to be a main contributor to the inflation of the 1970s and early 1980s. Two different measures of oil prices — one an index from the IMF, the other the producer price index for petroleum — appear to have some incremental forecasting power at shorter horizons, but are noticeably weaker than the broader commodity price indexes. In fact, these oil price measures do not improve upon the baseline case noticeably.

Producer price indices for crude and intermediate materials are also included in table 1. As measured by the RMSE, the crude materials PPI resembles the broader commodity indexes — better at shorter horizons — although it generally underperforms them.[37] The intermediate goods PPI, on the other hand, seems to shine at the 2– and 3–year horizons. This index may be a better longer–term indicator for the aggregate price level because changes in intermediate goods prices get passed along a "chain of production" into the prices of final goods. Clark (1995), however, reports that the chain of production view is neither strongly nor systematically supported in sectoral data.

Perhaps the most natural indicators of future inflation would be interest rates. The expectations hypothesis of the term structure implies that a long–short interest rate spread should provide information about expected future inflation. However, the various spread measures analyzed here (the 10–year government bond rate less the 3–month Treasury bill, the 10–year bond rate less the 6–month commercial paper rate, and the commercial paper rate less the T–bill rate) do not appear substantially different 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. In table 1, levels and spreads perform about equally well, and do not significantly improve upon the baseline case.

Similarly, monetary aggregates do not provide much information about future prices by this metric. M2 does appear to play a more pronounced role at longer horizons, but M1 and the monetary base (M0) are indistinguishable from the baseline case. These days few economists would propose tracking a broad monetary aggregate for the purposes of forecasting inflation over the horizons considered here.[38]

Some other indicators of inflation actually reduce the RMSE of the price level more significantly: in particular, average hourly earnings and the exchange rate. These two measures outperform all other series we consider at the longer horizons. Average hourly

[37] Sims and Zha (1998) use the crude materials PPI in place of a broad commodity price index.

[38] Benjamin Friedman (1997), amongst others, has noted that the relationship between money and inflation has broken down during the last decade or so.

earnings may be capturing wage increases that ultimately are incorporated into final goods prices, similar to the evidence for the intermediate goods PPI. Interestingly, average hourly earnings perform much better than the unemployment rate (despite the Phillips curve). The nominal exchange rate was originally included with commodity prices in Sims's (1992) analysis of the price puzzle, but has not been included in more recent SVAR models as an inflation indicator. The explanatory power of the exchange rate is consistent with either the pass–through or informational channels outlined above.

*Alternative Indicators and the Price Puzzle*

I turn next to estimating "augmented" SVAR models which include one indicator from a subset of those listed in table 1. To match the literature, the indicator variable is included in the non–policy block, $Y_t$, after output and the aggregate price level. However, as discussed in section 3 above, this may not be the most appropriate way to incorporate these indicators into a SVAR model. This is especially true for informational variables that might enter the reaction function contemporaneously yet also respond to policy changes within the period. This practice could introduce a separate source of misspecification.

With that caveat in mind, figures 3a through 3d report the responses of the CPI price level, industrial production, and the federal funds rate to a contractionary policy shock in several augmented models. The first row in 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.

[figure 3a here]

The baseline case was presented above: the price level response is significantly positive for about 15 months, while output falls after about half a year and the funds rate remains above its initial value for a year. The augmented cases tend to look very similar to 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 response, but does not completely eliminate it. This finding of a residual price puzzle is replicated in several recent monthly studies, including Leeper, Sims and Zha (1996), and Christiano, Eichenbaum and Evans (1998), and contrasts with previous quarterly results (see Christiano, Eichenbaum and Evans, 1996b). They also appear to reduce the length of the interest rate response to a little over one–half year. The responses of these indicators themselves to the policy shock are also plausible, although the reduction in the commodity prices is barely statistically significant. Nonetheless, the magnitudes of their point estimates are dramatically larger than the CPI responses.

Figure 3b illustrates the impact of including each of several producer price indexes. Both the crude materials PPI and the intermediate materials PPI reduce the significant response of prices from about 16 months to around a year. Thus, these indicators are not as successful as the broad commodity indexes. The responses to a monetary policy shock are not statistically significant for either measure, and are much smaller in magnitude than the commodity price indexes shown in figure 3a. The responses of output and the funds rate are unaffected by the inclusion of any of these PPI series.

[figure 3b here]

Both commodity prices and the PPI measures reflect prices of inputs into final goods, the PPI probably more so than the broad commodity indexes with their large agricultural share. On the other hand, the broad commodity price indexes report prices set in auction markets, and so are likely more flexible than the PPI measures; thus they should more quickly incorporate inflationary pressures. Contrasting figure 3a with figure 3b suggests this latter property is more relevant for resolving the price puzzle. It follows that other asset prices, which also should immediately reflect inflationary shocks, should perform better than indicators operating primarily through a pass–through channel.

A case in point is the price of oil. Supply shocks in the form of large, rapid increases in the price of oil are commonly believed to be a major factor in the U.S. inflation experience.[39] The final row of figure 3b reports the results with a nominal oil price (the petroleum PPI) included in the VAR. As can be readily verified, the responses of the other variables in the VAR are practically unaffected by its presence. This result undermines claims that commodity prices — or, indeed, oil prices — belong in a SVAR model to control for supply shocks. (Stranger still, the response of the oil price to a monetary contraction is *positive* — albeit barely significant.) The sub-sample analysis of the next section will present further evidence against the supply shock story.

[figure 3c here]

Figure 3c considers interest rate indicators. In contrast with the PPI indexes of figure 3b, the long–short spread (defined as the 10–year bond less the 3–month T–bill rate) resolves the price puzzle about as well as commodity prices, reducing the length of the significant policy response to slightly over one–half year. (Whether this constitutes a true "resolution" of the price puzzle is debatable.) But table 1 indicates that the long–short spread is *not* a particularly good inflation indicator, especially in comparison to the broad commodity price indexes — let alone oil prices (or, as we will discuss in a moment, the exchange rate). Examining the components of the spread in the third and fourth rows shows

[39] See Bernanke, Gertler and Watson (1997) for a recent investigation of this issue.

all the action is coming from the short end of the market (the three–month Treasury bill): augmenting the model with a long rate doesn't effect the price puzzle or any of the other responses to policy. Note that after a year or so, policy shocks do lead to a statistically significant increase in the long rate.

In his 1992 paper on the price puzzle, Sims added both a commodity price and the exchange rate to his baseline VAR models. The impulse responses in figure 3d indicate why exchange rates are no longer routinely included in SVAR models for the U.S.: including this variable has no effect on the estimated price response.[40] (There also appears to be no statistically significant effect of policy on the exchange rate itself, although the point estimate implies an appreciation in response to a contractionary shock, as expected.) Yet in table 1 the exchange rate is one of the better indicators shown, particularly at longer horizons. Notably, the forecasting power of the exchange rate equals or exceeds the broad commodity price indexes at every horizon we examine. This is perhaps the strongest piece of evidence against the information view of commodity prices in SVAR models.

[figure 3d here]

As noted above, a measure of average hourly earnings forecasted as well as the exchange rate at long horizons; impulse responses with this variable included in the VAR are shown in the third row of figure 3d. Like several of the other indicators examined here, including average hourly earnings has almost no effect on any of the reported responses, including the price level. (Policy also has little measurable effect upon average hourly earnings.)

The unemployment rate also performs poorly as an inflation indicator variable. Looking at the final column, the impulse response of the unemployment rate closely resembles that for industrial production: no significant response for over one–half year, followed by a persistent, statistically significant worsening of the variable. Thus the unemployment rate and industrial production may be roughly equivalent measures of real activity in a parsimonious VAR model.

The main conclusion I draw from this investigation is that the ability of a given indicator to resolve a price puzzle is not correlated with its forecast power (as measured by the RMSE of the price level forecast).[41] Figures 4a through 4c summarize this relationship at various horizons (of 6 months, 1 year and 2 years, respectively). Each point on these figures plots the forecasting power of the variable indicated (relative to the baseline case; see table 1) versus its ability to reduce the positive price level response below its level in the baseline model. The horizontal axis gives the percentage reduction in the RMSE of the price level

[40] This statement excludes, of course, papers that explicitly examine the exchange rate response.

[41] One interpretation of this experiment is that there are only very small differences in the IRFs across a wide variety of candidate indicators.

forecast from including the indicator variable shown, relative to the baseline (no indicator) model (i.e. the baseline model RMSE less the augmented model RMSE, at the frequency specified for the figure). Thus a bigger gap implies more incremental forecasting power. The vertical axis shows the reduction in the average level of the impulse response function (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, for the number of periods specified).

Variables with strong forecasting power will be positioned towards the right side of each graph. Those which notably reduce the positive price level response will sit near the top of the graph. If stronger forecasting power implies a less pronounced puzzle, the points plotted should generally lie along a line from the upper right corner to the lower left. Few of the graphs reveal such a line clearly, and there are numerous outliers in the opposite corners. For example, the trade–weighted nominal exchange rate (TWXR) regularly illustrates substantial forecasting power without appreciably reducing the price puzzle. On the other hand, several of the interest rate series (most notably the 10–year bond to 3–month T–bill spread, SPREAD) generate a noticeable decline in the price level response without exhibiting hardly any forecasting power at all. Even some of the commodity price measures perform poorly along one or both measures.

[figures 4a through 4c here]

These scatterplots nicely contrast the results in figures 3a through 3d with the data in table 1, and reveal numerous cases which undermine the conventional wisdom. First, none of the interest rate measures examined here produced noteworthy improvements for the forecast of inflation relative to the baseline model, yet the long–short spread (and, arguably, the T–bill rate) appears to nearly match commodity prices in resolving the price puzzle. This finding suggests high forecasting power of an indicator is not a necessary property for it to significantly reduce the positive price level response. Contrast this conclusion with that of the three PPI variables in figure 3b: the lack of congruence between the reasonably strong forecasting power of these indicators and their inability to resolve the price puzzle also is not consistent with the conventional view of the price puzzle, and reveals that a sizable reduction in the RMSE of the price level is not sufficient to reduce the price puzzle. (Note that at longer horizons the PPI measures actually forecast better than most of the broader commodity price indexes.)

The conventional wisdom underlying the inclusion of commodity prices in a SVAR model — that they serve as indicators of future inflation — is difficult to reconcile with the above results. Some variables which offer significant forecasting power do not affect the price puzzle, while others that appear not to possess significant forecasting power nonetheless do resolve the puzzle (to a similar degree as commodity prices). Moreover,

the especially poor performance of oil prices calls into question the claim that commodity prices belong in the VAR as a proxy for supply shocks.[42]

In light of the findings of this section, proponents of the conventional view must offer an alternative explanation why commodity prices belong in the SVAR models. Commodity prices very well may be serving a fairly unique role in the VAR, but one that cannot be attributed primarily to an inflation forecasting role. In other words, without offering a compelling justification for including commodity prices in a VAR model, the literature does not reveal why the price puzzle arises in the first place. To better understand both the existence and resolution of the price puzzle, I turn next to a different possible source of misspecification: unaccounted variation in the monetary policy reaction function.

## 5. Robustness to Sub–sample Estimation

A number of authors have discovered that their empirical results are sensitive to the sample period over which their monetary SVAR models are estimated.[43] They recognize that the conduct of monetary policy, embodied in the reaction function, has not remained unchanged over the past forty years. In the literature, the most commonly cited change in monetary policy occurred in October 1979. This policy shift actually impacts upon my estimates along two dimensions: a change in the policy instrument ($m_t$ in equation (1)), and a change in the reaction function itself (the form of $f(\cdot)$).

First, identification of the policy shocks in the SVAR models I consider depends upon assumptions made about the policy instrument of the Federal Reserve. Between 1979 and 1982, the Fed shifted from effectively targeting the federal funds rate to explicitly targeting nonborrowed reserves. Thus it would be inappropriate to use model "F" in a period when the Fed more tightly controlled reserves and allowed the funds rate to fluctuate significantly, such as the 1979 to 1982 period.[44] For this reason, I focus my attention on the pre–1979 and post–1982 sub–samples.

Recall from the discussion in section 2 that assuming an invariant policy rule could lead to false inference about the role of the policy shocks in the economy. As mentioned in section 2, if the econometrician falsely estimates an unchanged reaction function, the estimated reaction function $g(\cdot)$ may not closely match the true reaction function $f(\cdot)$.

[42] The relative success of the long–short interest rate spread also is unlikely due to its usefulness as a proxy for supply shocks.

[43] Examples include Brunner (1994), Gordon and Leeper (1994), Strongin (1995), Bagliano and Favero (1998), and Bernanke and Mihov (1998).

[44] Bernanke and Mihov (1998) report that their version of model "F" does not fit this period very well. They also demonstrate that model "N" (of Christiano, Eichenbaum and Evans, 1996a,b) performs better over the 1979 – 1982 period than any other period they examine.

Such misspecification could contribute toward generating unexpected impulse responses, such as a price puzzle.

Second, although both of these sub–samples can be characterized as periods in which the funds rate was the policy instrument, the response of the funds rate to variables in the Fed's information set need not be uniform across these samples. In particular, several authors have recently argued that the Fed accommodated inflation to a much greater extent prior to 1979 than since 1982. This outcome may be a consequence of trying to exploit a Phillips curve or accommodating adverse supply shocks (particularly oil price shocks).[45] Taylor (1997) finds that, in contrast to the rule he estimated over the post–1982 period (see Taylor, 1993.) the estimated rule for the 1960 – 1979 period implied less than one–for–one adjustment of the nominal rate in response to increases in inflation, so that the real rate actually fell. Clarida, Galí and Gertler (1998) confirm these results with a more generalized version of a Taylor rule that incorporates interest rate smoothing and forward–looking behavior. Both studies conclude that the estimated rule for the 1960 – 1979 period led to an unstable inflation rate, and that sub–optimal policy was therefore responsible for the poor inflation performance during the pre–1979 period.

Whether or not the particular form of the monetary policy rule is the culprit, inflation does appear to be much more inertial in the pre–1979 period than in post–1982. Table 2 indicates how the stochastic nature of inflation has varied over the sample. It is well known that inflation was trending upwards between 1960 and 1979, and generally has been trending down since 1982. Unit root tests suggest that inflation has a highly persistent (integrated) component over the full sample, as well as in a sub–sample ending in October, 1979. In a sub–sample beginning in November, 1982, however, inflation appears to be much closer to stationary. Several test statistics that support this conclusion are reported in table 2. The first set of results reports the augmented Dickey–Fuller test statistic for the (logarithmic) CPI price level.[46] For each of the three monthly sample periods examined — 1959:01 to 1997:06, 1959:01 to 1979:10, and 1982:11 to 1997:06 — the null hypothesis of a unit root in the price level cannot be rejected. In each case, the largest autoregressive root is estimated to be in excess of 0.99. The second set of results uses the ADF test to test the inflation rate (equal to the differenced log price level). Here it is not possible to reject that inflation contains a unit root for the full 1959 – 1997 sample or the 1959 – 1979 sub–sample. The null can be (strongly) rejected for the 1982 – 1997 sub–sample, however.

It is generally known that these unit root tests have low power, especially against

[45] De Long (1997) and Sargent (1998) propose that the temptation to try to exploit a Phillips curve relationship led the Fed to pursue policies in the 1960s and 1970s that induced inflation to accelerate. Bernanke, Gertler and Watson (1997) suggest that the endogenous response of monetary policy to the oil price shocks is responsible for the subsequent inflation experience rather than the shocks themselves.

[46] The number of lags in the test are chosen by the Schwarz Information Criterion.

local alternatives. But the estimated values of the largest autoregressive root of inflation are consistent with the conclusion that inflation does not contain a unit root over the latter sub–sample: the estimated root is only 0.42 since 1982, whereas it exceeds 0.9 for both other samples. Second, testing a null of stationarity provides (weak) confirmation of these conclusions. The unit root test of Leybourne and McCabe (1994) indicates very strong rejection of the null that inflation is stationary in both the 1959 – 1997 and 1959 – 1979 sample periods. The rejection is not nearly so strong for the 1982 – 1997 period: the test statistic is 0.667 versus a critical value of 0.453 at the 5% level. So while not overwhelming, these results point towards the same conclusion: inflation appears to be very persistent when measured over the full sample, but closer to mean reverting in the latter sub–sample.[47]

Recall that the price puzzle is a pronounced, *sustained* increase in prices following a contractionary monetary policy shock. Anything which leads to more persistence in the process for prices will exacerbate any tendency for a price puzzle to occur in the data. Given this sub–sample variation in the stochastic process of the price level, it seems prudent to allow the coefficients of the reaction function to vary across the sub–samples as well. Whether the form of the monetary policy reaction function is ultimately responsible for the greater persistence in the early sample, or merely a reflection of it, is still an open question.[48]

That the reaction function has changed can be seen by comparing the impulse responses across the two sub-periods. Figure 5 replicates the baseline model "F" for the full sample and sub–samples covering 1959:01 – 1979:10 and 1982:11 – 1997:06. The break dates are chosen to correspond with the period of experimentation with nonborrowed reserves targeting and thus match those used elsewhere in the literature.[49] A comparison down the second column makes it clear that the price puzzle is a regime–specific phenomenon: the price level response is not statistically significant in the post–1982 sub–sample, while

[47] The results for 1960 – 1979 correspond to those reported in Barsky (1987). The conclusions from all of these unit root tests are unchanged when alternative price level series are examined (namely the personal consumption expenditure (PCE) deflator and the CPI excluding shelter). The largest autoregressive roots are also estimated to have similar magnitudes.

[48] The recent literature cited here implies that a shift in the nature of the reaction function — a "regime change" — could have caused the apparent change in the stochastic process of inflation. Such a conclusion is not unprecedented: Mankiw, Miron and Weil (1987) claimed that the policies introduced by the Fed after its founding changed the stochastic process describing nominal interest rates. Barsky (1987) found that the stochastic process for inflation has varied significantly over the past century, and its degree of persistence was correlated with different monetary regimes. While the 1979 "regime shift" did not involve a change in institutions, a significant modification to the form of the reaction function — as suggested by Taylor (1997) and Clarida, Galí and Gertler (1998) — could have resulted in a change in the dynamics of inflation. This is a question for future research.

[49] See Strongin (1995) and the references cited within for a lengthy discussion of changes in the Federal Reserve's monetary policy reaction function since 1959.

the pre–1979 puzzle is quite protracted (nearly twice as long as in the full sample) and relatively large in magnitude.[50] Notice that each of these graphs plots responses to an initial 50 basis point increase in the federal funds rate.

There are other important differences between the sub–sample results. 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. Part of this pattern is likely attributable to an apparent reversal of policy: the funds rate actually *falls significantly* about two to three years following the initial contractionary shock. Contrast this behavior with that estimated for the full sample — in which the funds rate declines (more or less) smoothly back to zero a little over a year after the initial shock, and that for the later sub–sample, in which the funds rate shocks are much less persistent (remaining positive for about a year). Consistent with this lower persistence, policy shocks do not have statistically significant effects upon industrial production or the consumer price index, as the final row of figure 5 illustrates.

This final result should not be interpreted as meaning monetary policy has not been an important factor in the macroeconomy since 1982. On the contrary, this observation reinforces the importance of the reaction function — that is, the "endogenous" component of policy — vis–à–vis the "exogenous" policy shocks. The role of the reaction function has been under–appreciated in the SVAR literature, where it is still common to conclude that monetary policy must not be important if the policy *shocks* are judged to be inconsequential for explaining variation in output or prices. As has been argued elsewhere, this conclusion is simply not warranted: even in the absence of unanticipated policy shocks, changes in the reaction function would change the processes for other variables of interest in the system.[51] However this result does suggest that under Volcker and Greenspan, the Fed has not deviated dramatically from the estimated policy rule. As a result, it may be difficult to infer the effects of policy shocks when they are not an important source of variation in this sub–sample.

Additionally, to the extent that variation across the two sub–sample periods represents variation in the reaction function, my results indicate such changes are at least as important a factor in the price puzzle as indicators of future inflation. Figure 6 reproduces the scatterplot of figure 4b (one year horizon), but for the 1959 –1979 sub–sample only. Notice that the scale of the vertical axis is nearly *one–hundredth* that of figure 4b: incorporating any potential inflation indicator — including commodity prices — has a minimal effect upon the positive price response. Were these points plotted on figure 4b, they would

[50] Qualitatively similar results were obtained with other price level variables, and with quarterly data.

[51] Bernanke, Gertler and Watson (1997) illustrate the importance of the Fed's reaction to oil price shocks for understanding the U.S. inflation experience. Sims and Zha (1998) also investigate changes in the endogenous response of policy to the state of the economy.

appear to lie on a horizontal line: there is no discernible reduction in the price puzzle regardless of how much the RMSE is reduced. (The scales of the horizontal axis are comparable across the two figures.)[52]

Considering the large price puzzle illustrated in figure 5 in conjunction with the paltry impulse response reductions of figure 6 suggests that regardless of the indicator variable included in the augmented model, a pronounced, statistically significant price puzzle still exists for this sub–sample. Figure 7 makes this fact more explicit by plotting a few augmented models. The second row reports the impulse responses when commodity prices (the price of sensitive materials) are included. The price puzzle is not as large as in the baseline (5 variable) case, but is still statistically significant for more than one and one–half years. Augmented models with the other broad commodity price indexes considered above (the CRB index and the IMF index) produce even longer price puzzles.

The third row of figure 7 shows the impulse responses when the long–short interest rate spread is used as an indicator variable. Reflecting figure 6, the price puzzle is smaller in this case than with commodity prices. Still a sizable puzzle (well over a year) remains. The fourth row reveals that, as in the full sample case, including the exchange rate does not resolve the price puzzle, despite its larger relative improvement in the RMSE of the price level forecast.

The final row reports the consequences of including oil prices (the petroleum PPI) in the VAR. If the need to control for supply shocks were ever necessary, it would be so for the 1959 – 1979 period. However, despite the large oil price shocks that occurred over this period, including the oil price to capture this effect reveals *no discernible effect* on the response of the price level (or any other variable). For the sample period in which the supply shock rationalization should be most appropriate, neither commodity prices nor oil prices are able to resolve the sizable estimated price puzzle.

For completeness of comparison, figure 8 illustrates the same impulse responses as figure 7, but for the later 1982 – 1997 sub–sample. In general, the 1982 – 1997 results do not show any significant price puzzle. While the inclusion of several indicators, including commodity prices, makes the drop in output (marginally) significant, they have at best no effect on the price level response. In some cases, including one of the candidate indicator variables actually worsens the point estimate of the price response, and tends to lead to wider confidence intervals.

The clear implication of all of these estimates is that the cross–sample variation in the endogenous component of monetary policy is a more important omission than any candidate indicator variable. This is not to say that additional variables do not belong in the VAR, only that the justification for their inclusion cannot be based on resolving the

[52] Excluded for brevity, the 6 month and 2 year scatterplots for 1959 – 1979 (which correspond to figures 4a and 4c, respectively) look very similar to figure 6.

price puzzle. Indeed, the sub–sample results further undermine the conventional wisdom that commodity prices belong in a SVAR model because they are proxies for future inflation or supply shocks. The 1959 – 1979 period features both oil price shocks, so if supply shocks were the key underlying factor in the price puzzle, one should expect a large puzzle over this sample which is resolved by the inclusion of commodity prices (or, arguably even better, oil prices). While a very significant price puzzle is observed in the data for this period, there is little evidence to support the supply shock story. Rather, the evidence cited here is more consistent with important variation in the monetary policy reaction function over time, a view which has often been ignored in the literature.

## 6. Conclusion

This essay highlights the importance of accounting for changes in the endogenous component of monetary policy — the policy reaction function — when attempting to measure the significance of monetary policy upon the economy. Two general implications for the broader monetary SVAR literature are discussed: first, one should not focus solely on the estimated policy shocks (the "exogenous" component of policy) when attempting to understand the role of monetary policy for economic fluctuations. Indeed, changes in the reaction function (the "endogenous" component of policy) are often more important. Second, the exogenous policy shocks themselves cannot be properly identified without taking into account the changes over time in the reaction function.

I examine a common puzzle that arises in earlier SVAR models in the empirical monetary policy literature: a significant, protracted increase in aggregate prices following what the researcher has labeled a contractionary monetary policy shock. The conventional wisdom, following a suggestion by Sims (1992), is to include commodity prices in the VAR as an indicator of future inflation. The results in section 4 demonstrate the difficulties with that proposed explanation of the price puzzle: there does not appear to be a systematic relationship between the ability of a given indicator to forecast inflation (as measured by the RMSE of the price level forecast from the VAR) and its ability to prevent a price puzzle in standard SVAR models. My results also suggest that commodity prices do not proxy for supply shocks.

In section 5 I investigate how variation in the reaction function impacts the estimated impulse response functions in a SVAR model. The results are quite striking: the existence of a price puzzle is strongest in the 1959 – 1979 sample, and is not resolved by any indicator variables, including commodity prices. On the other hand, there does not appear to be much evidence of a price puzzle since 1982. These results provide further evidence that proxies for supply shocks or other missing information are not the *primary* source of misspecification in the SVAR literature. Rather, failure to account for a change in the Federal Reserve's reaction function may have led to the more fundamental causes of the

puzzle being obscured.

This project suggests several directions for further research. First, economic conclusions in these models are intricately related to the way in which they are identified. As discussed above, the common practice of including commodity prices — or other forward–looking informational variables, such as interest rates or asset prices — as predetermined with respect to the current period's monetary policy changes is both undesirable and untenable. Future work in the SVAR literature must address the joint endogeneity between these variables and monetary policy. This is clearly a limitation of the recursive identification schemes employed by the models I examine, although Christiano, Eichenbaum and Evans (1998) point out that non–recursive models to date still impose similar restrictions.

Second, commodity prices arguably belong in a SVAR model, although not necessarily for reasons most commonly believed. While this essay cannot answer definitively what special role commodity prices are playing in the VAR, it highlights the importance of investigating the role of commodity prices in the macroeconomy more generally. A theoretical investigation may be able to motivate the empirical findings reported herein. I intend to pursue this avenue in related work.

Finally, the results in section 5 are consistent with the view that the nature of the monetary policy reaction function may have led to unstable behavior of some important macroeconomic aggregates, including the price level. In assessing the role of monetary policy in economic fluctuations, it is desirable to determine the extent to which suboptimal policy ultimately may be responsible for the observed patterns in the data. Policy makers and academic economists alike should be interested in learning the sources behind the economic outcomes in the pre–1979 period.

## References


Bagliano, Fabio C. and Favero, Carlo A. 1998. "Measuring Monetary Policy with VAR Models: An Evaluation." *European Economic Review*. 42. 1069–1112.

Balke, Nathan S. and Kenneth M. Emery. 1994. "Understanding the Price Puzzle." Federal Reserve Bank of Dallas *Economic Review*. 4th quarter. 15–26.

Barsky, Robert B. 1987. "The Fisher Hypothesis and the Forecastability and Persistence of Inflation." *Journal of Monetary Economics*. 19:1. 3–24.

_____ and Lutz Kilian. 1998. "Money, Stagflation, and Oil Prices: A Re–Interpretation." manuscript. University of Michigan. February.

Barth, Marvin J. and Valerie Ramey. 1999. "The Cost Channel of Monetary Transmission." manuscript. August.

Bernanke, Ben S. and Alan S. Blinder. 1992. "The Federal Funds Rate and the Channels of Monetary Policy." *American Economic Review*. 82:4. 901–921.

_____ , Mark Gertler and Mark Watson. 1997. "Systematic Monetary Policy and the Effects of Oil Price Shocks." in W. C. Brainard and G. L. Perry, Eds. *Brookings Papers on Economic Activity*. 1997:1. 91–157.

_____ and Ilian Mihov. 1998. "Measuring Monetary Policy." *Quarterly Journal of Economics*. 113(3). 869–902.

Brunner, Allan. 1994. "The Federal Funds Rate and the Implementation of Monetary Policy: Estimating the Federal Reserve's Reaction Function." Board of Governors of the Federal Reserve System International Finance Discussion Paper No. 466. May.

Cecchetti, Stephen G. 1995. "Inflation Indicators and Inflation Policy." *1995 NBER Macroeconomics Annual*.

Christiano, Lawrence J., Martin Eichenbaum, and Charles Evans. 1996a. "The Effects of Monetary Policy Shocks: Evidence from the Flow of Funds" *Review of Economics and Statistics*. 78:1. 16–34.

_____ . 1996b. "Identification and The Effects of Monetary Policy Shocks," in *Financial Factors in Economic Stabilization and Growth*. Eds: Mario I. Blejer, Zvi Eckstein, Zvi Hercowitz, and Leonardo Leiderman. Cambridge University Press: New York. 1996. 36–74.

_____ . 1998. "Monetary Policy Shocks: What Have We Learned and to What End?" NBER Working Paper No. 6400. February.

Clarida, Richard, Jordi Galí, and Mark Gertler. 1998. "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory." National Bureau of Economic Research Working Paper No. 6442. March.

Clark, Todd E. 1995. "Do Producer Prices Lead Consumer Prices?" Federal Reserve Bank of Kansas City *Economic Review*. 80:3. 25–39.

Cochrane, John H. 1994. "Shocks." *Carnegie Rochester Conference Series on Public Policy*. 41. 295–364.

Cushman, David O. and Tao Zha. 1997. "Identifying Monetary Policy in a Small Open Economy under Flexible Exchange Rates." *Journal of Monetary Economics*. 39(3). 433–48.

De Long, J. Bradford. 1997. "America's Peacetime Inflation: The 1970's." *Reducing Inflation*. Eds: Christina D. Romer and David Romer. NBER: Cambridge, MA.

Eichenbaum. Martin. 1992. "Comment on 'Interpreting the Macroeconomic Time Series Facts: The Effects of Monetary Policy.'" *European Economic Review* 36. 1001–1011.

Faust, Jon. 1998. "The Robustness of Identified VAR Conclusions about Money." *Carnegie–Rochester Conference Series on Public Policy*. forthcoming.

_____ and Eric Leeper. 1997. "When Do Long–Run Identifying Restrictions Give Reliable Results?" *Journal of Business and Economic Statistics*. 15:3. 345–353.

Friedman, Benjamin M. 1997. "The Rise and Fall of Money Growth Targets as Guidelines for U.S. Monetary Policy" in *Towards More Effective Monetary Policy*. Ed: Iwao Kuroda. New York: St. Martin's Press. 137–164.

Galí, Jordi. 1992. "How Well Does the IS–LM Model Fit Post–War U.S. Data?" *Quarterly Journal of Economics*. 92. 709–738.

Gerlach, Stefan and Frank Smets. 1995. "The Monetary Transmission Mechanism: Evidence from the G–7 Countries." Centre for Economic Policy Research Discussion Paper, no. 1219. February.

Gordon, David B. and Eric M. Leeper. 1994. "The Dynamic Impacts of Monetary Policy: An Exercise in Tentative Identification." *Journal of Political Economy*. 102:6. 1228–1247.

Hall, Robert. 1996. Comment on Leeper, Sims and Zha, "What Does Monetary Policy Do?" in W. C. Brainard and G. L. Perry, Eds. *Brookings Papers on Economic Activity*. 1996:2. 64–68.

Hanson, Michael S. 1999. "Evaluating U.S. Monetary Policy Regimes in a Structural VAR Context." manuscript. June.

Johansen, Soren and Katarina Juselius. 1994. " Identification of the Long–Run and the Short–Run Structure: An Application to the IS–LM Model." *Journal of Econometrics*. 63:1. 7–36.

Keating, John W. 1996. "Structural Information in Recursive VAR Orderings." *Journal of Economic Dynamics and Control*. 20: 9–10. 1557–1580.

Kilian, Lutz. 1998. "Small Sample Confidence Intervals for Impulse Response Estimates." *Review of Economics and Statistics*. 80(2). 218–230.

Kim, Soyoung. 1996. "Does Monetary Policy Matter in the G–6 Countries? Using Common Identifying Assumptions about Monetary Policy across Countries." manuscript. Yale University. March.

Kozicki, Sharon. 1997. "Predicting Real Growth and Inflation with the Yield Spread." Federal Reserve Bank of Kansas City *Economic Review*. Fourth quarter. 39–57.

Kwiatkowski, Denis, Peter C. B. Phillips, Peter Schmidt, and Yongcheol Shin. 1992. "Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root." *Journal of Econometrics*. 54. 159–178.

Lastrapes, William D. and George Selgin. 1995. "The Liquidity Effect: Identifying Short–Run Interest Rate Dynamics Using Long–Run Restrictions." *Journal of Macroeconomics*. 17:3. 387–404.

Leeper, Eric M. 1995. "Reducing Our Ignorance about Monetary Policy Effects." Federal Reserve Bank of Atlanta *Economic Review*. July/August. 1–38.

_____, Christopher A. Sims and Tao Zha. 1996. "What Does Monetary Policy Do?" in W. C. Brainard and G. L. Perry, Eds. *Brookings Papers on Economic Activity*. 1996:2. 1–63.

Leybourne, S. J. and B. P. M. McCabe. 1994. "A Consistent Test for a Unit Root." *Journal of Business and Economic Statistics*. 12:2. 157–166.

Mankiw, N. Gregory, Jeffery A. Miron and David N. Weil. 1987. "The Adjustment of Expectations to a Change in Regime: A Study of the Founding of the Federal Reserve." *American Economic Review*. 77:3. 358–374.

Pagan, Adrian R. and Robertson, John C. 1998. "Structural Models of the Liquidity Effect." *Review of Economics and Statistics*. 80:2. 202–217.

Sargent, Thomas. 1998. *The Conquest of American Inflation*. manuscript. Hoover Institution.

Sims, Christopher A. 1989. "Models and Their Uses." *American Journal of Agricultural Economics*. 71:2. 489–494.

_____ . 1992. "Interpreting the Macroeconomic Time Series Facts: The Effects of Monetary Policy." *European Economic Review* 36. 975–1000.

_____, James H. Stock and Mark W. Watson. 1990. "Inference in Linear Time Series Models with Some Unit Roots." *Econometrica*. 58:1. 113–144.

_____ and Tao Zha. 1998. "Does Monetary Policy Generate Recessions?" Federal Reserve Bank of Atlanta Working Paper 98-12. July.

Strongin, Steven. 1995. "The Identification of Monetary Policy Disturbances: Explaining the Liquidity Puzzle." *Journal of Monetary Economics*. 35:3. 463–497.

Taylor, John B. 1993. "Discretion Versus Policy Rules in Practice." *Carnegie–Rochester Conference Series on Public Policy*. 39. 195–214.

_____ . 1997. "An Historical Analysis of Monetary Policy Rules." manuscript. Stanford University and NBER. December.

Zha, Tao. 1998. "Identifying Monetary Policy: A Primer" *Federal Reserve Bank of Atlanta Economic Review*. 82:2. 26–43.

**Table 1: RMSE for CPI Forecasts, 1959:01 – 1997:06**

| | *Forecast Horizon* | | | | |
|---|---|---|---|---|---|
| *Indicator* | *3 mo.* | *6 mo.* | *1 yr.* | *2 yr.* | *3 yr.* |
| Baseline | 0.5118 | 0.7939 | 1.4074 | 2.6629 | 3.7713 |
| | | | | | |
| *Percent improvement over Baseline:* | | | | | |
| CRB Spot Index | 0.0567 | 0.1128 | 0.1704 | 0.0484 | -0.1479 |
| CRB Raw Industrials | 0.0575 | 0.1155 | 0.1783 | 0.0623 | -0.1542 |
| CRB Foodstuffs | 0.0254 | 0.0358 | 0.0273 | -0.1096 | -0.1862 |
| Price of Sensitive Mat'ls | 0.0660 | 0.1363 | 0.2608 | 0.2131 | -0.0018 |
| Gold Price | 0.0290 | 0.0672 | 0.1727 | 0.2413 | 0.1221 |
| IMF Overall Index | 0.0464 | 0.1044 | 0.1885 | 0.1051 | -0.0820 |
| IMF Foodstuffs | 0.0318 | 0.0679 | 0.0980 | -0.0471 | -0.1795 |
| IMF Agr. Raw Mat'ls | 0.0704 | 0.1653 | 0.2881 | 0.2313 | -0.0203 |
| IMF Metals | 0.0171 | 0.0390 | 0.0412 | -0.0769 | -0.1780 |
| IMF Oil Index | 0.0113 | 0.0057 | -0.0129 | -0.0440 | -0.0433 |
| | | | | | |
| Petroleum PPI | 0.0156 | 0.0352 | 0.0618 | 0.0922 | 0.0204 |
| Crude Materials PPI | 0.0288 | 0.0498 | 0.0850 | 0.0314 | -0.0158 |
| Intermediate Mat'ls PPI | 0.0239 | 0.0420 | 0.0947 | 0.2635 | 0.4591 |
| | | | | | |
| 3 mo. Treasury Bill | 0.0063 | 0.0015 | 0.0091 | 0.0198 | -0.0135 |
| 6 mo. Commercial Paper | 0.0064 | -0.0025 | -0.0205 | -0.0483 | -0.0538 |
| 10 year Gov't Bond | 0.0121 | 0.0172 | 0.0213 | 0.0148 | -0.0194 |
| Bond – T-bill Spread | 0.0102 | 0.0237 | 0.0275 | 0.0241 | -0.0138 |
| Bond – CP Spread | 0.0105 | 0.0200 | 0.0225 | 0.0360 | 0.0316 |
| CP – T-Bill Spread | 0.0122 | 0.0304 | 0.0445 | 0.0453 | 0.0272 |
| | | | | | |
| Monetary Base | 0.0127 | 0.0114 | 0.0137 | -0.0155 | -0.0360 |
| M1 (level) | 0.0141 | 0.0234 | 0.0291 | 0.0182 | 0.0118 |
| M2 (level) | 0.0132 | 0.0335 | 0.0804 | 0.1902 | 0.3126 |
| | | | | | |
| Exchange Rate | 0.0321 | 0.0890 | 0.2597 | 0.6241 | 0.7403 |
| | | | | | |
| Ave. Hourly Earnings | 0.0138 | 0.0328 | 0.1092 | 0.4088 | 0.9194 |
| Unemployment Rate | 0.0034 | 0.0002 | -0.0040 | 0.0012 | 0.1149 |
| Capacity Utilization | 0.0183 | 0.0435 | 0.0654 | -0.0085 | -0.0877 |
| S&P 500 Index | 0.0080 | 0.0253 | 0.0763 | 0.2431 | 0.4344 |

Note: 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 indicators that worsen the forecast relative to the baseline.

## Table 2: Unit Root Tests for Inflation

Tests of Consumer Price Index

| | 1959:01 – 1997:06 | 1959:01 – 1979:10 | 1982:11 – 1997:06 |
|---|---|---|---|
| *ADF Test: $H_0$: price level is I(1)* | | | |
| $Z_t$ test statistic | -0.9019 | +4.4675 | -1.6631 |
| Largest AR root | 0.9999 | 1.0035 | 0.9986 |
| No. lags used in test | 10 | 3 | 2 |
| *ADF Test: $H_0$: inflation is I(1)* | | | |
| $Z_t$ test statistic | -2.0833 | -0.9931 | -7.0787*** |
| Largest AR root | 0.9185 | 0.9492 | 0.4203 |
| No. lags used in test | 9 | 6 | 1 |
| *Leybourne–McCabe Test:* | | | |
| *$H_0$: inflation is I(0)* | | | |
| $\eta_\mu$ test statistic | 6.8596*** | 8.4745*** | 0.6669** |
| No. lags used in test | 10 | 10 | 1 |

*Notes*: "Price level" is logarithm of CPI; "inflation" is difference of log price level. *** denotes significance at 1% level, ** at 5% level, * at 10% level. Critical values for ADF test from Hamilton (1994); for Leybourne–McCabe test from Kwiatkowski, et al. (1992). Lag lengths chosen via Schwarz Information Criterion. All tests include a constant term.

**Figure 1: Impulse Responses for Model F, 1959:01 - 1997:06**

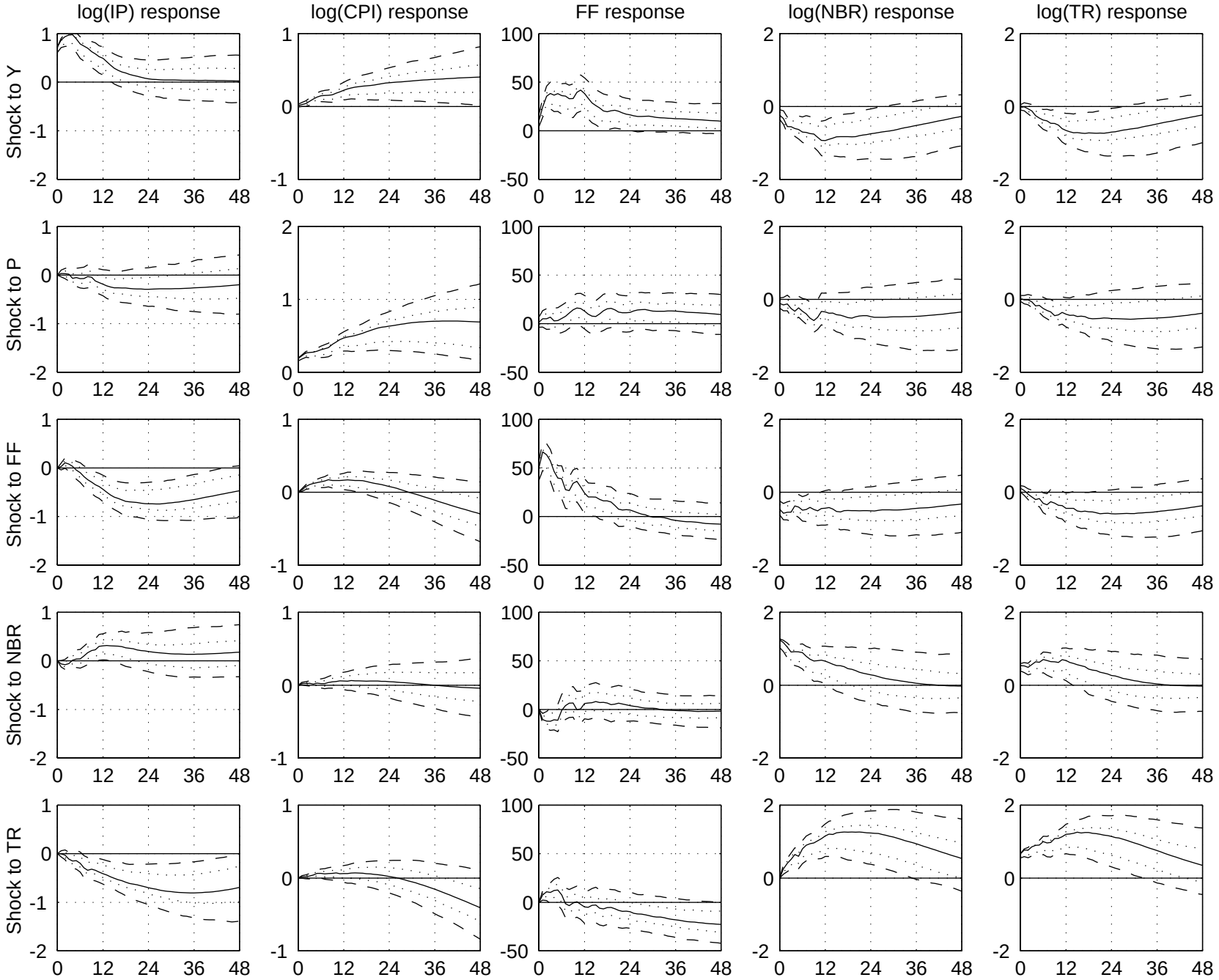

**Figure 2: Impulse Responses for Model S, 1959:01 - 1997:06**

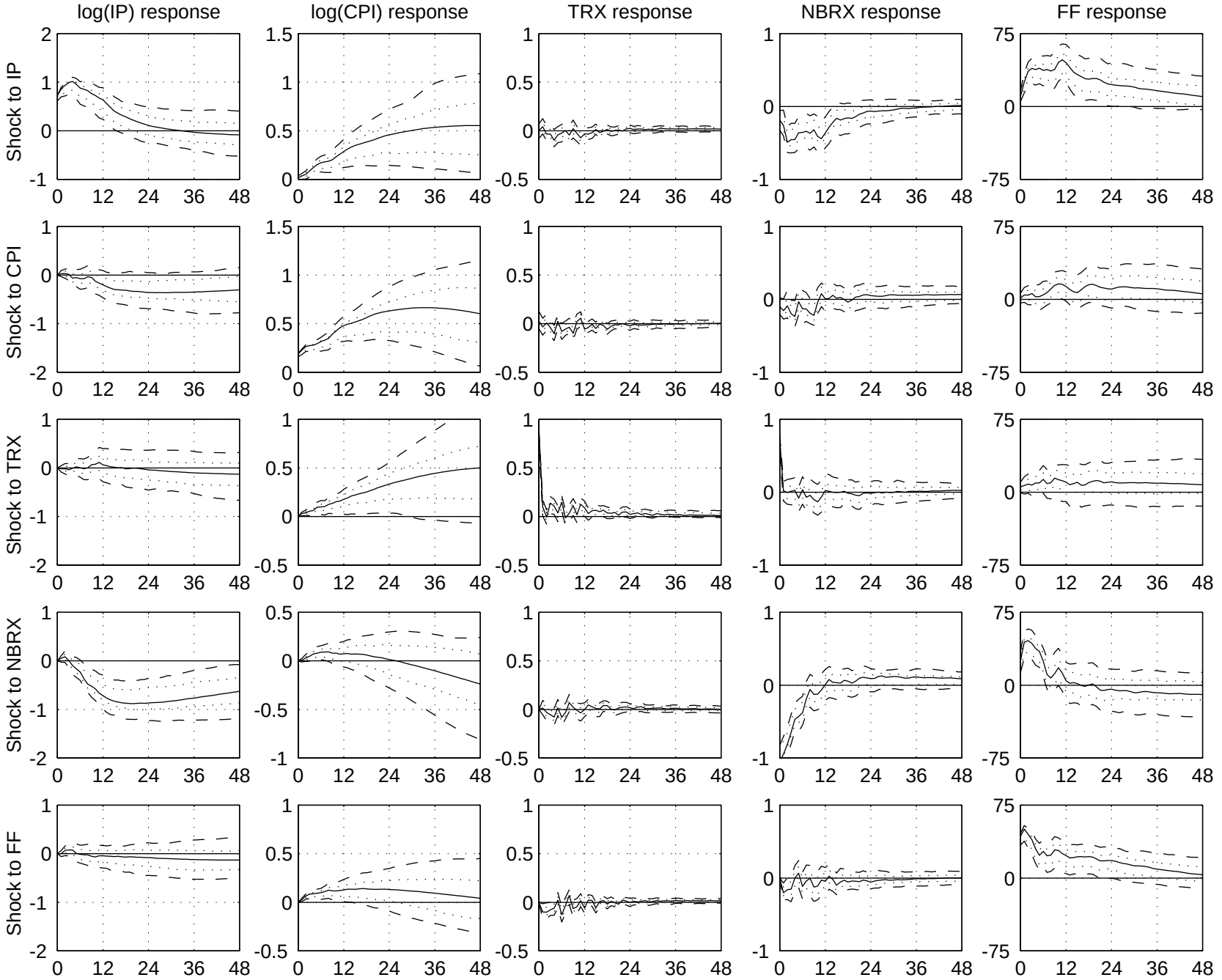

**Figure 3a: Impulse Responses for Model F, 1959:01 - 1997:06**

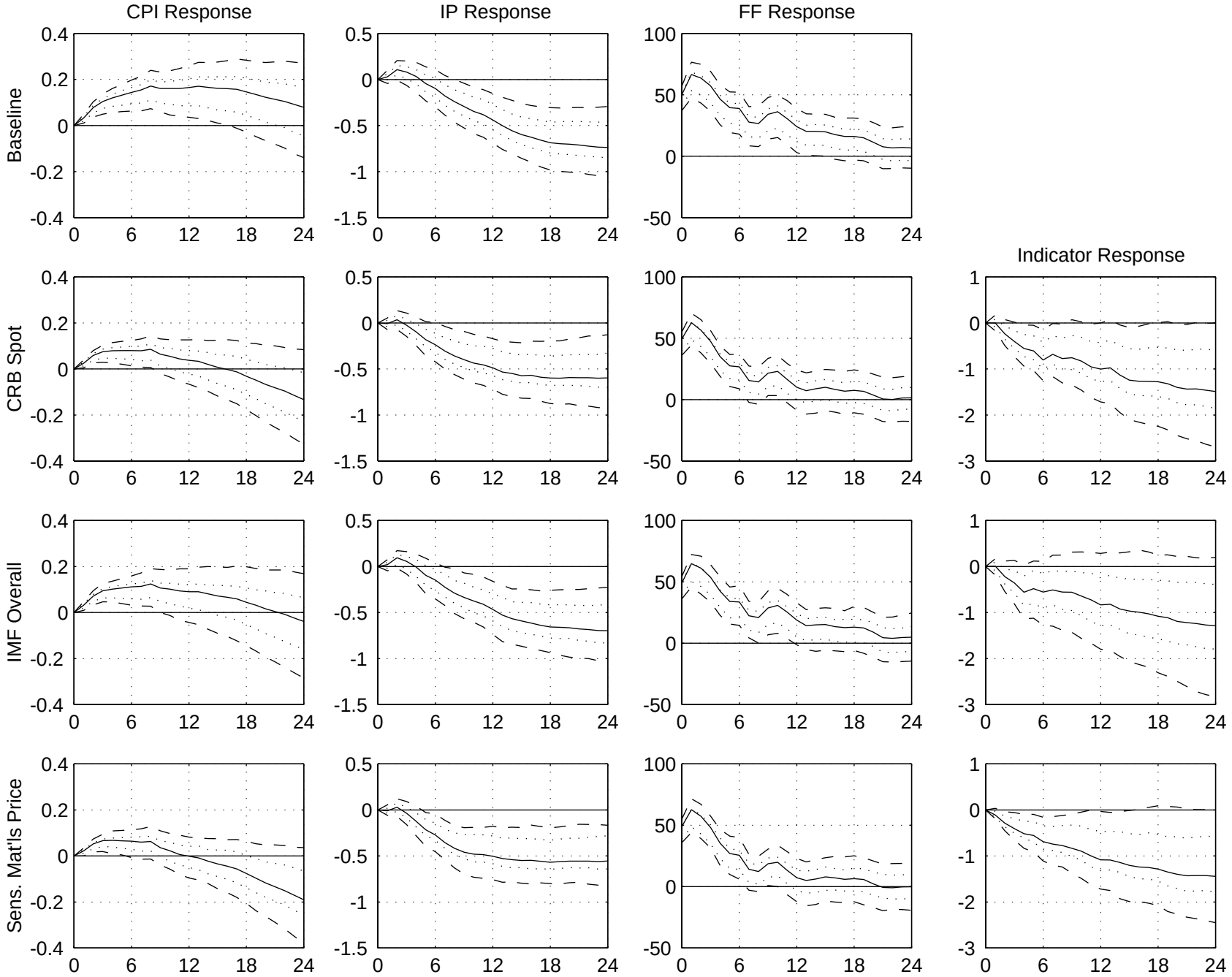

**Figure 3b: Impulse Responses for Model F, 1959:01 - 1997:06**

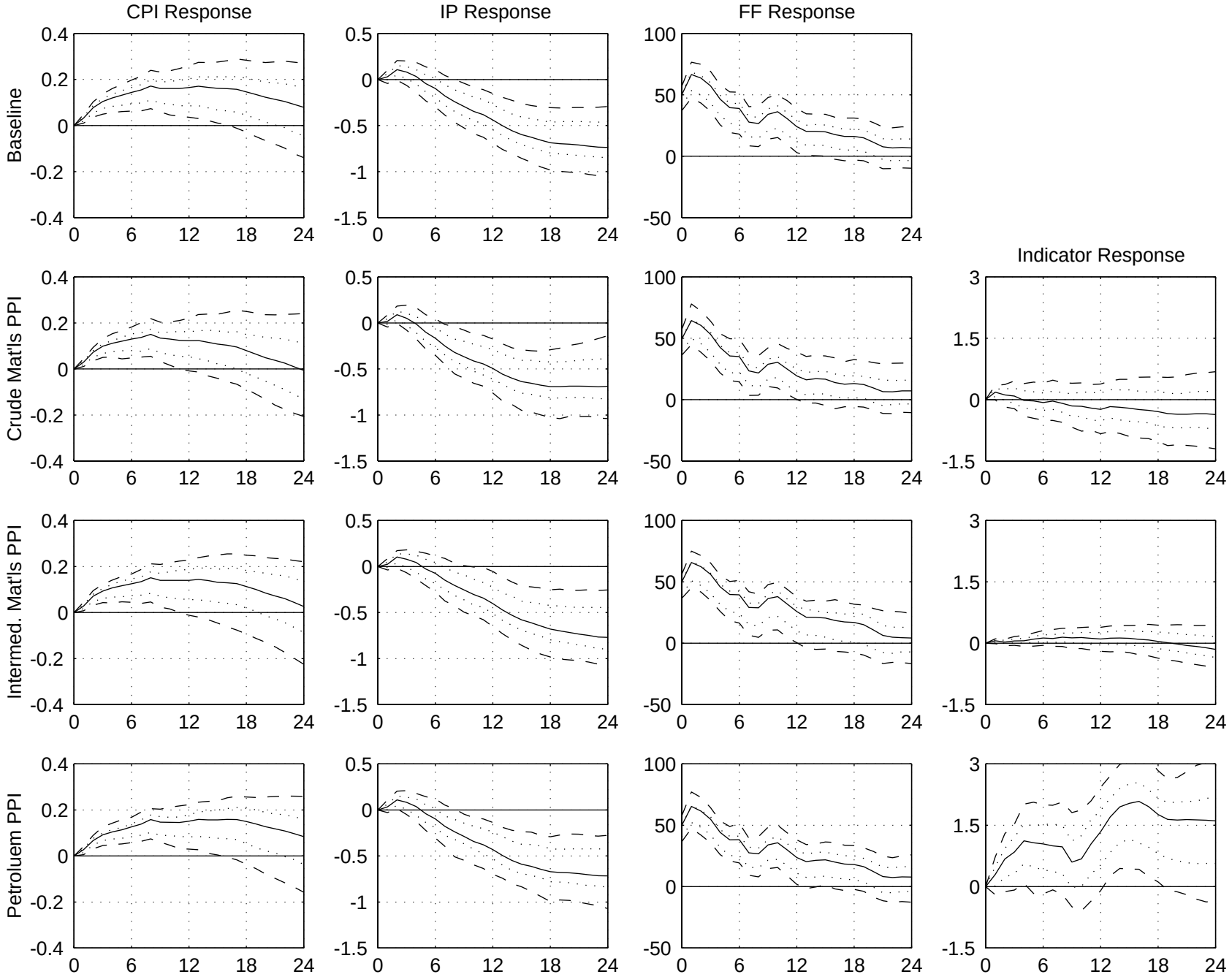

**Figure 3c: Impulse Responses for Model F, 1959:01 - 1997:06**

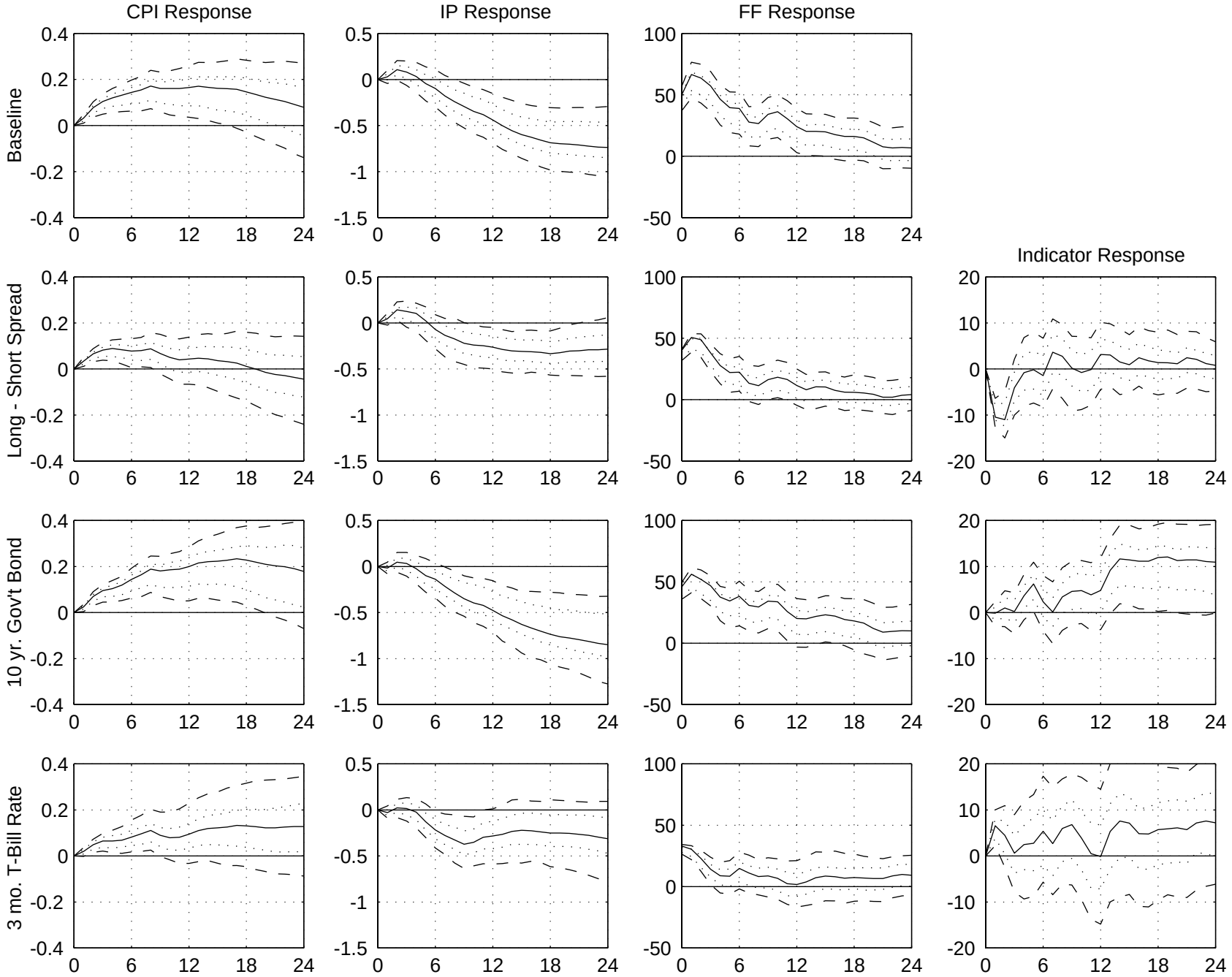

**Figure 3d: Impulse Responses for Model F, 1959:01 - 1997:06**

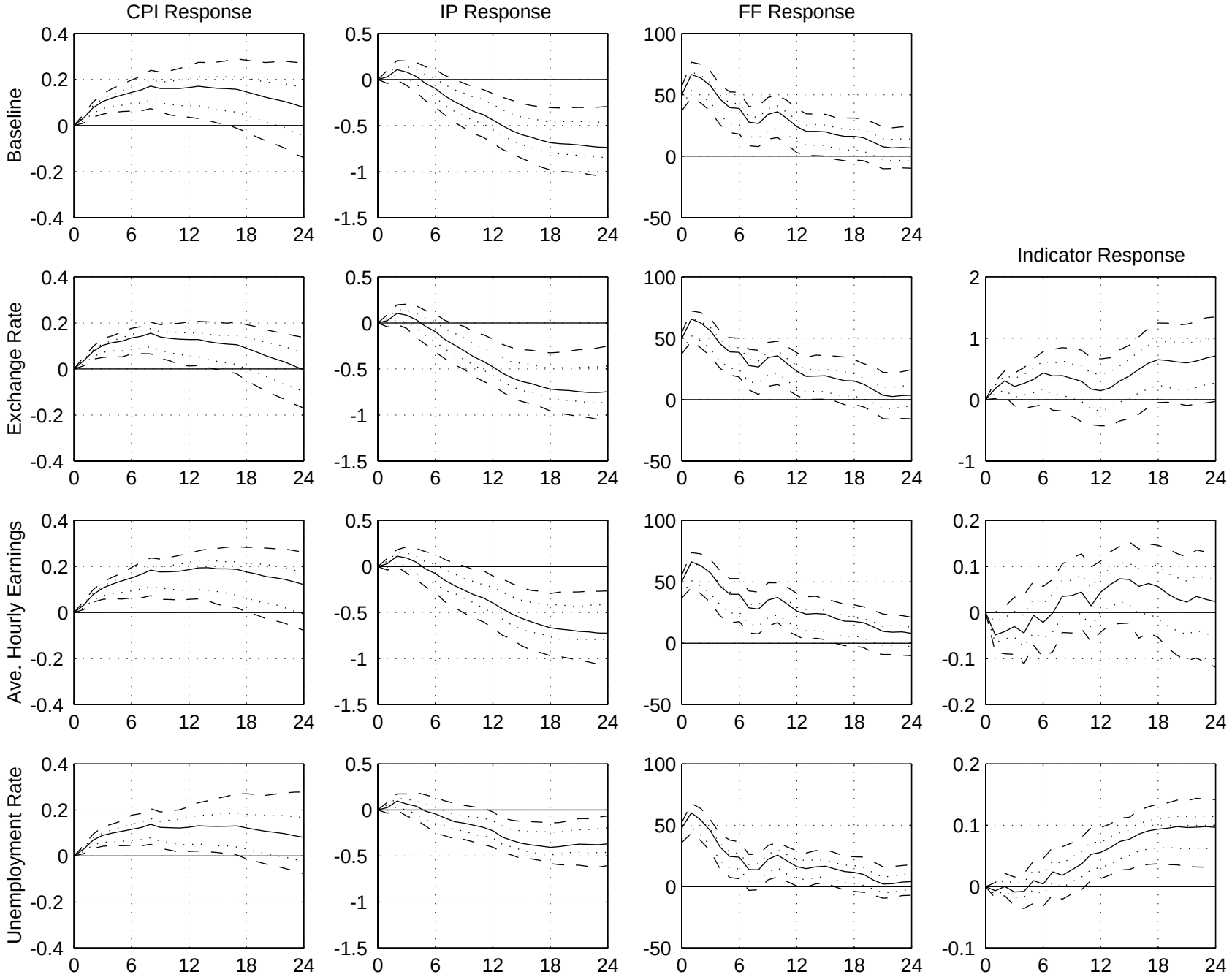

**Figure 4a: Forecasting Power vs. Resolution of Price Puzzle: 6 mo. horizon, 59 - 97**

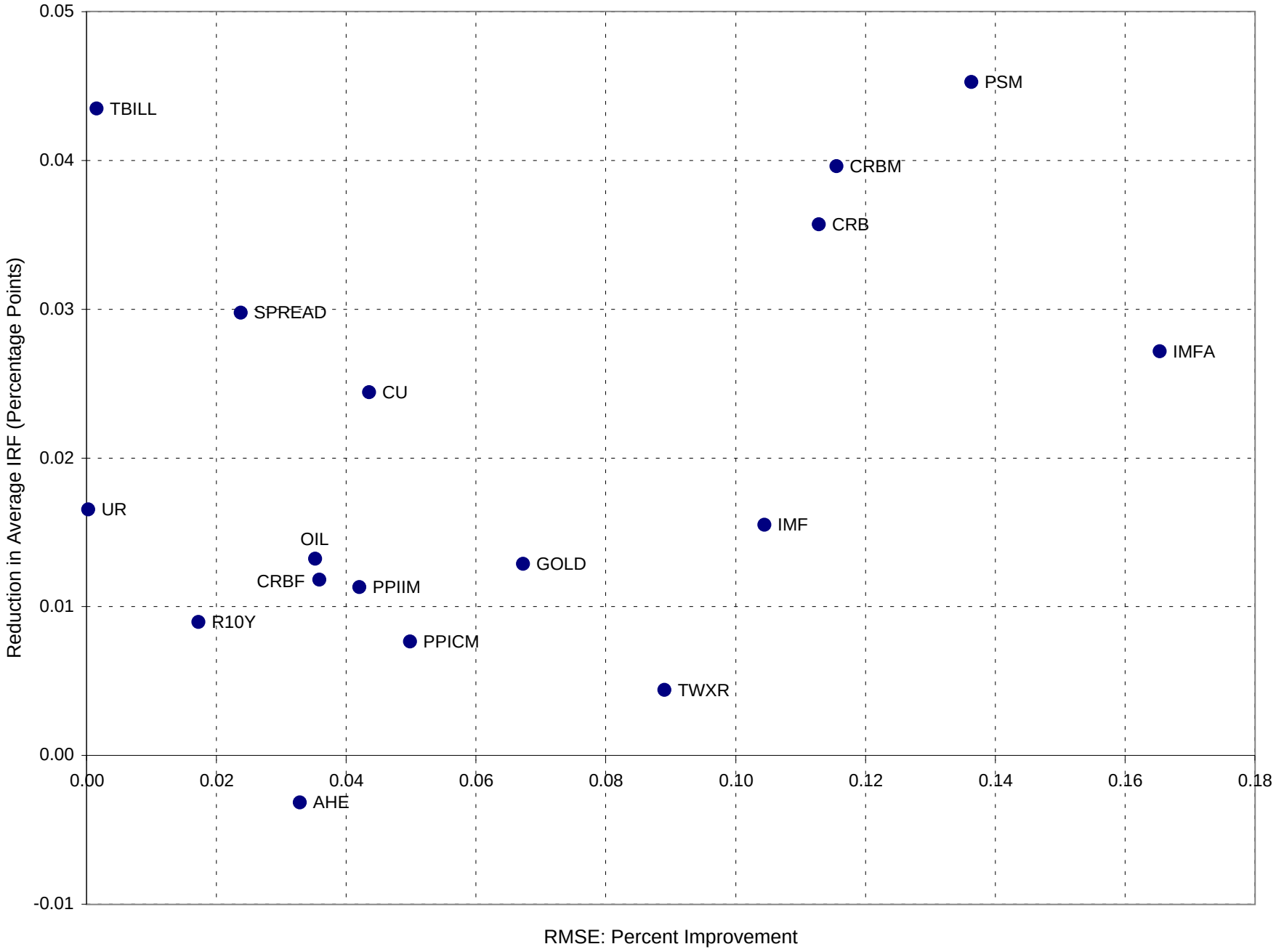

**Figure 4b: Forecasting Power vs. Resolution of Price Puzzle: 1 yr. horizon, 59 - 97**

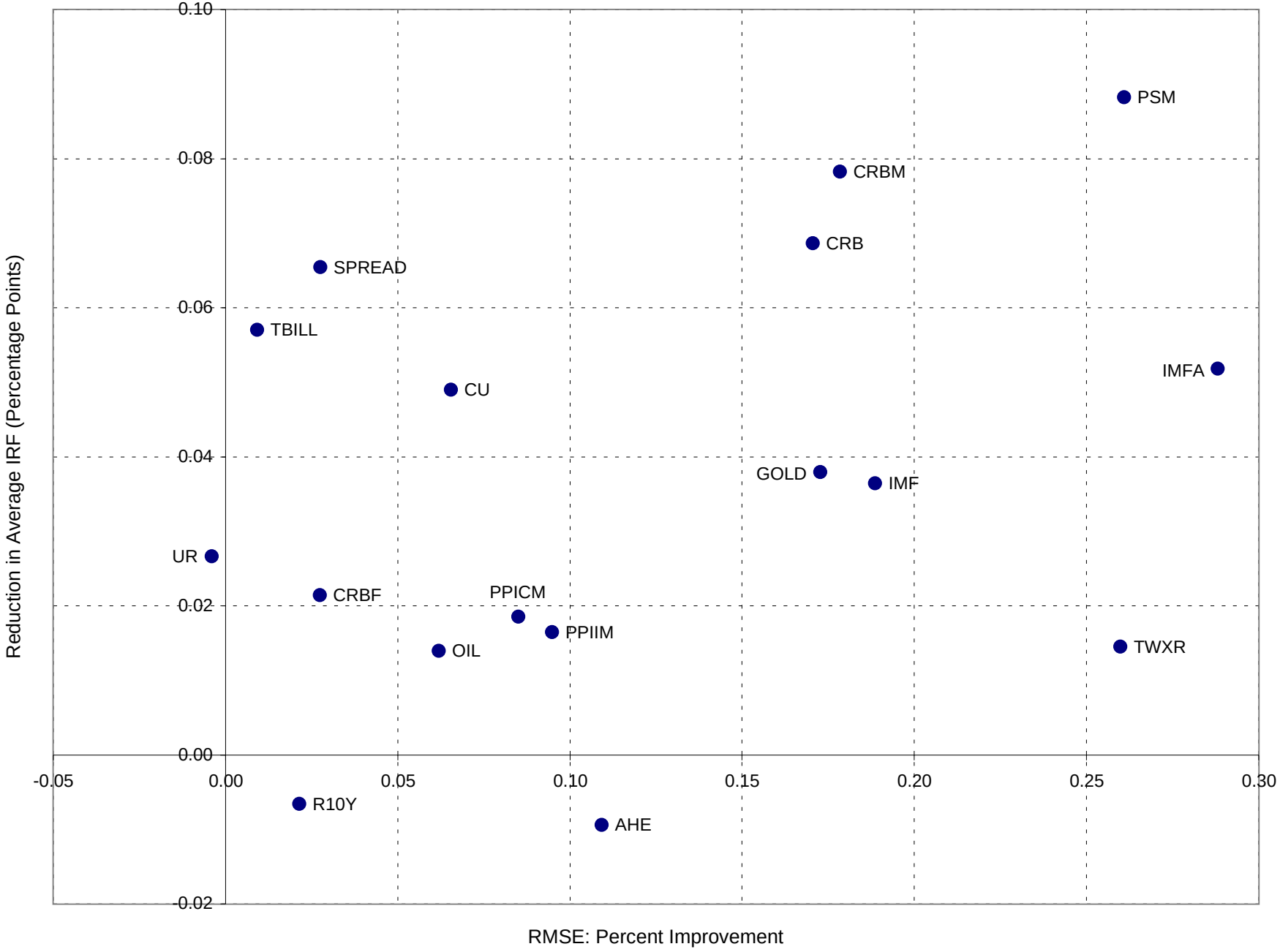

**Figure 4c: Forecasting Power vs. Resolution of Price Puzzle: 2 yr. horizon, 59 - 97**

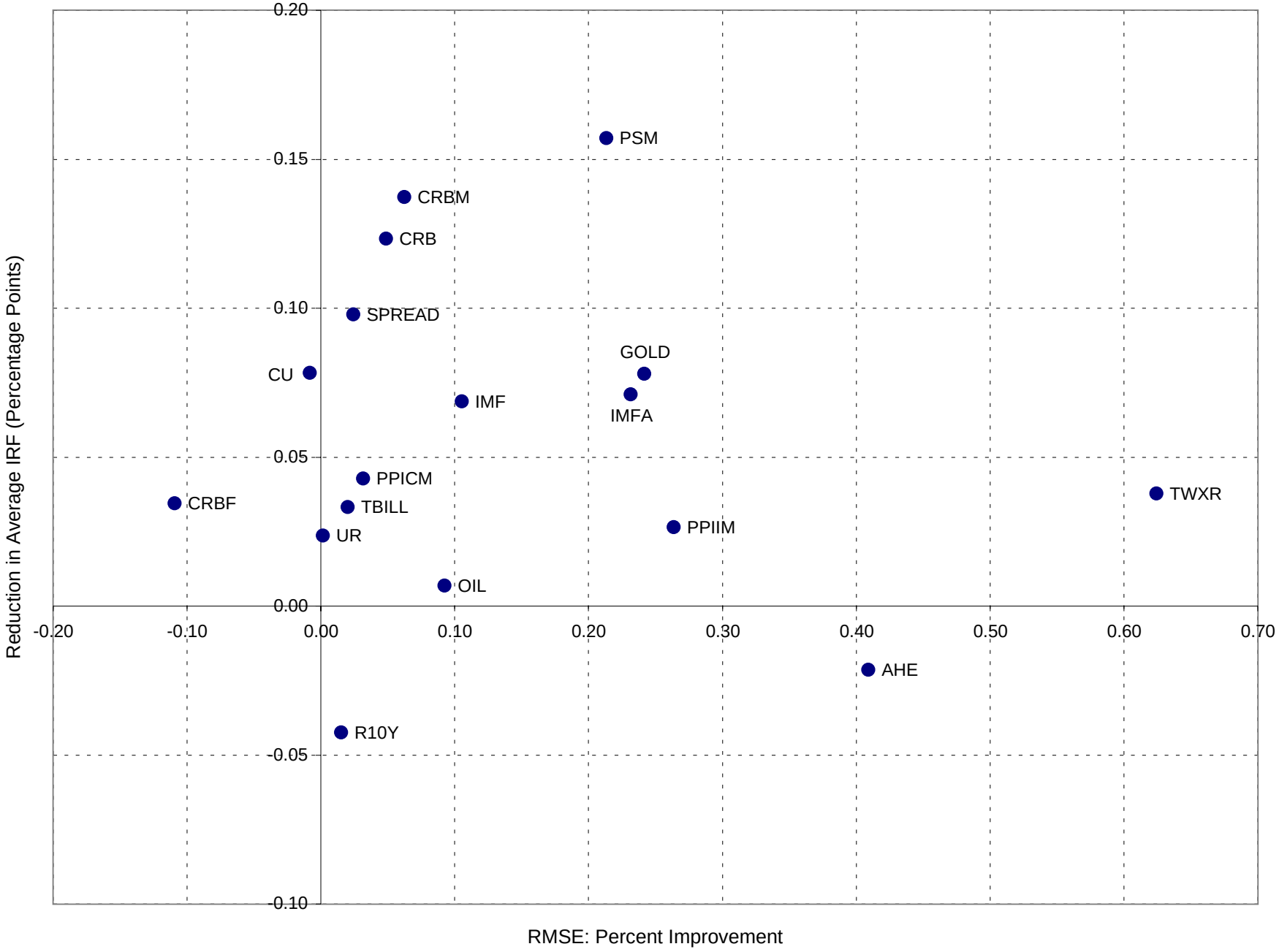

**Figure 5: Impulse Responses to 50 Basis Point Policy Shock (Model F, Baseline)**

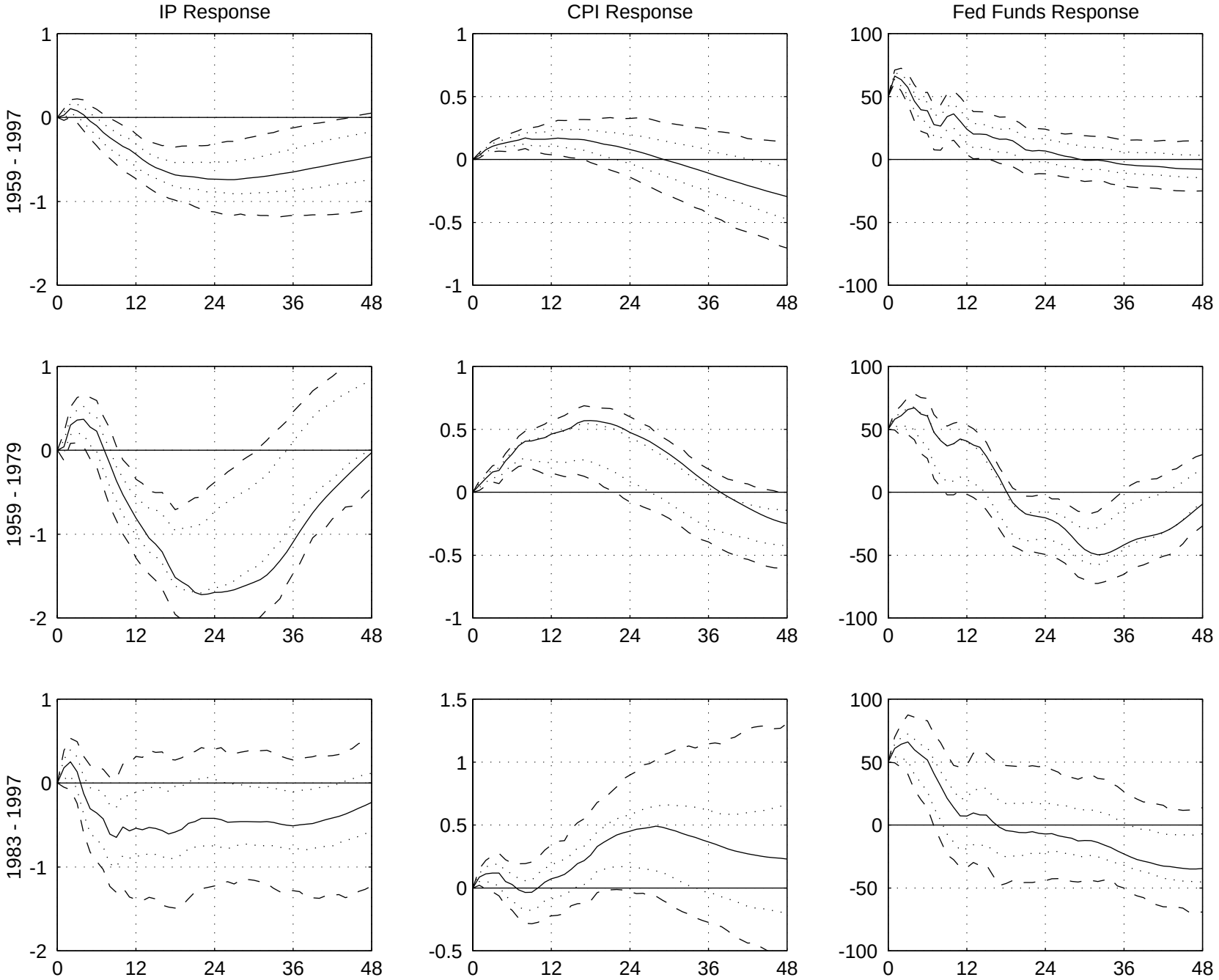

**Figure 6: Forecasting Power vs. Resolution of Price Puzzle: 1 yr. horizon, 59 - 79**

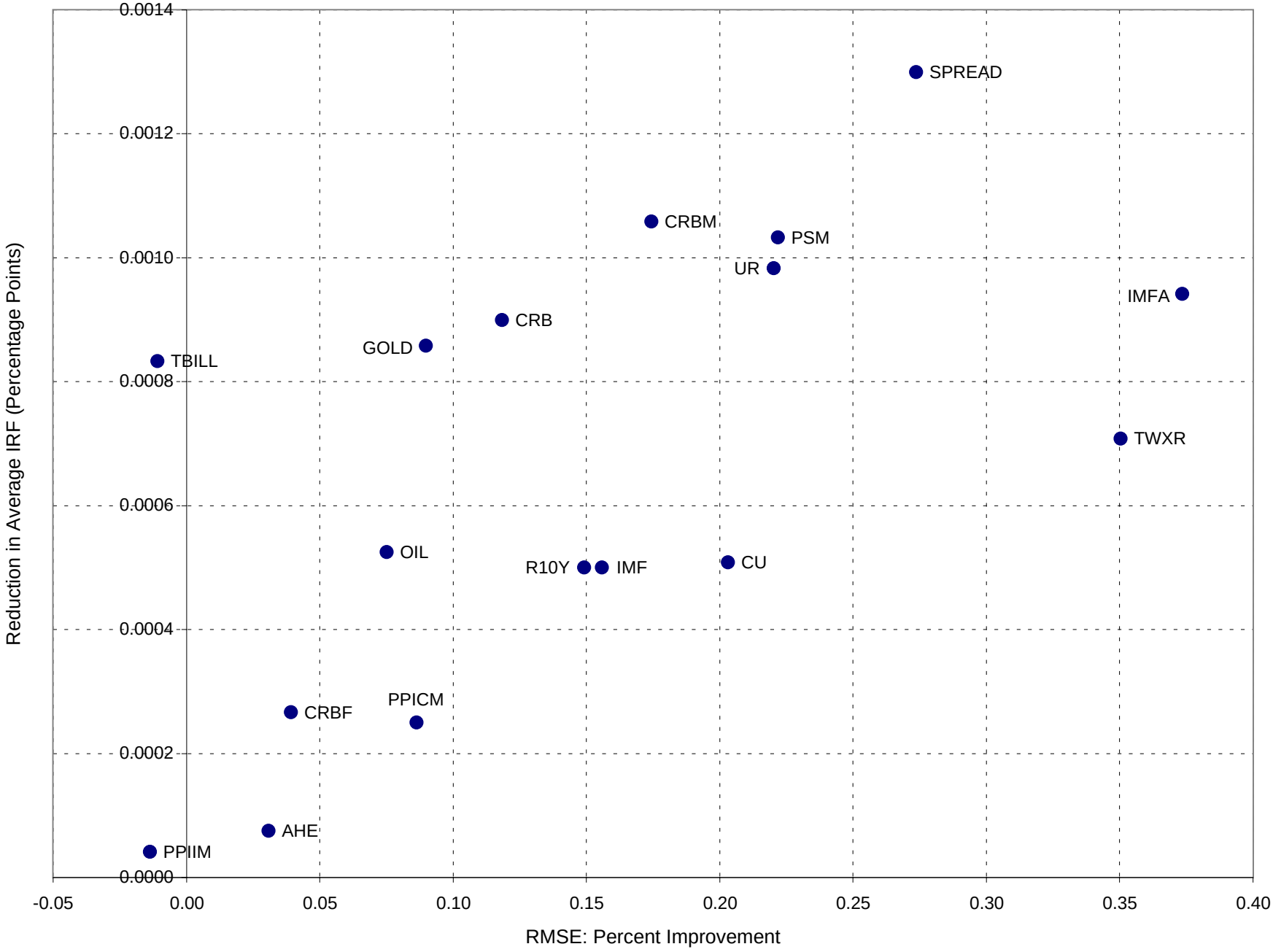

**Figure 7: Impulse Responses to Policy Shock in Model F, 1959:01 - 1979:09**

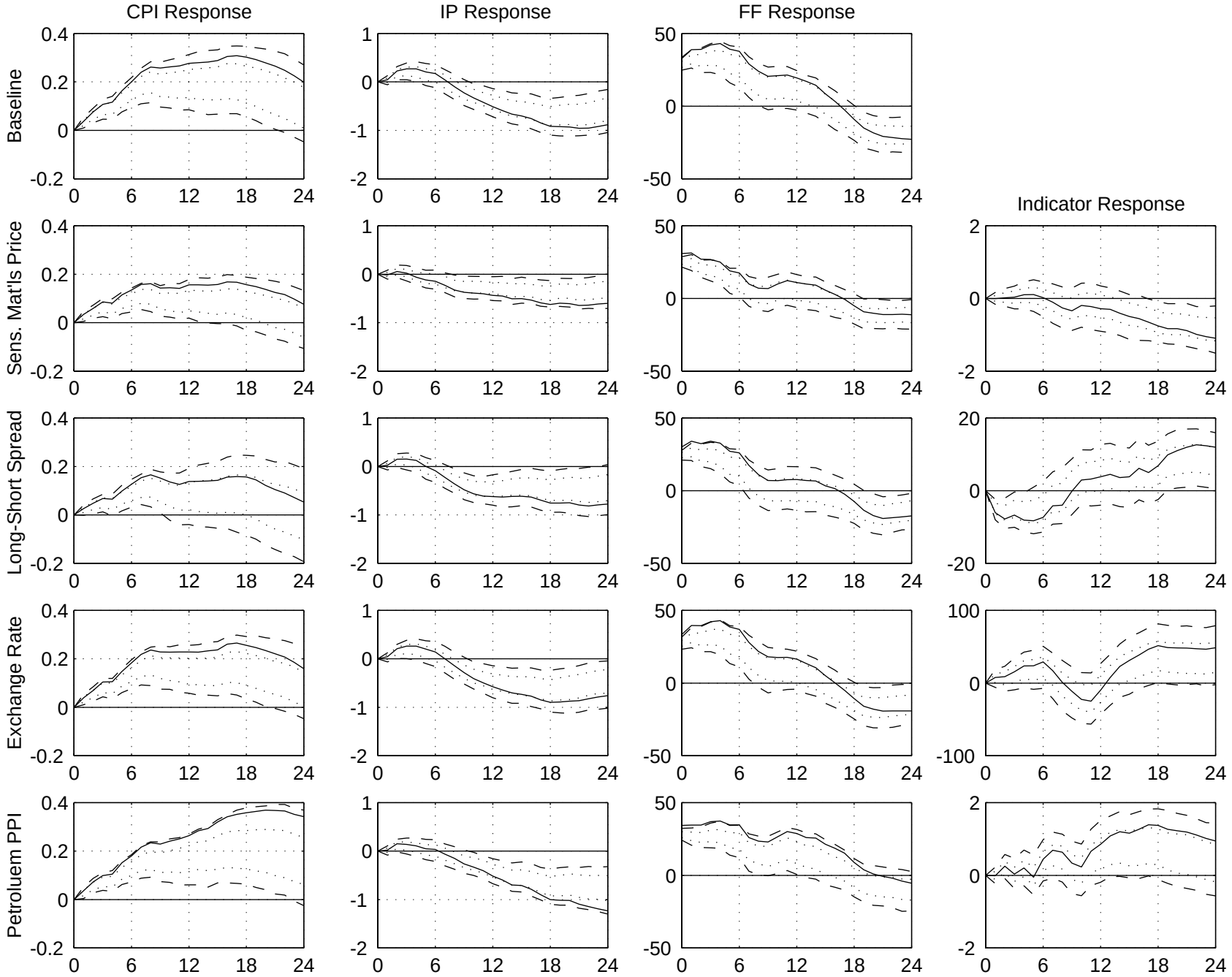
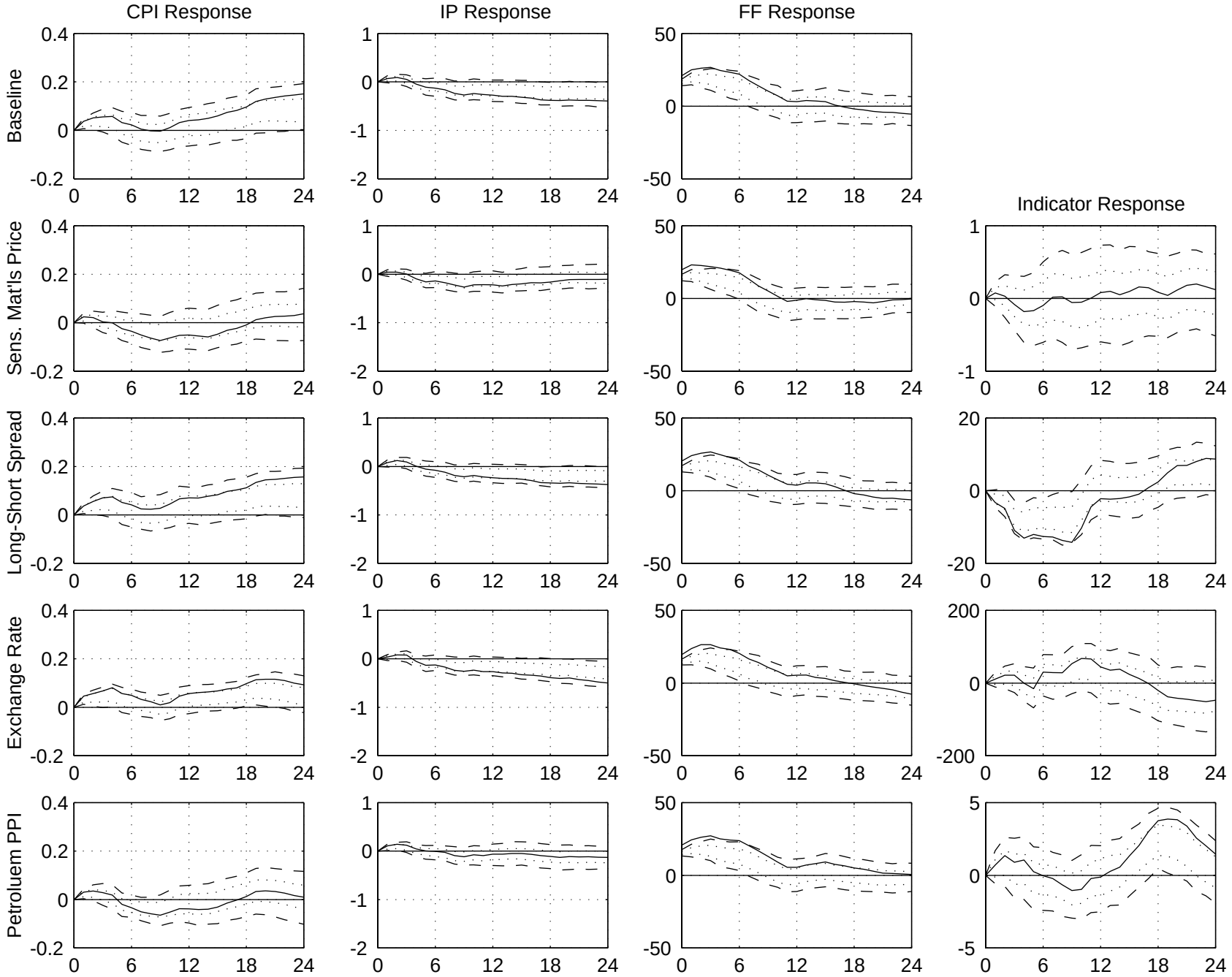
**Figure 8: Impulse Responses to Policy Shock in Model F, 1983:01 - 1997:12**