VECTOR AUTOREGRESSIVE MEASURES OF MONETARY POLICY: ISSUES AND CRITIQUES

by

Michael Steven Hanson

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in The University of Michigan
2000

Doctoral Committee:

Professor Robert B. Barsky, Co-chair
Professor Matthew D. Shapiro, Co-chair
Associate Professor Kathryn M. Dominguez
Professor E. Philip Howrey

To Corina and Alexander

ACKNOWLEDGEMENTS

For a project this many years in gestation, a great many people contributed in one way or another to its final form. Firstly, the members of my committee, each of whom challenged my thinking in one way or another: Bob Barsky, who provided much inspiration for my research agenda; Matt Shapiro, who provided guidance on all matters professional; Phil Howrey, who patiently worked through many econometrics issues with me; and Kathryn Dominguez, who provided invaluable, candid advice.

Susanto Basu also was a main source of inspiration throughout graduate school. I learned agreed deal from him as a student, research assistant, teaching assistant and Ph.D. candidate. His early influence was instrumental in my choice to pursue research in macroeconomics. I am grateful for the many hours he invested in me.

Many friends from graduate school also helped along the way. Two of whom were particularly influential early on were Chris Foote and Jonathan Parker. They each brought a unique intellectual excitement to the study of economics during my early years as a graduate student. On the job market, I benefited from interactions with Erik Hurst, Jon Bakija, Stephen Holland, Peter Boberg and Ufuk Demiroglu.

Special thanks to the International Finance Division of the Federal Reserve Board of Governors for offering a fellowship to continue my research in the summer of 1998. A good number of people there provided comments and insights, including but not limited to Jon Faust, John Rogers, Dale Henderson, Andy Levin, Chris Erceg, John Fernald, Allan Brunner, Charles Fleischman, and especially my fellow interns, Mark

Aguiar and Petra Gegraats. (The views expressed within are solely those of the author, and not necessarily those of the Federal Reserve Board.)

Additional thanks to my colleagues at Wesleyan, who offered much nonjudgmental support (and editorial comments) as I completed this project. They have impressed upon me how this project is a beginning rather than an end.

Finally, I must offer most sincerest gratitude to my wife Corina for her unwavering patience and love. Without her support, this project might not ever have reached fruition. I owe her a debt I may never be able to repay.

And to little Alexander, who in the past nine months has changed my life in ways too wondrous for me to understand, let alone express in words. He simultaneously distracted and motivated me. Seeing his beautiful smile makes any task a little easier.

PREFACE

This dissertation is comprised of three distinct but related essays on the empirical measurement of monetary policy. One common thread through each chapter is the use of a vector autoregressive (VAR) methodology, arguably the predominant paradigm in modern monetary economics. Each essay offers a critique of how this technique is employed by the current academic literature that quantifies the effects of monetary policy for the U.S. economy. The first chapter investigates the role of certain "indicator" variables to address a common empirical puzzle, and finds limited support for the conventional approach to resolving this puzzle. The second examines three concepts that deserve more emphasis in the monetary VAR literature: the greater significance of the endogenous monetary policy reaction function vis-à-vis the exogenous policy shocks, the impact of changes in the form of the policy rule over time, and the contribution of parameter instability residing in the non-policy portion of the model. The final essay exploits intuitive but non-traditional restrictions to identify long-run co-movement between commodity prices and consumer prices, and finds that a sizable proportion of each can be attributed to nominal (i.e. monetary) shocks.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

CHAPTER I

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

1.1 Introduction

Over the past decade, a large number of researchers have estimated structural vector autoregressive (SVAR) models to establish "stylized facts" about U.S. 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 largely disappeared if commodity prices were included in his VAR.[2] He proposed that commodity prices served as an "information variable" in the Federal Reserve's policy reaction function, i.e. as an indicator of nascent inflation.

[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 the estimated disturbances (or some subset thereof) with a particular economic interpretation.

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

Despite subsequent advances in SVAR modeling, 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.[3] As this tactic has since evolved into a "conventional wisdom," many authors now make only a passing reference to the problem commodity prices are intended to resolve. Nor is any *a priori* rationale given for including commodity prices in an otherwise parsimonious VAR model. And while a VAR system often is meant to correspond with a theoretical business cycle model, 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. Proponents of this view include Zha (1997), Sims (1998a), and Christiano et al. (1998).[4] Viewed this way, the price puzzle goes to the very heart of empirical measurement of the effects of monetary policy. The exclusion of a measure of future inflation is the particular misspecification cited.

This essay 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."[5] My results cast doubt on both of these proposed ex-

[3] Recent empirical work on the form of the Federal Reserve's reaction function, outside of the VAR literature, have included commodity prices for reportedly similar reasons. (See, e.g., Clarida et al., 1998; Sims, 1999.)

[4] 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).

[5] Christiano et al. (1996b) also propose that commodity prices may belong in the VARs to account

planations. The results of this essay 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. 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 significantly 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 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

for oil price "supply shocks."

mounting evidence that the reaction function before 1979 differs from that since.[6]

Whether or not a change in policy is ultimately responsible for these findings, my subsample 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 essay discusses the monetary policy reaction function as an organizing framework. The recent literature is also reviewed. Identification is taken up in section 1.3, where estimated results from common models in the literature are presented. In section 1.4 I then investigate the inclusion of various candidate indicators of expected inflation in an SVAR model. Section 1.5 presents evidence that the price puzzle is a sample-specific phenomenon. Section 1.6 concludes.

1.2 The Identification of Monetary Policy

While I focus upon a particular empirical puzzle for much of this essay, the central issue is the 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") that summarizes how the monetary authorities respond to the state of the economy (past, present or expected future), and an exogenous part that captures shifts in policy objectives (or other features of policy

[6] In Hanson (2000) 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 et al. (1998). De Long (1997) and Sargent (1998) investigate the pre-1979 policy situation in more depth.

that cannot be captured by the estimated rule).[7]

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 essay I wish both to evaluate these "stylized facts" and to infer something about the role for endogenous policy, in light of sub-sample estimation results.

1.2.1 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 posits one equation that can be interpreted as the monetary policy reaction function. In general terms, this reaction function can be written as:

$$m_t = f(\mathcal{I}_t) + \mu_t \, , \tag{1.1}$$

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, μ_t, represents the component of policy that is not a function of any current or past information contained in $\mathcal{I}_t$, and is in that sense

[7] 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 et al. (1998).

"exogenous."[8] 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 have been 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 (1.1) typically only observes some subset of the information utilized by the monetary authority: $\mathcal{H}_t \subset \mathcal{I}_t$.[9] Thus, the econometrician estimates

$$m_t = g(\mathcal{H}_t) + \eta_t\,, \tag{1.2}$$

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 that should contain information about future inflation — could result in a poor estimate of $g(\cdot)$. Suppose the excluded inflationary indicator variable signaled a rise in future inflation, to which the Fed endogenously reacted

[8] Hall (1996) refers to this component as the "spontaneous random element" of policy. Unfortunately such a characterization has lead to some confusion regarding SVAR analysis: that it implies truly "random" — and by implication, irrational or sub-optimal — behavior by the monetary policy makers. Rather, the μ_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.

[9] Parsimonious estimation generally will require a much smaller information set than $\mathcal{I}_t$. An example of data that 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 that preclude $\mathcal{H}_t$ from being a subset of $\mathcal{I}_t$.

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 prices 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 μ_t shock — should not lead to an increase in the price level.[10]

1.2.2 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 that measured exogenous changes in money supply, and therefore monetary policy. Numerous variations on broad money aggregates were commonly proposed. Bernanke and Blinder (1992) countered by presenting evidence in favor of the federal funds rate 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*

[10]The source of misspecification given by this proposed explanation 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 that are otherwise absent from the VAR. Consideration of such a hypothesis is beyond the scope of this paper. See Barth and Ramey (1999) for an investigation of sectoral, rather than aggregate, price indexes.

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 monetary policy from other shocks. Models that evolved from that investigation include those developed by Strongin (1995) and Christiano et al. (1996a,b). Bernanke and Mihov (1998) compare several models in this class within an encompassing framework. Further investigation has been undertaken by Leeper et al. (1996), Christiano et al. (1998), and Bagliano and Favero (1998), to name but a few.[11] With the exception of Strongin (1995), all of these papers 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 portion of the model unrestricted. Christiano et al. (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 report that the inclusion of commodity prices tends to eliminate the puzzle.[13]

[11] Similar models have been estimated by Cochrane (1994), Kim (1996) and Bernanke et al. (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 (as 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).

Non-recursive semi-structural models also have been estimated in the literature. Recent examples include Gordon and Leeper (1994), Leeper (1995), Leeper et al. (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 that 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.

In this essay I focus on models whose supply and demand shocks originate in the reserves market. These models of the economy have several features that 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 et al. (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]

[14] One reason for the popularity of these models undoubtedly is the ease of estimating recursive models in many statistical packages. However, Christiano et al. (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." 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*. In this sense 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) critique this class of models.

1.3 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 essay. This section concludes with the estimated impulse response functions for several "baseline" models: models that do not include any of the candidate inflation indicators.

1.3.1 Overview of SVAR Estimation

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

$$X_t = \Theta(L)\,\varepsilon_t\,. \tag{1.3}$$

X_t is an n-dimensional vector of observed endogenous variables, and ε_t a vector of unobserved structural disturbances. These are white noise primitives whose history accounts for the fluctuations of the endogenous variables. The monetary policy shock, μ_t, is an element of ε_t. $\Theta(L)$ is a one-sided (infinite-order) matrix polynomial in the lag operator L.[16] Thus equation (1.3) is the vector moving average (VMA) representation of X_t in the structural disturbances ε_t. The estimated $\Theta(L)$ will give the impulse response functions of interest. The covariance matrix for the structural disturbances is given by

$$E\big[\varepsilon_t\varepsilon_t'\big] = \Omega\,. \tag{1.4}$$

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

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

Under standard regularity conditions, the model in equation (1.3) also has a structural vector autoregressive (SVAR) form:

$$\Phi(L)\, X_t = \varepsilon_t \,, \tag{1.5}$$

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 (1.5) 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:

$$A(L)\, X_t = v_t \,, \tag{1.6}$$

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

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

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,

$$E\big[v_t v_t'\big] = \Sigma \,, \tag{1.7}$$

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

Given the definitions and the normalization discussed above, the reduced-form VAR system (1.6) 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 ε_t span the same linear space, identification consists of finding the $n \times n$ matrix S such that

$$v_t = S\,\varepsilon_t\,. \tag{1.8}$$

Recall that the elements of ε_t are orthogonal by construction.

Equation (1.8) implies the following relationship between the structural and reduced-form covariance matrices:

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

This follows from equations (1.4) and (1.7), and the previous normalization of $\Omega = I$. Recall that Σ 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 (1.8) 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.

1.3.2 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 into 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 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 (1.5) can now be divided into policy and non-policy blocks and rewritten as

$$\begin{bmatrix} \phi_{YY0} & \phi_{YM0} \\ \phi_{MY0} & \phi_{MM0} \end{bmatrix} \begin{bmatrix} Y_t \\ M_t \end{bmatrix} = \begin{bmatrix} \Phi_{YY}(L) & \Phi_{YM}(L) \\ \Phi_{MY}(L) & \Phi_{MM}(L) \end{bmatrix} \begin{bmatrix} Y_t \\ M_t \end{bmatrix} + \begin{bmatrix} \varepsilon_t^Y \\ \varepsilon_t^M \end{bmatrix}. \tag{1.9}$$

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 (1.6), then can be expressed as

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

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

An essential identifying assumption is the predetermination of the non-policy block within a period. That is, policy can 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 μ_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 instrument model (m_t = FF), and a nonborrowed reserves instrument model (m_t = NBR). The blocks that 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 instrument model, as in Bernanke and Blinder (1992) and Bernanke and Mihov (1998),[20] the Federal Reserve (perfectly) targets the funds rate, making the effective supply of reserves (perfectly) elastic: any change in demand for reserves would be accommodated by the Fed to return the funds rate to its targeted level. Changes in the funds rate then would be identified as policy shocks, whereas changes in the quantity of reserves would reflect shocks to both supply and

[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 interpretation of 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.

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 I examine.

To make this model operational, the monetary policy shock is identified by placing the funds rate first in the policy block 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.[21]

Alternatively, the Fed could target the quantity of reserves in the market by supplying reserves (perfectly) inelasticly; 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 et al. (1996a) argue in favor of nonborrowed reserves as being under tighter control by the Fed than any of the broader measures of money.[22] One might expect this specification 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 variant of this model has been proposed by Strongin (1995), who identifies the policy instrument as that portion of nonborrowed reserves that 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.[23] He views

[21] A variant of this model has also been estimated by Christiano et al. (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] This is also the "NBR" model of Christiano et al. (1996b).

[23] 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.

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 is less appropriate 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 the above nonborrowed reserves model is a special case of the shock from Strongin's model.[24]

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 a fed funds instrument model 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. Strongin's model is predicated upon the idea that orthogonal shocks to nonborrowed reserves fully and uniquely measure the monetary policy shocks.

1.3.3 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 1.1 presents the estimated impulse response functions for the funds rate model. 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

[24]Technically, Strongin's model does not nest the nonborrowed reserves model as both are just identified.

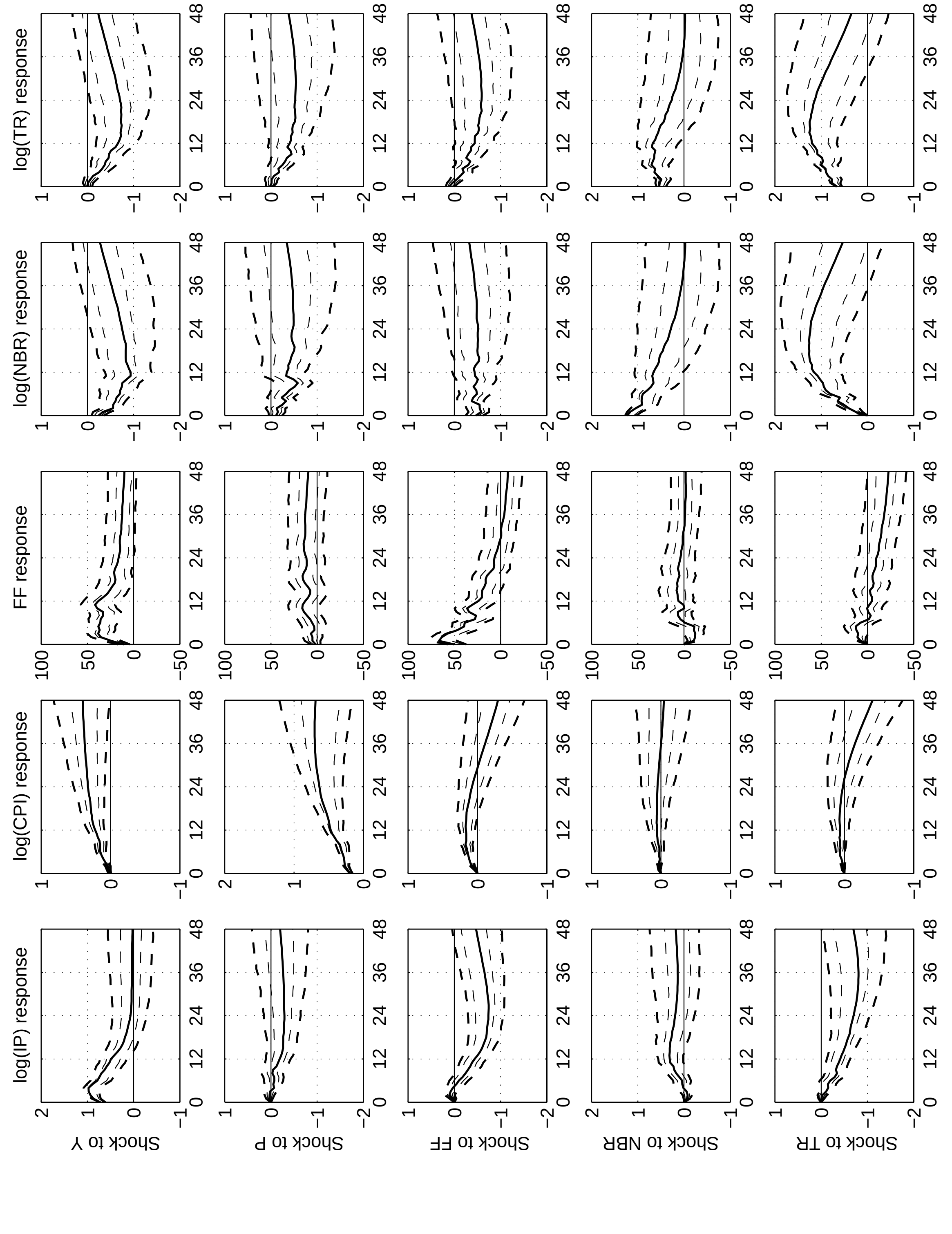


Figure 1.1: Impulse Responses for Baseline Model with Fed Funds Instrument, 1959:01 - 1997:06

index.[25]

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 at 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 thin dashed lines report 68% bootstrapped confidence bands, while the thick dashed lines represent 95% bootstrapped confidence bands.[26] While 95% error bands are becoming more common in the literature, I also report the 68% bands for comparison with earlier papers.[27]

The third row of figure 1.1 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 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 that 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.

[25]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. According to Sims et al. (1990), estimation in (log) levels will be consistent (but not necessarily efficient).

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

[27]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.

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 et al., 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. Few economists 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.[28]

An alternative assumption to targeting the federal funds rate is targeting nonborrowed reserves. Figure 1.2 reports the impulse responses for Strongin's model (described above).[29] 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. A contractionary

[28]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 1.1.

[29]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.

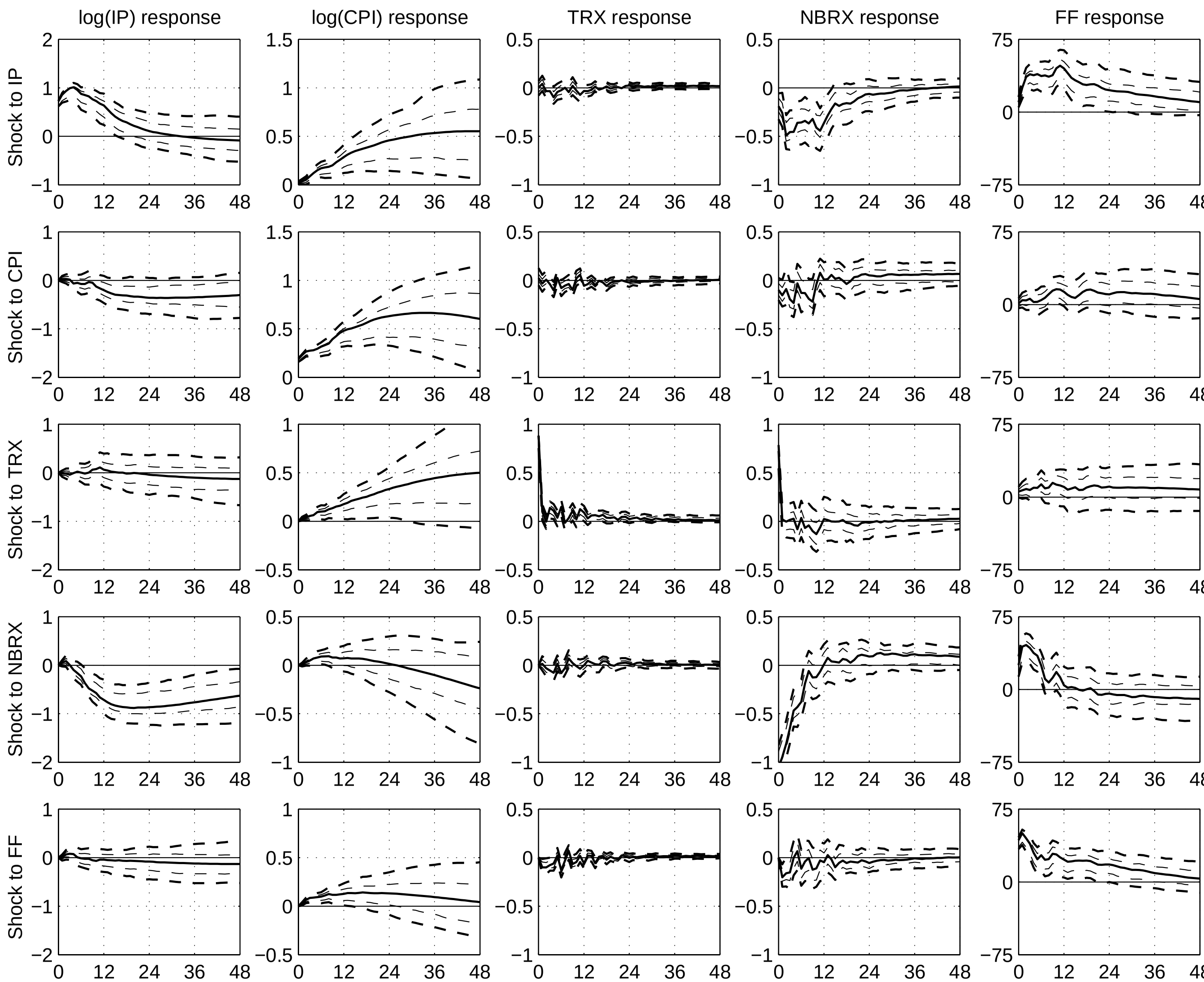


Figure 1.2: Impulse Responses for Baseline Model with Nonborrowed Reserves Instrument, 1959:01 - 1997:06

policy shock is a one-time *decrease* in nonborrowed reserves.[30]

Similar to the effects in the funds rate model, a contractionary policy shock in Strongin's model 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 the funds rate model, 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 Strongin's model over a funds rate model. Indeed, one possible interpretation of these results is that the true source of misspecification is not excluding certain information variables but choosing the wrong variable as the policy instrument.[31]

Before accepting that conclusion it is worthwhile to examine the role of innovations to the federal funds rate in Strongin's model. The fifth row of figure 1.2 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

[30] By contrast, the fourth row of figure 1.1 records the effect of a positive shock to nonborrowed reserves.

[31] 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 et al. (1996b, 1998). Strongin reports that his model does not exhibit a price puzzle, consistent with our findings. On the other hand, Leeper et al. (1996) conclude that Strongin's model generates a significant price puzzle. Their version of Strongin's model differs from my estimates in several respects, most importantly in 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 estimate over a slightly different sample: 1960:01 – 1996:03.

shock. However, the impulse responses to a FF shock in his model mirror those of the more common funds rate instrument model. In particular, one observes a price puzzle of 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 the orthogonalization implied by either model generates similar "structural" shocks.

Not only does the orthogonalized funds rate shock in Strongin's model 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 Strongin's model is in fact attributable to monetary policy. Moreover, as emphasized by Leeper et al. (1996), models that target a reserve variable imply potentially dramatic swings in the federal funds rate in response to unexpected changes in discount window borrowing, which is strongly at odds with actual practice by the monetary authorities.[32]

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. 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 empirical monetary literature — vector autoregressive or otherwise — 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

[32]Note that these criticisms apply with equal force to the basic nonborrowed reserves instrument model. Leeper et al. (1996) offer additional criticism of this approach.

given to interest rate rules in the theoretical literature. For these reasons, plus the fact that the phenomenon I wish to examine is more pronounced in the fed funds instrument model, I focus on that specification in the following section.

1.4 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 et al. (1996b), the information variable most commonly added in the literature is commodity prices. Yet other variables might plausibly contain information about future inflation.[33] Consideration of the channels through which a particular variable might be expected to help forecast inflation suggests a number of candidate indicators. 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.

[33]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.

One prominent channel is a "chain of production" or "pass-though" one: an increase in the costs of intermediate inputs may lead to an increase in the prices of final goods. In this case, indicators which directly measure such costs play a causal role for the aggregate price level. Relatedly, such variables could measure "supply shocks." Examples include producer price indices, the price of oil or (arguably) general commodity price indices, measures of labor costs and capacity measures. Import prices or exchange rates may also fit this story.[34]

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 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 1.1 lists the particular series I examine. This list of potential indicators is meant only to be representative, not exhaustive.[35]

[34]Or the shocks in question could be purely monetary, in which case a monetary aggregate might be an appropriate indicator.

[35]Cecchetti (1995) provides a complementary investigation of indicators of inflation. Christiano et al. (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. I analyze each of these indicators in greater detail below.

1.4.1 Forecasting Power of Candidate Indicators

Table 1.1 reports the root mean squared error (RMSE) for a VAR-based forecast of the price level at various horizons. From equation (1.3), the forecast error for the price level h-periods ahead, conditional on information available at time t, can be written as:

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

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 Θ_k. (The price level always is 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

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

since $E[\varepsilon_t \varepsilon_t'] = I$. The RMSE is simply the (positive) square root of $\mathrm{MSE}(\hat{p}_{t+h|t})$. 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 1.1 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). Subsequent rows display the percentage reduction in the RMSE of the price level forecast from including the indicator variable listed at left, relative to the baseline model (i.e. the baseline model RMSE less the augmented model RMSE). Thus a bigger gap implies more incremental forecasting power for the particular indicator at the specified forecast horizon.

Table 1.1: Root Mean Squared Error 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

Notes: Table reports difference between RMSE of the baseline model (no indicator) and RMSE of augmented model (model includes indicator listed the lefthand column). Negative numbers indicate those indicators that worsen the forecast 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.[36]

Several interesting results appear in table 1.1. 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.[37] Finally, no single indicator (or class of indicators) produces the lowest RMSE at all horizons between three months and three years.

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.[38]

[36] 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 1.1.

[37] Computational considerations precluded testing multiple indicators simultaneously.

[38] 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.

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 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 that do not impact the price level. There is limited 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 appreciably 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 1.1. As measured by the RMSE, the crude materials PPI resembles the broader commodity indexes — better at shorter horizons — although it generally underperforms them.[39] The intermediate goods PPI, on the other hand, seems to shine at the 2- and 3-year horizons. This index may be a better long-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,

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

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 1.1, 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.[40]

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 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 could be consistent with either the pass-through or informational channels outlined above.

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

1.4.2 Alternative Indicators and the Price Puzzle

I turn next to estimating "augmented" SVAR models that include one indicator from a subset of those listed in table 1.1. To match the approach most common in the literature, the indicator variable is included in the non-policy block, Y_t, after output and the aggregate price level. However 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 1.3 through 1.6 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 each of these figures reproduces the 5-variable baseline case as in figure 1.1. The remaining rows plot the responses to the policy shock when the estimation is augmented with the indicator variable listed at the left of the row. The final column gives the response of that indicator variable to the policy shock.

The baseline case was presented 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 price response but does not completely eliminate it. This finding of a residual price puzzle is replicated in several recent monthly studies, including Leeper et al. (1996), and Christiano et al. (1998), and contrasts with previous quarterly results (see Christiano et al., 1996b). Commodity prices

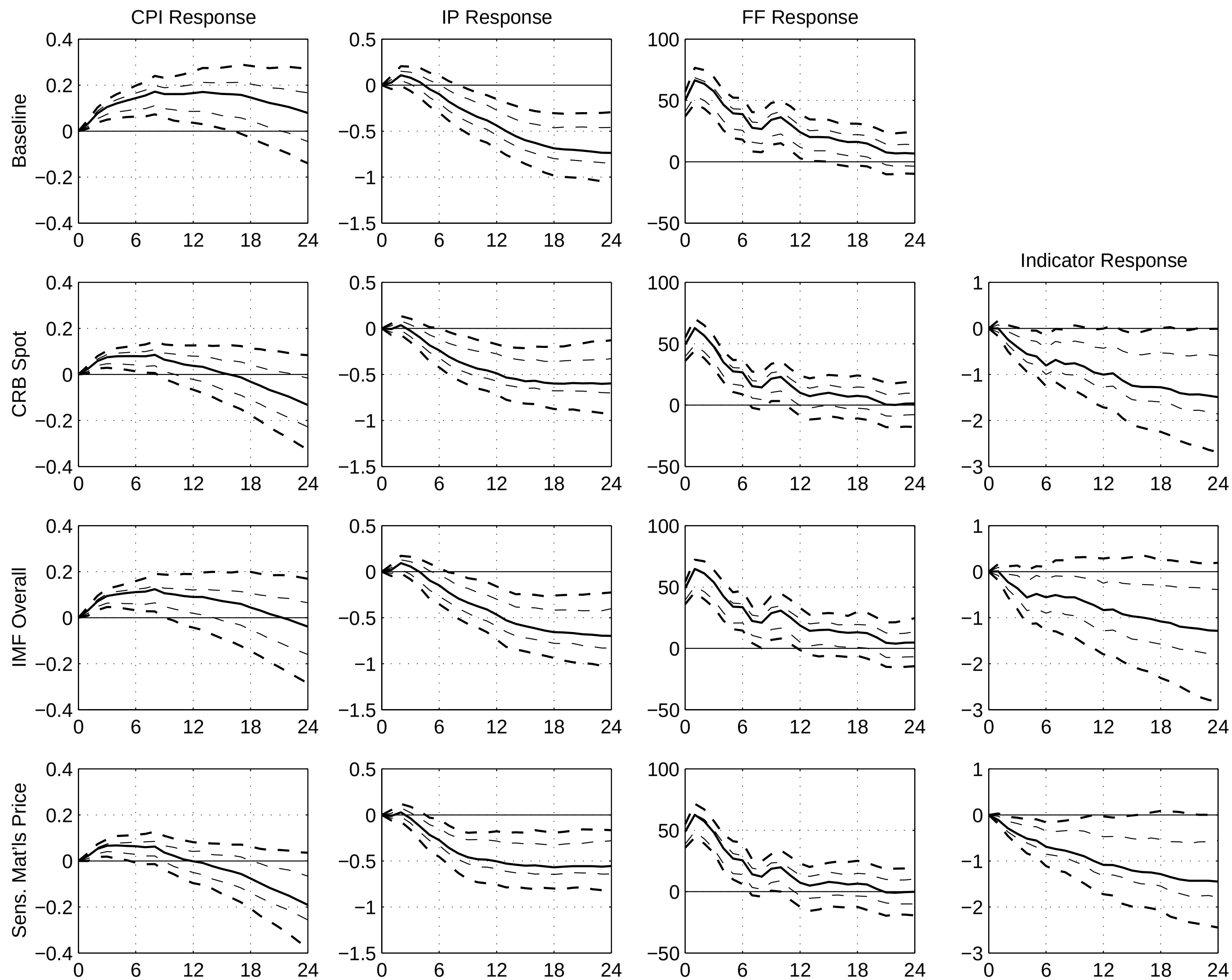


Figure 1.3: Impulse Responses to a Policy Shock for Augmented Model with Fed Funds Instrument, 1959:01 - 1997:06

also appear to reduce the length of the funds 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 commodity prices is barely statistically significant. Nonetheless, the magnitudes of their point estimates are dramatically larger than the CPI responses.

Figure 1.4 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 1.3. The responses of output and the funds rate are unaffected by the inclusion of any of these PPI series.

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 1.3 with figure 1.4 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.

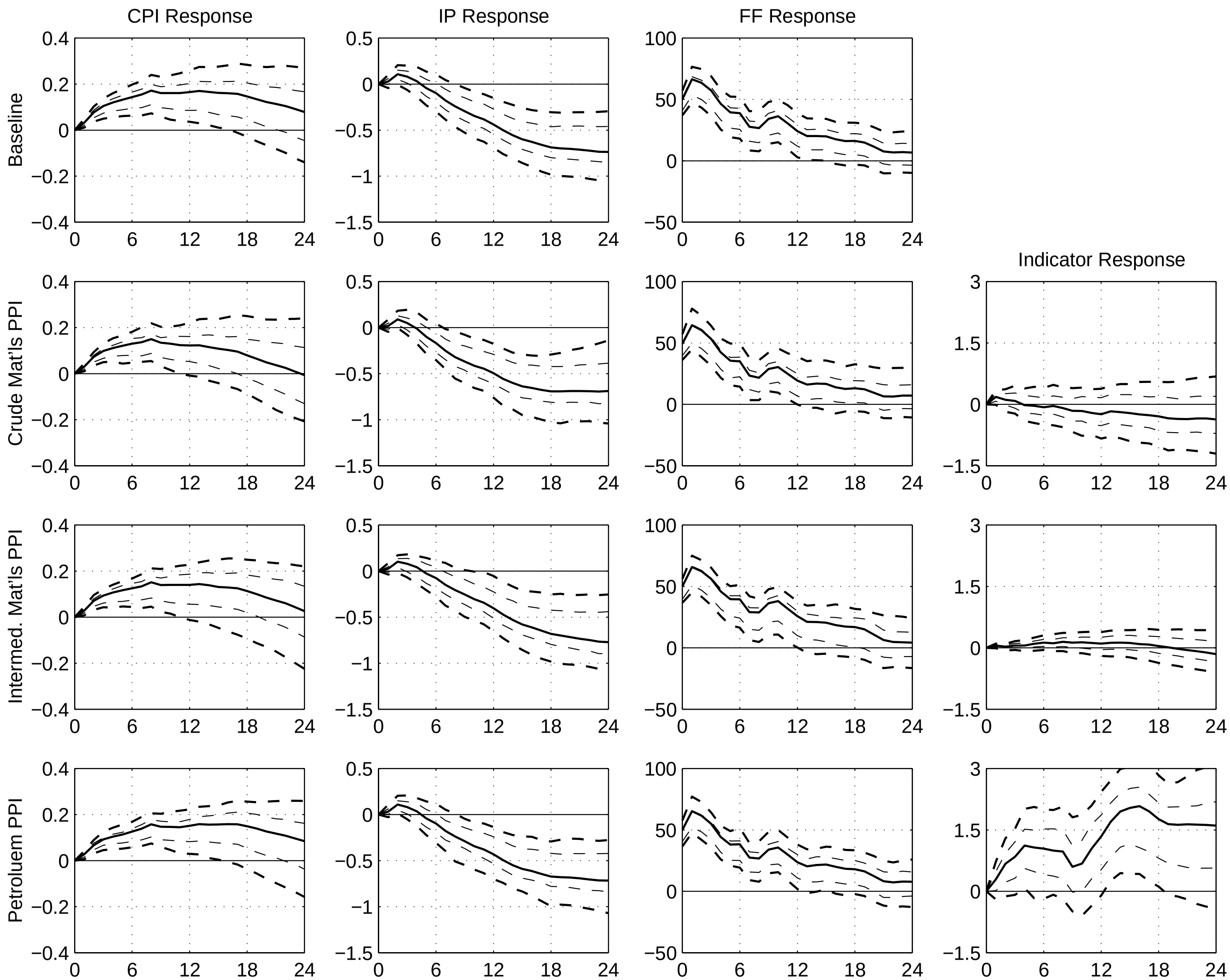


Figure 1.4: Impulse Responses to a Policy Shock for Augmented Model with Fed Funds Instrument, 1959:01 - 1997:06 (continued)

inflation experience.[41] The final row of figure 1.4 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 presents further evidence against the supply shock story.

Figure 1.5 considers interest rate indicators. In contrast with the PPI indexes of figure 1.4, 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 1.1 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 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 1.6 indicate why exchange rates are no longer routinely included in SVAR models for

[41]See Bernanke et al. (1997) for a recent investigation of this issue.

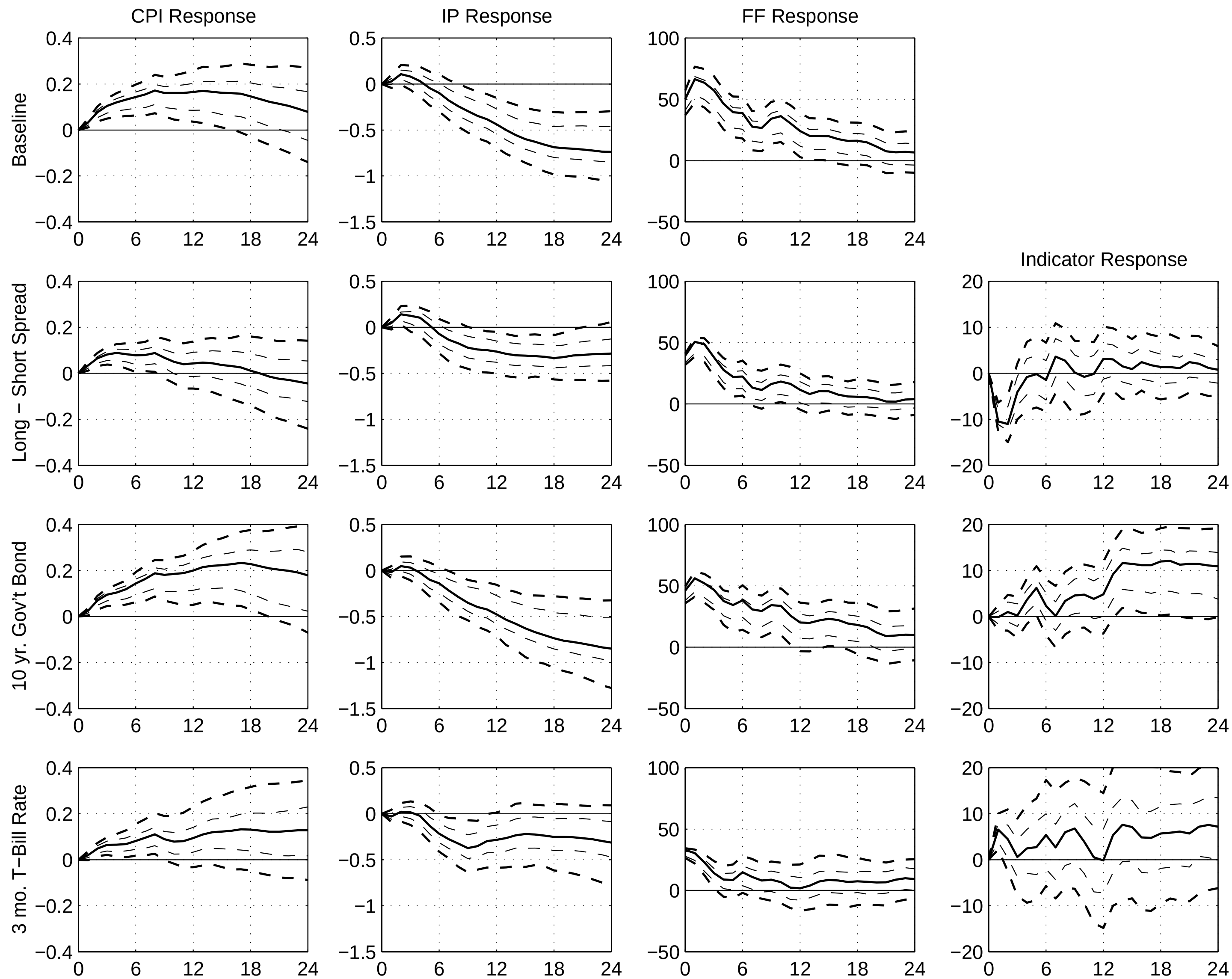


Figure 1.5: Impulse Responses to a Policy Shock for Augmented Model with Fed Funds Instrument, 1959:01 - 1997:06 (continued)

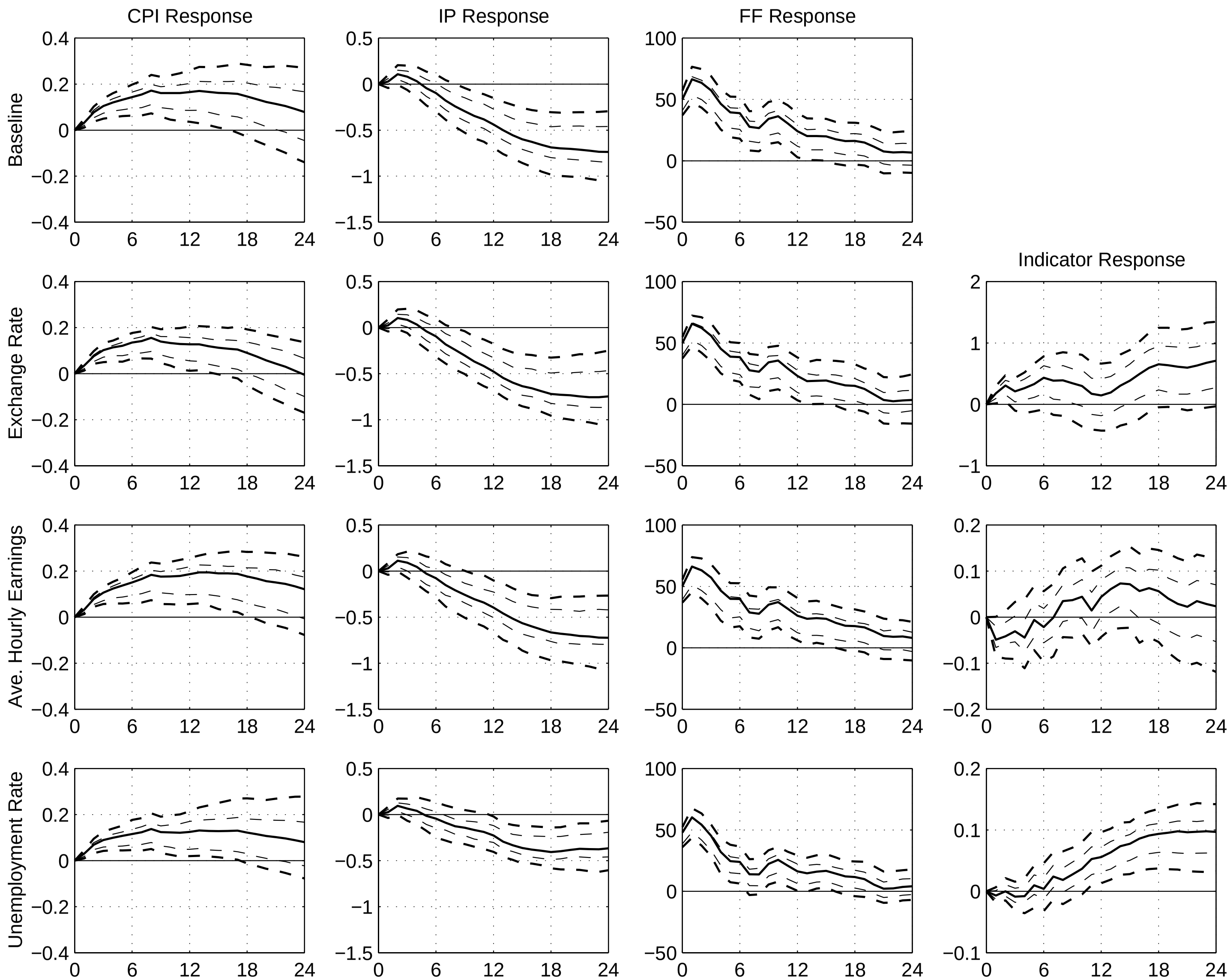


Figure 1.6: Impulse Responses to a Policy Shock for Augmented Model with Fed Funds Instrument, 1959:01 - 1997:06 (continued)

the U.S.:[42] including this variable has no effect on the estimated price response. (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 1.1 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.

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 1.6. 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).[43] Figures 1.7 through 1.9 summarize

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

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

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 1.1) 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 forecast from including the indicator variable shown, as reported in table 1.1. 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 that significantly reduce the positive price level response will sit near the top of the graph. If greater forecasting power implies a less pronounced puzzle, the points plotted should generally lie in a region from the upper-right corner to the lower-left. Few of the graphs reveal such a relationship 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.

These scatterplots nicely contrast the results in figures 1.3 through 1.6 with the data in table 1.1, and reveal numerous cases that undermine the conventional wisdom. First, none of the interest rate measures examined produced noteworthy improve-

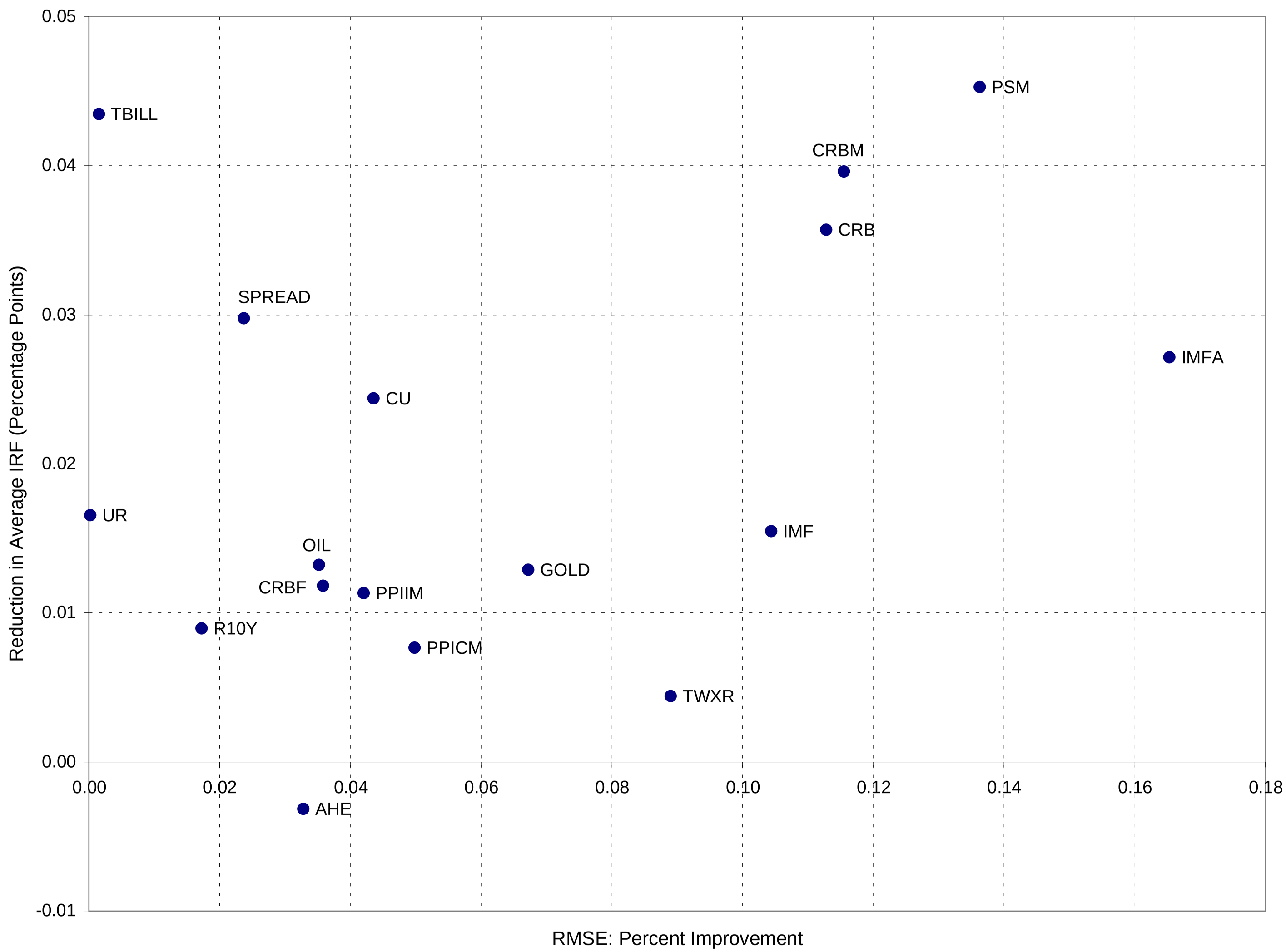


Figure 1.7: Forecasting Power versus Resolution of Price Puzzle, 6 month horizon (1959:01 – 1997:06)

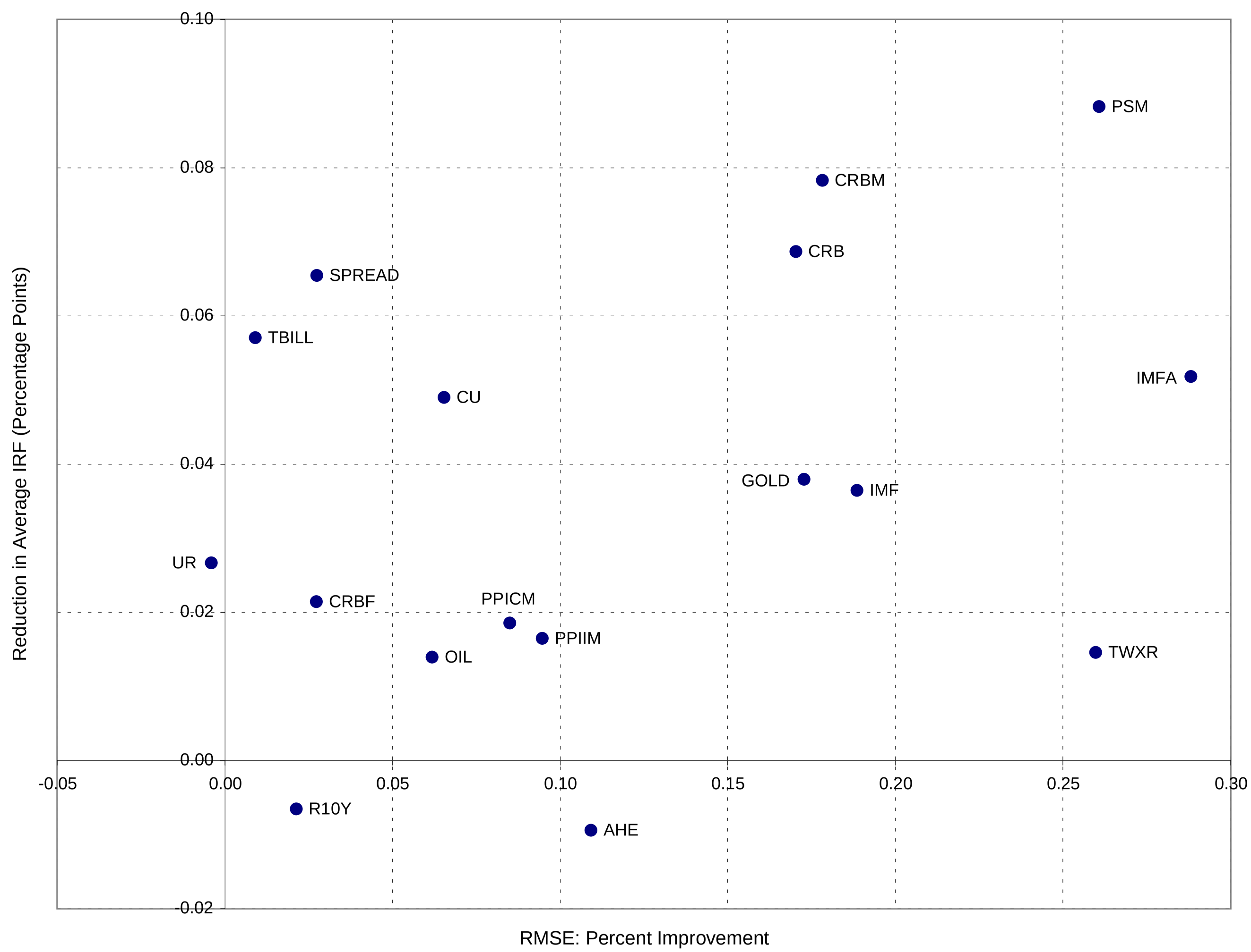


Figure 1.8: Forecasting Power versus Resolution of Price Puzzle, 1 year horizon (1959:01 – 1997:06)

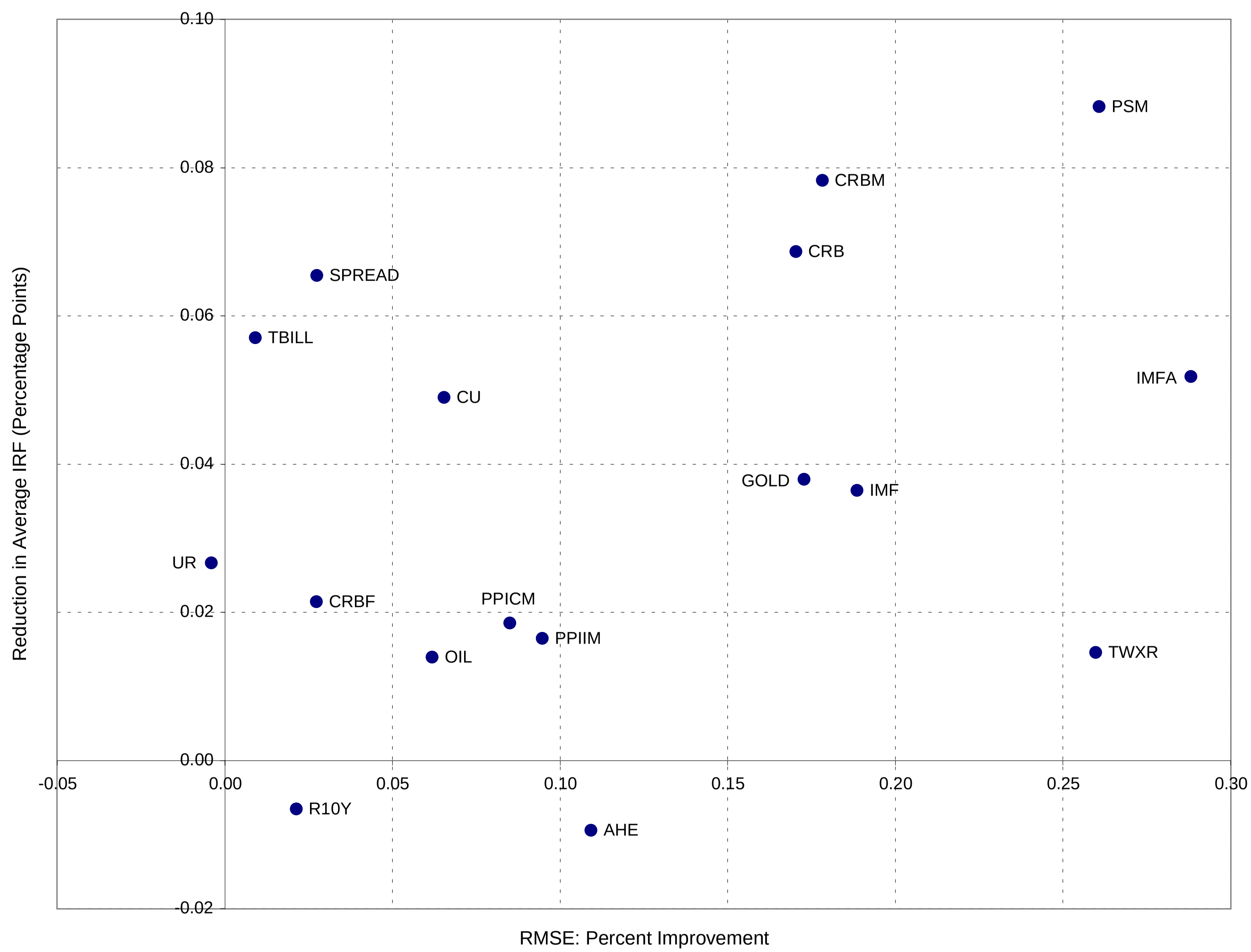


Figure 1.9: Forecasting Power versus Resolution of Price Puzzle, 2 year horizon (1959:01 – 1997:06)

ments 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 1.4: the lack of congruence between the reasonably strong forecasting power of these indicators and their inability to resolve the price puzzle also is inconsistent 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 these results. Some variables that 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. The relative success of the long-short interest rate spread also is unlikely due to its usefulness as a proxy for supply shocks.

In light of the findings of this section, proponents of the conventional view must offer an alternative justification for including commodity prices in their VAR models. Commodity prices may be serving a fairly unique role in the VAR, but one that cannot be attributed primarily to a forecasting future price changes. By breaking the link between indicators and the resolution of the price puzzle, the proposed mechanism that generates a puzzle in the first place is called into question. To better understand both the existence and resolution of the price puzzle, I next investigate sub-sample estimates and ask whether unmodeled variation in the monetary policy reaction function could be a separate source of misspecification.

1.5 Robustness of Sub-sample Estimates

Section 1.3 highlighted the importance of the choice of policy instrument for identifying the effects of monetary policy in a SVAR model. As mentioned in the comparison of the fed funds rate instrument model with Strongin's model, throughout most of 1959 – 1997 sample period the Federal Reserve used the funds rate as its primary policy instrument. However, in October 1979 Chairman Paul Volcker announced a shift from effectively targeting the federal funds rate to explicitly targeting nonborrowed reserves. Thus it would be inappropriate to use a funds rate instrument model 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 explore the empirical measures of monetary policy for the pre-1979 and post-1982 sub-samples.[45]

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 periods. Several researchers recently have argued that the Fed accommodated inflation to a much greater extent

[44] Bernanke and Mihov (1998) report that their version of the funds rate model does not fit this period very well. They also demonstrate that nonborrowed reserves model of Christiano et al. (1996a,b) performs better over the 1979 – 1982 period than any other period they examine.

[45] Several other authors report that their empirical results are sensitive to the sample period over which their monetary SVAR models are estimated. Examples include Brunner (1994), Gordon and Leeper (1994), Strongin (1995), Bagliano and Favero (1998), and Bernanke and Mihov (1998).

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).[46] 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 et al. (1998) confirm these results with a 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.

If the pre-1979 reaction function did take a different form than that post-1982, failure to account for this "regime change" would lead to a misspecified empirical model. From the discussion in section 1.2, if the econometrician incorrectly assumes an unchanged reaction function, the estimated reaction function $g(\cdot)$ may not closely match the true reaction function $f(\cdot)$. As a consequence, the endogenous and exogenous policy components would be mismeasured. Such misspecification could contribute toward generating unexpected impulse responses, perhaps even a price puzzle.

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.[47] Table 1.2 indicates how the stochastic nature of inflation has varied over the

[46]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 et al. (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.

[47]Taylor (1997) and Clarida et al. (1998) suggest that a shift in the nature of the reaction function

Table 1.2: Unit Root Tests for Consumer Prices

	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)			
η_μ 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 the 1% level, ** at the 5% level, and * at the 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 Criteria. All tests include a constant term.

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 an 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 1.2. The first set of results reports the augmented

could have resulted in a change in the dynamics of inflation. Such a conclusion is not unprecedented: 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. Mankiw et al. (1987) claimed that the policies introduced by the Fed after its founding changed the stochastic process describing nominal interest rates. As the 1979 "regime shift" did not involve a change in institutions, it remains an open question whether the nature of the change to the parameters of the reaction function was significant enough to have caused the apparent change in the stochastic process of inflation.

Dickey-Fuller test statistic for the (logarithmic) CPI price level.[48] 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 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, tests with a null of stationarity weakly confirm 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. 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.[49]

Factors other than monetary policy may account for the different behavior of inflation over time illustrated in table 1.2. Nonetheless, comparing the impulse responses across the two sub-periods provides some evidence that the reaction function

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

[49] The results for 1960 – 1979 are consistent with those reported in Barsky (1987). The conclusions from all of these unit root tests are unchanged when alternative price level series are examined. The largest autoregressive roots also are estimated to have similar magnitudes.

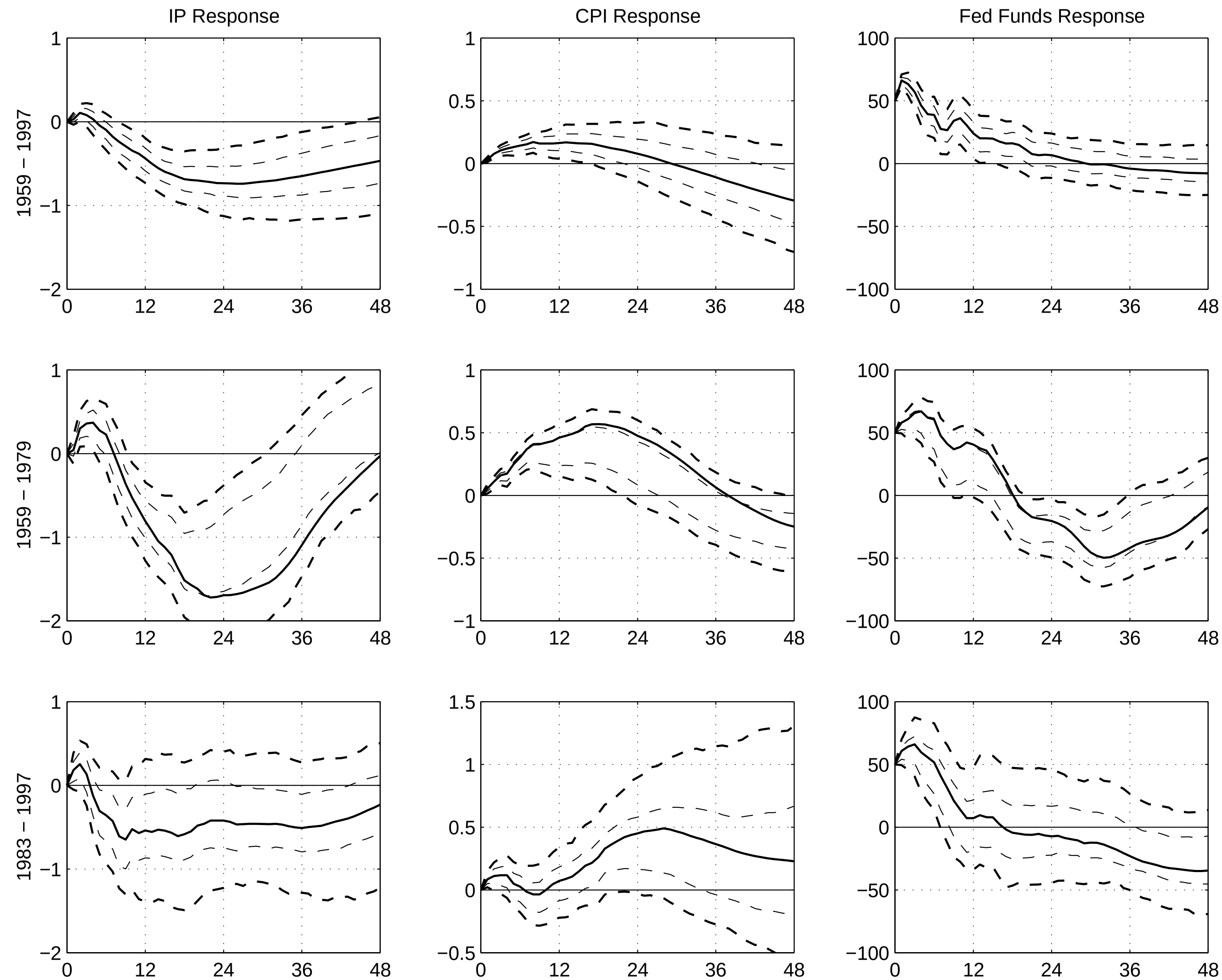


Figure 1.10: Sub-sample Comparison of Impulse Responses to a 50 Basis Point Policy Shock (Baseline Model with Fed Funds Target)

has changed. Figure 1.10 replicates the baseline funds rate instrument model 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 non-borrowed reserves targeting and thus match those used elsewhere in the literature.[50] A comparison down the second column suggests that the price puzzle is a regime-specific phenomenon: the price level response is not statistically significant in the post-1982 sub-sample, while the pre-1979 puzzle is quite protracted (nearly twice as long as in the full sample) and relatively large in magnitude.[51] Each of these graphs in figure 1.10 plots responses to an initial 50 basis point increase in the federal funds rate.

There are other important differences between these 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 with that for the latter 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 1.10 illustrates.

While these results are consistent with the hypotheses of Taylor (1997) and Clarida et al. (1998), they cannot rule out the possibility that other changes in the

[50]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.

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

economy, not directly attributable to monetary policy, have caused the differences across sub-samples. (For further investigation of this issue, see Hanson, 2000.) To the extent the sub-sample differences are due to changes in the policy rule, these results highlight the importance of the reaction function — that is, the "endogenous" component of policy — vis-à-vis the "exogenous" policy shocks. With a few exceptions (Bernanke et al., 1997; Sims and Zha, 1998; Sims, 1998b; Hanson, 2000), the SVAR literature has not attempted to systematically analyze the role of the monetary reaction function in fluctuations.

The lack of any significant response of output or prices to monetary policy shocks in the 1983 – 1997 period may reflect that the Fed has not deviated much from its policy rule during this period. This interpretation would imply an inability to identify the effects of policy shocks with much statistical accuracy. But it may also suggest that the difference in policy between the two periods is due in part to a greater commitment to the estimated policy rule in the later period — or fewer instances that led the Fed to deviate from this rule.

In either case, my results indicate that factors causing sub-sample variation are at least as important for understanding the price puzzle as indicators of future inflation. Figure 1.11 reproduces the scatterplot of figure 1.8 (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 1.8: incorporating any potential inflation indicator — including commodity prices — has a minimal effect upon the positive price response. Were these points plotted on figure 1.8, they would 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

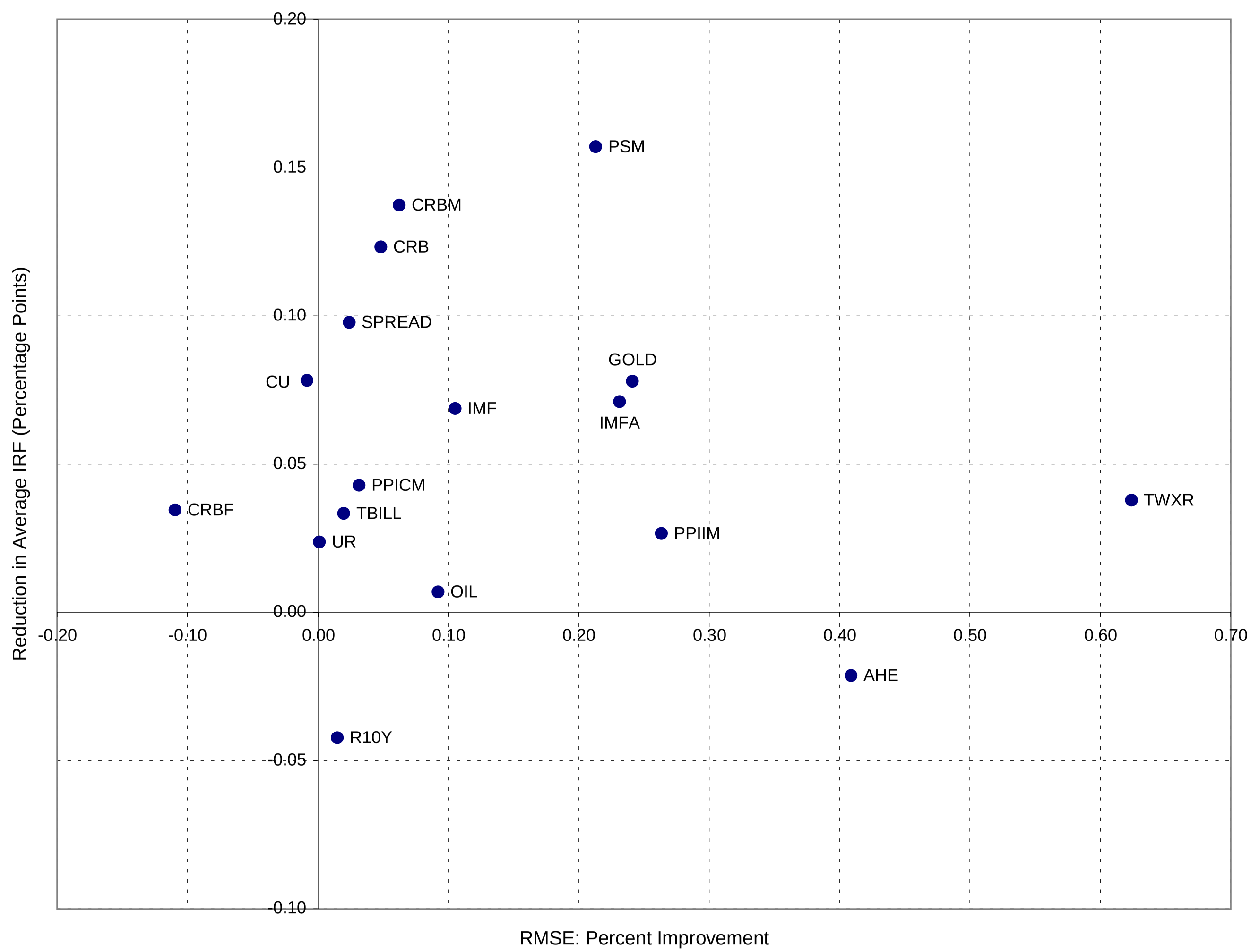


Figure 1.11: Forecasting Power versus Resolution of Price Puzzle, 1 year horizon (1959:01 – 1979:09)

figures.)[52]

Considering the large price puzzle of 1959 – 1979 illustrated in figure 1.10 in conjunction with the paltry impulse response reductions of figure 1.11, a pronounced, statistically significant price puzzle still exists for this sub-sample regardless of the indicator variable included in the augmented model. Plotting a few augmented models in figure 1.12 makes this fact more explicit. 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 1.12 shows the impulse responses when the long-short interest rate spread is used as an indicator variable. Reflecting figure 1.11, the price puzzle is smaller in this case than with commodity prices. Yet a sizable puzzle of 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. In this case, there is *no discernible effect* on the response of the price level (or any other variable) — despite the large oil price shocks that occurred over this period. The "supply shock" story of the price puzzle — that failing to control for these events with an appropriate indicator variable — cannot be rectified with these results. For the sub-sample period in which the supply shock rationalization should be most appropriate, neither commodity prices nor oil prices are able to resolve the

[52]Excluded for brevity, the 6 month and 2 year scatterplots for 1959 – 1979 (which correspond to figures 1.7 and 1.9, respectively, for the full sample) look very similar to figure 1.11.

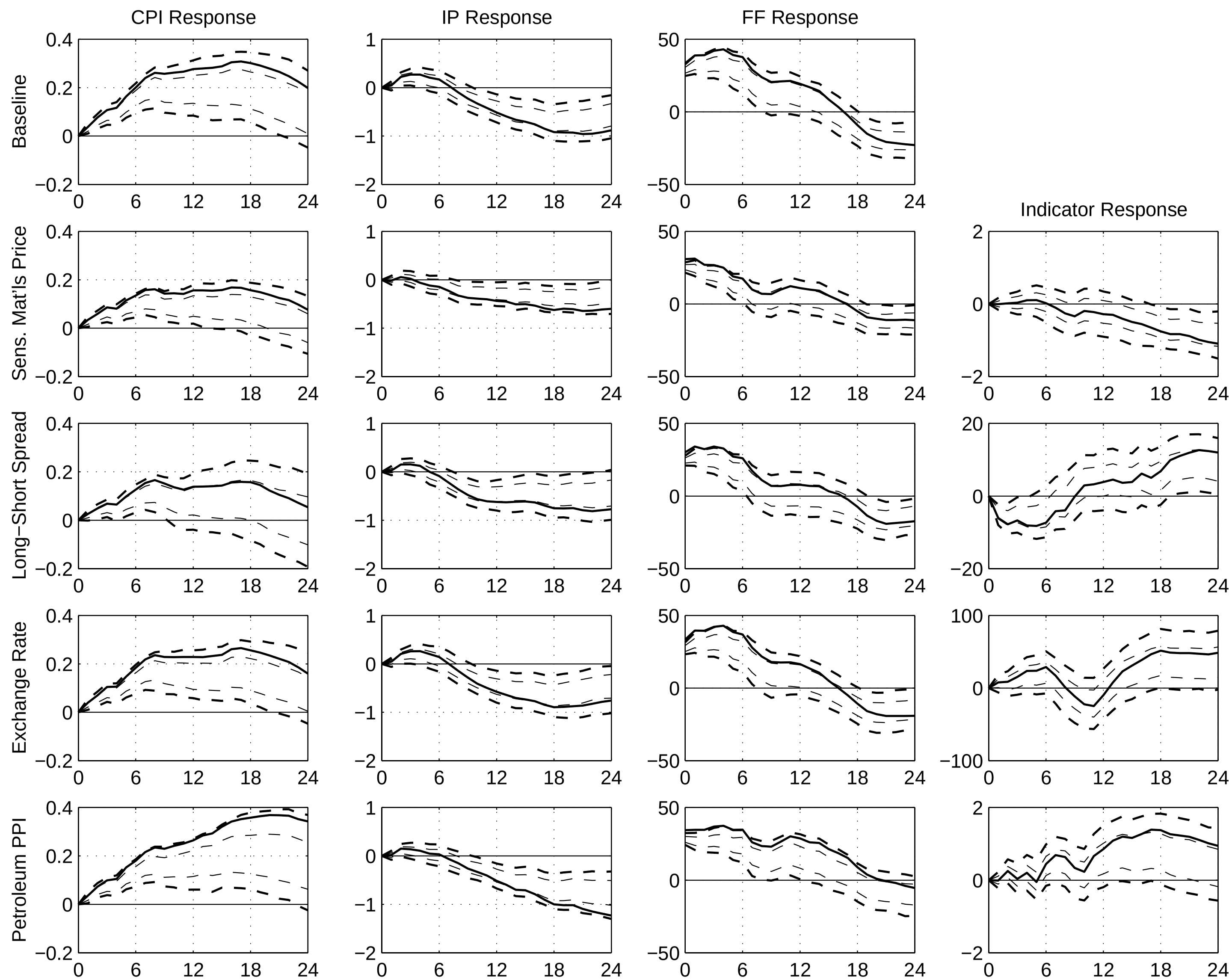


Figure 1.12: Impulse Responses to a Fed Funds Policy Shock, 1959:01 – 1979:09

sizable estimated price puzzle.

For completeness, figure 1.13 illustrates the same impulse responses as figure 1.12, but for the latter 1982 – 1997 sub-sample. In general, the 1982 – 1997 results do not reveal any statistically 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 — perhaps due to changes 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 price puzzle. In particular, 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 that is resolved by the inclusion of commodity prices (or, arguably, 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 for its resolution.

1.6 Conclusion

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

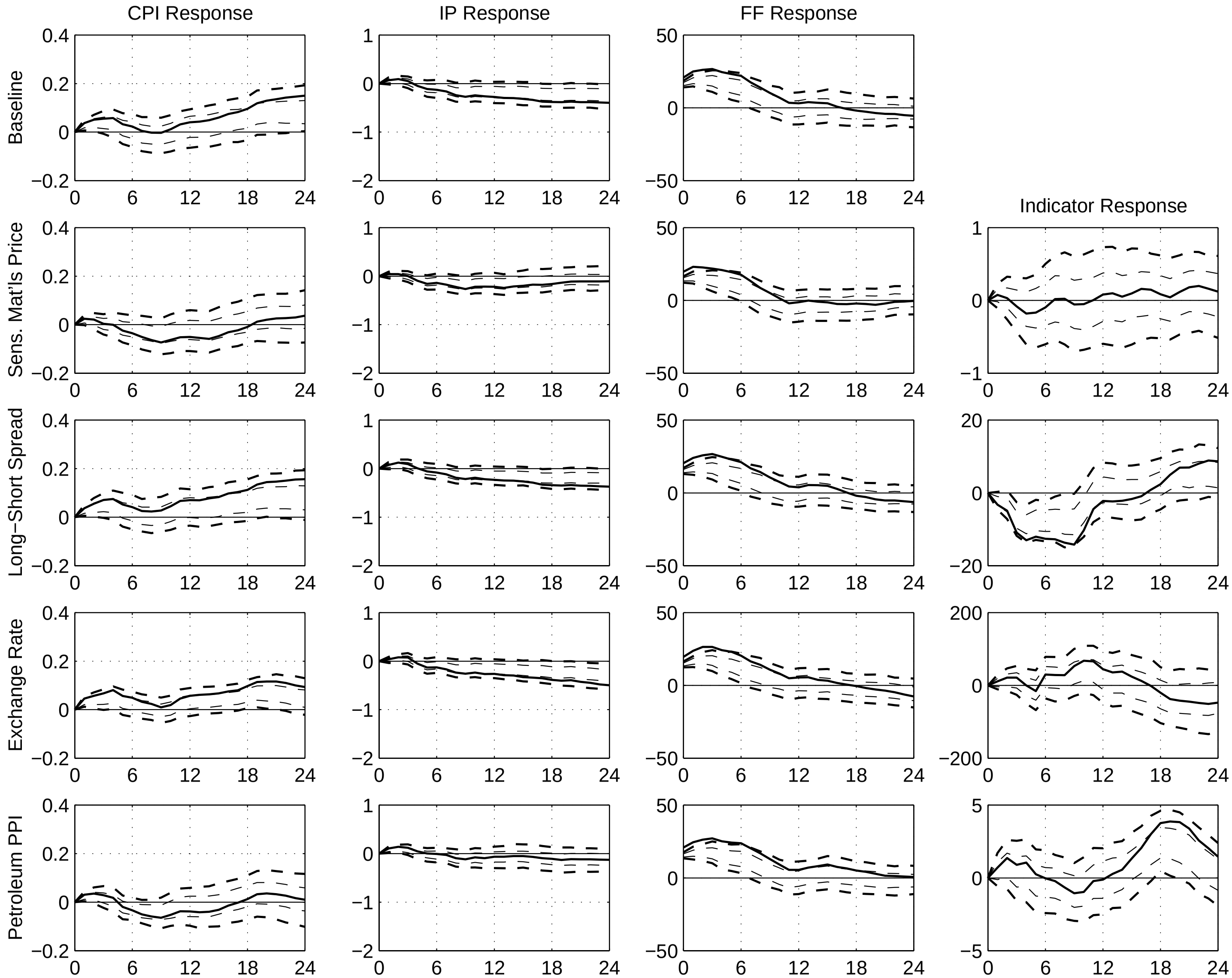


Figure 1.13: Impulse Responses to a Fed Funds Policy Shock, 1983:01 – 1997:12

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 1.4 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 prices (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 1.5 I investigate how the estimated impulse response functions vary within the sample period. 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. One plausible interpretation is that 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 a limitation of the recursive identification schemes employed by the models I examine. Even so, Christiano et al. (1998) point out that non-recursive models to date still impose similar identifying 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.

Finally, the results in section 1.5 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 sub-optimal 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.

CHAPTER II

Varying Monetary Policy Regimes: a Vector Autoregressive Investigation

2.1 Introduction

Has the prolonged U.S. economic expansion of the last decade and a half been due to good fortune or good policy? In the popular and business press, the Federal Reserve Board — and Chairman Alan Greenspan in particular — often receives nearly exclusive accolades for the sustained growth and moderate to low inflation experienced since the mid–1980s.[1] Recent academic work also has focused upon the Fed, crediting monetary policymakers for a share of the recent economic outcomes while blaming them for the poor performance experienced in the late 1960s and 1970s. (De Long, 1997; Taylor, 1997; Sargent, 1998.)

In contrast, a conventional view of U.S. economic performance during that period posits an unavoidable policy dilemma in the face of adverse supply shocks. By implication, the absence of such shocks would be largely responsible for current conditions. Thus this recent literature, by focusing on the contribution of Federal Reserve policy to economic fluctuations, raises a pair of questions. First, is an "improvement" in monetary policy the primary cause of the more recent economic

[1] Time Magazine's February 15, 1999, cover article on "The Committee to Save the World" (featuring Alan Greenspan, Robert Rubin and Lawrence Summers) is but one recent example.

experience, or have policy makers benefited from some transformation in the nature of the economy itself? Put more glibly, has Alan Greenspan been smart or lucky — or both?

Once attention is focused upon monetary policy, a second question naturally arises: what attributes of policy have changed? In assessing the role of monetary policy, it is important (indeed, from the standpoint of econometric identification, necessary) to distinguish between *exogenous* (unforecastable) innovations to monetary policy and the *endogenous* response of policy to the state of the economy. This dichotomy allows researchers to determine whether movements in the policy instrument can be prescribed to factors other than policy actions by the monetary authority. Interestingly, the two main empirical literatures on monetary policy have placed very different emphasis on each of these components.

The vector autoregressive (VAR) literature, which arguably represents the predominant empirical paradigm for monetary economics, has focused almost exclusively upon the effects of monetary policy shocks.[2] VAR practitioners claim to identify policy shocks with a limited set of assumptions, which accounts for much of the popularity of this approach. However, the general lack of attention to the estimated policy reaction function in these models has been a subject of recent criticism (e.g. Rudebusch, 1998a; McCallum, 1999). By narrowly focusing upon the estimated shocks to policy, Cochrane (1994, p. 331) muses that existing VAR research "may simply be asking the wrong *question*." [Emphasis in original.] Although some authors recently have investigated the role of the estimated policy rule in semi-structural VAR models (see Bernanke et al. (1997), Sims (1998b), and Sims and Zha (1998) for examples),

[2]Recent VAR papers that investigate post-war U.S. monetary policy include Bernanke and Blinder (1992), Strongin (1995), Christiano et al. (1996a,b), Bernanke et al. (1997), Bernanke and Mihov (1998), Bagliano and Favero (1998), and Sims and Zha (1998). Cochrane (1994) and Christiano et al. (1998) present overviews of this literature.

such work remains the exception.

Beginning with Taylor (1993), a distinct literature has focused on simple specifications of a feedback rule for monetary policy, with the federal funds rate serving as the policy instrument. No attempt is made to interpret or explain the residuals of the estimated rule. By and large, this "Taylor rules" literature and the monetary VAR literature have evolved separately.[3] Recent work in the Taylor rules literature (e.g. Taylor, 1997; Clarida et al., 1998; Judd and Rudebusch, 1998) provides a single direct answer to the pair of questions posed at the opening of this paper: by failing to (endogenously) respond forcefully enough to inflationary pressures, "inappropriate" Federal Reserve policy introduced instability into the economy prior to 1980.

This answer contrasts notably with the more common, "textbook" view of Fed policy during these years: in the face of circumstances largely beyond its control — adverse supply shocks and (possibly) domestic political pressure — the Fed had to choose between high inflation or high unemployment. While the choice facing the Fed may have been difficult and undesirable, the conventional view does not regard its policy response as negligent.

It is important to note that much of the Taylor rules literature *presumes* monetary policy to be the culprit: the question of changing dynamics for the broader economy is not explored. In fact, since many researchers estimate the Taylor rule as a single equation model (or within a small, stylized macroeconomic model) this question *cannot* be investigated in this context. Were a Taylor rule a representation of a structural behavioral equation, this concern would not be important. But since the specification of the rule is sensitive to the equations that govern the broader

[3]A third literature employs large-scale macroeconometric models to estimate the effects of alternative policy regimes. This approach is commonly employed by policymakers (see, e.g. Bryant et al., eds, 1993) but less so among academic researchers; Ray Fair's FAIRMODEL is one notable exception.

economy, failure to account for instability residing elsewhere in the model could lead to improper inference about the role of the policy rule. By positing a more general specification for the economy, in principle an SVAR approach can overcome this shortcoming of much of the current literature that utilizes Taylor rule specifications.

This paper attempts to reconcile the motivation of the recent Taylor rules literature with the VAR approach, by focusing attention on the estimated policy reaction function and its stability over time. Section 2.2 develops the statistical tools for this investigation, and discusses the interpretation of the components of an estimated semi-structural VAR model. Section 2.3 explores the properties of the estimated reaction function and policy shocks across various sub-sample periods. But this paper also moves beyond either current literature by exploring the possibility of significant instability residing in the non-policy part of the model. Section 2.4 pursues methods to quantify the economic significance of some of the instability discovered in section 2.3. To foreshadow the main result, despite evidence of instability in the practice of monetary policy, these changes do not appear able to explain the varying economic outcomes observed in the data. Section 2.5 concludes.

2.2 Estimation Technique

As stated in the introduction, the vector autoregressive model constitutes the most common econometric approach in modern empirical macroeconomics. Its widespread use is due in part to the fact that it stresses simultaneity and joint endogeneity of the variables under consideration. In trying to disentangle the question posed in the introduction, it is essential to take into account the feedback of policy into the other (non-policy) variables — particularly in a dynamic setting. Despite their other shortcomings, VAR models do allow a prominent role for the dynamic interaction of variables rather than imposing strong (and often data-inconsistent) restrictions upon the equations of the model.[4] The sheer volume of VAR-oriented work suggests it is worth investigating whether this paradigm in some form can accommodate questions about changes in the policy rule.

2.2.1 Identification and Estimation of a Monetary VAR

Begin with the following representation for the structural model of the economy:

$$X_t = \Theta(L)\,\varepsilon_t \tag{2.1}$$

where $X_t = (x_{1t}, x_{2t}, \ldots, x_{nt})'$ is an n-vector of observed endogenous variables, ε_t is an n-vector of unobserved structural disturbances, and $\Theta(L)$ is a (possibly infinite-order) polynomial in the lag operator L:

$$\Theta(L) = \Theta_0 + \Theta_1 L + \Theta_2 L^2 + \cdots + \Theta_q L^q + \cdots .$$

In this vector moving average form, the ε_t form the "white-noise building blocks" of the economy, while $\Theta(L)$ reflects how these shocks are propagated through the economy. Thus each variable in X_t is written in equation (2.1) as an infinite sum of all current and past innovations to the economic system. The covariance matrix for the structural disturbances is given by

$$E\big[\varepsilon_t \varepsilon_t'\big] = \Omega\,. \tag{2.2}$$

Since the structural disturbances are mutually and serially uncorrelated by assumption, the covariance matrix Ω will be diagonal.

[4]In essence this is the same critique originally raised by Sims (1980): a VAR freely estimates the dynamic relations of the data, for which any particular restriction would lack theoretical justification.

Under standard regularity conditions, the system in equation (2.1) is invertible and can be written in vector autoregressive (VAR) form as:

$$\Phi(L)\, X_t = \varepsilon_t\,, \tag{2.3}$$

where $\Phi(L) = \Theta(L)^{-1}$ is also a polynomial in the lag operator:

$$\Phi(L) = \Phi_0 - \Phi_1 L - \Phi_2 L^2 - \cdots - \Phi_p L^p\,. \tag{2.4}$$

Let p be the (assumed finite) lag order of $\Phi(L)$, chosen to be large enough to capture the dynamics of the data.

The structural form of a VAR, as with any simultaneous equations model, cannot be estimated directly. Instead, one must estimate a reduced-form model and then impose identifying restrictions that allow the structural parameters of interest to be uncovered. The reduced-form model is

$$A(L)\, X_t = v_t\,, \tag{2.5}$$

where v_t is an $(n \times 1)$ vector of reduced-form residuals.[5] $A(L)$ is a lag polynomial in the lag operator L, also of order p:

$$A(L) = I - A_1 L - A_2 L^2 - \cdots - A_p L^p. \tag{2.6}$$

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 is the $n \times n$ identity matrix. The leading identity matrix in equation (2.6) reflects the normalization of placing the variable x_{it} on the left-hand side of the i^{th} equation of the system. The simultaneity of the reduced-form system is summarized by the covariance matrix,

$$E\big[v_t v_t'\big] = \Sigma\,, \tag{2.7}$$

[5]This notation ignores intercept terms for convenience; they are included in the estimation.

in which some $\sigma_{ij}\,(i,j=1,\,\ldots,\,n;\; i\neq j)$ are non-zero.

Pre-multiplying the structural VAR polynomial in equation (2.4) by Φ_0^{-1} yields an expression that is observationally equivalent to the reduced-form VAR polynomial given in equation (2.6). Thus,

$$A_k = \Phi_0^{-1}\Phi_k \text{ for } k=1,\,\ldots,\,p\,. \tag{2.8}$$

Then it follows from equations (2.3) and (2.5) that

$$v_t = \Phi_0^{-1}\varepsilon_t\,, \tag{2.9}$$

which in turn implies the following relationship between the structural covariance matrix of equation (2.2) and the reduced-form covariance matrix of equation (2.7):

$$\Sigma = E\big[v_t v_t'\big] = E\big[\Phi_0^{-1}\varepsilon_t\varepsilon_t'\Phi_0^{-1\prime}\big] = \Phi_0^{-1}\Omega\,\Phi_0^{-1\prime}\,. \tag{2.10}$$

The reduced-form VAR model can be estimated equation-by-equation via OLS. The resulting parameter estimates $\widehat{A(L)}$ and $\widehat{\Sigma}$ can be used to recover the structural parameters once sufficient restrictions are imposed to just identify the model. To see this, recall from equation (2.10) that $\Sigma = \Phi_0^{-1}\Omega\,\Phi_0^{-1\prime}$. Since Σ is symmetric, it has $\frac{n(n+1)}{2}$ unique elements. However, $\Phi_0^{-1}\Omega\,\Phi_0^{-1\prime}$ has n^2 elements,[6] thus requiring $\frac{n(n-1)}{2}$ restrictions. Given $\widehat{\Sigma}$ and a suitable set of restrictions, Φ_0 can be calculated; from equation (2.8) the remaining Φ_k, $k=1,\,\ldots,\,n$, (and therefore the Θ_k) matrices then can be computed.

Placed on the off-diagonal elements of Φ_0, these are restrictions on the mapping from the estimated reduced-form residuals v_t into the structural innovations ε_t.

[6]This total is due to an implicit normalization of the structural model: either the diagonal elements of Φ_0 or the variances of the structural innovations can be normalized to one. In the latter case, $\Omega \equiv I$.

(See equation (2.9).) In other words, these restrictions govern the contemporaneous co-movement of the variables of X_t. (From equation (2.10), these also have the interpretation as covariance restrictions.)

One set of restrictions employed by the class of models examined in this paper have the interpretation of adjustment lags of the private sector to monetary policy innovations. These have a number of possible interpretations, including lags in the ability to gather and process information about monetary policy, sluggish adjustment of prices and/or production — anything that will prevent private sector agents from responding to an unexpected policy innovation within the period it occurred. Obviously, the plausibly of such restrictions depends critically on the length of a "period;" in this paper I consider data sampled at a monthly frequency from 1966:01 –1997:12.[7]

To make this set of restrictions operational, let $1 \leq m < n$ variables represent monetary policy instruments or intermediate targets and the remaining $(n-m)$ variables represent "activity" variables, such as output and the price level. Partition the X_t vector as $X_t = [Y_t \;\; M_t]'$, where Y_t, the $(n-m)$-vector of activity variables, is ordered before M_t, the m-vector of policy variables. With the variables ordered in this fashion, the restrictions induce a block-recursive structure upon the model. Writing out the reduced-form model (2.5) in terms of the structural parameters yields the block-matrix expression:

$$\begin{bmatrix} (\Phi_0^{-1})_{YY} & 0 \\ (\Phi_0^{-1})_{MY} & (\Phi_0^{-1})_{MM} \end{bmatrix} \Phi(L) \begin{bmatrix} Y_t \\ M_t \end{bmatrix} = \begin{bmatrix} (\Phi_0^{-1})_{YY} & 0 \\ (\Phi_0^{-1})_{MY} & (\Phi_0^{-1})_{MM} \end{bmatrix} \begin{bmatrix} \varepsilon_{Y,t} \\ \varepsilon_{M,t} \end{bmatrix},$$

where the restrictions discussed above are reflected in setting the upper right-hand block of the Φ_0^{-1} matrix to zero. This assumption results in $m \times (n-m)$ restrictions

[7]Many of the other studies in this literature use quarterly data, which implies a full three months between the time a monetary policy action is undertaken and the time agents are allowed to respond. While these timing restrictions are clearly only an approximation to reality, the approximation is arguably worse with quarterly than with monthly data.

imposed on the system.

With the upper right-hand block $(\Phi_0^{-1})_{YM} = 0$, the activity variables cannot react within a period to innovations in any of the policy variables, $\varepsilon_{M,t}$. Conversely, so long as the (i,j) element of $(\Phi_0^{-1})_{MY}$ is non-zero, the i^{th} policy variable will respond to an innovation in the j^{th} activity variable. By leaving this block unrestricted, each of the policy variables is allowed to react to contemporaneous movements in all of the activity variables. In particular, the Y_t vector includes monthly measures of output (log of the industrial production index), the price level (log of the consumer price index) and an index of commodity prices (also in logs).

Additional restrictions placed on the lower right-hand block of the Φ_0^{-1} matrix allow one of the estimated equations in the policy block to represent the monetary policy reaction function. As has been common since Bernanke and Blinder (1992), the federal funds rate serves as the policy instrument. As in Christiano et al. (1996a,b) and Bernanke and Mihov (1998) it is ordered first in the M_t block, followed by measures of nonborrowed and total reserves (both in logs). A recursive structure for the $(m \times m)$ $(\Phi_0^{-1})_{MM}$ block implies that the funds rate does not respond to contemporaneous innovations to the reserves market variables. It does respond to lagged movements in these variables, which reflect a combination of supply and demand shocks within the market for reserves.

With these restrictions in place, the residuals of the funds rate equation are identified as *exogenous* monetary policy shocks. That is, given current and lagged information about the state of the economy, the residuals capture the unforecastable component of changes in the policy instrument. In semi-structural models, the funds rate equation commonly is associated with the monetary policy reaction function: the component of the policy instrument that responds *endogenously* to the state of the economy.

2.2.2 Exogenous and Endogenous Components of Policy

Most VAR research to date has focused on exogenous innovations to monetary policy. Despite nearly two decades of published research, this literature apparently has not yet reached a consensus as to whether exogenous monetary policy shocks have any economic effects: indeed, Sims (1998a) and Cochrane (1994, 1998) each claim to have found consensus in the literature — and reach exactly *opposite* conclusions regarding the nature of this consensus! While Sims (1998a, p. 933) argues that "[r]esponses of real variables to monetary policy shifts [i.e. shocks] are estimated as modest or nil, depending on the specification" (see also Sims and Zha, 1998), Cochrane claims the real effects are generally recognized as sizable.[8] Even the effect of policy on certain nominal variables is still open to debate, as recent work on the liquidity puzzle (e.g. Pagan and Robertson, 1998) and the price puzzle (e.g. Hanson, 1998) demonstrate.

The research in this area regularly uses the phrase "monetary policy" as a shorthand for "exogenous monetary policy shocks." By focusing on the implications of variance decompositions and impulse response functions attributable to these shocks, this literature largely has ignored the estimated monetary policy rule or reaction function.[9] This is an odd approach for several reasons. First, the shocks account for

[8] Rudebusch (1998b, p. 943) counters that Sims's view on this point is "idiosyncratic." Bernanke et al. (1997, pp. 95–6) write that "despite ongoing debates about precisely how the policy innovation should be identified [in a VAR], the estimated responses of key macroeconomic variables to a policy shock are reasonably similar across a variety of studies and suggest that monetary policy shocks can have significant and persistent real effects." Bernanke and Mihov (1998) also provide support for this latter view. My own reading of the literature is that the preponderance of research supports an economically significant role for exogenous policy shocks.

[9] In this paper I use these two terms interchangeably. Any regular pattern of policy responses to the economy will constitute a "rule" for my purposes, whether or not such responses are codified into law.

a very small fraction of the total variation of the policy instruments: the majority is accounted for by the endogenous component of policy. Second, the form of the policy rule is as much a decision of policy makers as the shocks to that rule, if not more so.[10]

Even if the debate concerning the role of the exogenous monetary policy shocks were to be resolved against the finding of significant effects (and as commented above, more evidence points in the opposite direction), it would say nothing as to the importance of the reaction function. To assert that monetary policy "does not matter" requires not simply an analysis of the exogenous shocks to monetary policy, but also a claim that the policy rule that characterizes the endogenous response of policy to other shocks in the economy is not an important propagation mechanism. The VAR literature largely has ignored the role the policy reaction function may play in amplifying or mitigating the various shocks of the system, thereby ignoring a sizable — perhaps the main — part of the story. In the next two sections, I investigate measures of the importance of the estimated reaction function relative to the policy shocks.

2.3 Structural Breaks and Policy Regimes

Both narrative and statistical studies regularly cite October 1979 as the date of a major change in U.S. monetary policy. During that month, Chairman Paul Volcker announced a shift in the operating procedures of the Board of Governors designed to dramatically reduce inflation and instill credibility. This experiment with monetary targeting occurred over several years, and commonly is viewed as

[10]In recognizing the importance of the reaction function in a VAR context, Bernanke et al. (1997) briefly discuss sub-sample results, but do not investigate how the policy rule may have varied over time, or whether such variation may have significantly affected the propagation of shocks in the models.

concluding in the early 1980s. Meulendyke (1998) provides an historical account of this shift to targeting money and using nonborrowed reserves as the primary instrument. In light of the identifying assumptions of section 2.2, there are *a priori* reasons to isolate this period from the rest of the sample: as the policy instrument shifted from the funds rate to nonborrowed reserves, the residual of the funds rate equation no longer would represent the exogenous policy shocks.

Henceforth I will use the phrase "Volcker disinflation" to refer to the 1979:10 to 1982:12 period. Some recent empirical research has explicitly recognized the uniqueness of this period as well.[11] In one such study, Fair (1999) has estimated a model with a monetary reaction function that bears some similarity to a Taylor-type rule. Excluding the 1979Q4 to 1982Q3 period from his full sample (1954Q1 to 1999Q3), Fair reports that he cannot reject the null of parameter constancy of his policy rule. This conclusion contrasts sharply with other results in the Taylor rules literature.[12]

The above discussion highlights one main unresolved issue in the current empirical monetary policy literature, broadly construed: accounting for the "Volcker disinflation," does one find significant differences in the practice of policy before and after this experiment? Given the conflicting research to date, as well as the questions raised in the introduction of this paper, I begin by considering evidence of instability for the VAR model as a whole. Then I turn my attention to the estimated

[11] For example, Strongin (1995) investigates the period 1979:11 to 1982:10; Bernanke and Mihov (1998) use 1979:10 to 1982:10. Some in the Taylor rules literature have excluded the 1979:10 to late-1982 period as well; see Taylor (1997). All of these authors recognize the Volcker disinflation as a fundamentally different policy regime (in terms of the form of the reaction function or of the policy instrument), and not as a discrete event (perhaps due to private agents only learning slowly about the new regime, or the length of time necessary for the Fed to acquire credibility). On the other hand, some authors have argued against any change in policy over the post-war sample; see Sims (1998a,b) and Christiano et al. (1998).

[12] It is unclear the extent to which Fair's results depend on his sample (namely excluding the Volcker disinflation) as opposed to his specification of the reaction function. See Fair (1999) for details.

components of monetary policy.[13]

2.3.1 Overall Stability of the VAR Model

I take up the question of a change in regime first by investigating the parameter stability of the full (reduced-form) VAR model. Table 2.1 examines the Akaike and Schwarz Information Criteria for VAR models estimated across several combinations of sample period and lag length. The information criteria are log likelihood statistics, with penalties for additional parameters. (The SIC penalizes additional parameters more heavily than the AIC.) Heuristically, the lower the AIC or SIC measure, the better the "fit" of the equation. The information criteria are computed for the reduced-form models as described in section 2.2.

Sims (1998a,b) has argued in favor of using these measures to "test" for overall stability of a VAR model. By this metric, a model estimated over a full sample is preferred to one estimated separately over two sub-sample periods if the former has a lower value of the AIC or SIC than a weighted-average of the information criteria for the latter case. The weights are determined by the proportion of the full sample in each sub-sample estimation.[14]

Table 2.1 permits an analysis of structural instability for two variants of the reduced-form model. The first is estimated over the full 1966:01 – 1997:12 sample, and breakpoints after October 1979 and December 1982 are considered. Based on the above discussion, these dates reflect the onset and conclusion of the "Volcker disinflation" experiment. The second investigates the joint sample, with the intervening

[13]I do not undertake a search for the date of the regime break, since there is no serious disagreement in the literature about the *candidate* breakpoint. (Moreover, tests of unknown breakpoints tend to have low power against local alternatives.) A preliminary investigation suggests that my results are insensitive to associating the end of the Volcker disinflation with any month between 1982:09 and 1982:12, inclusive.

[14]This can be shown by comparing the log likelihood functions for the two cases. I thank Phil Howrey for helping to clarify this point.

Table 2.1: Comparison of Akaike and Schwarz Information Criteria

	4 lags		*6 lags*		*8 lags*		*12 lags*	
Sample Period	AIC	SIC	AIC	SIC	AIC	SIC	AIC	SIC
1966:01 – 1997:12:								
Fixed Coefficients	−49.7759	−49.2471	−49.8312	−49.1754	−49.9992	−49.2154	−50.0105	−48.9676
Break after 1979:09	−50.5136	−49.6385	−50.6560	−49.5670	−51.1353	−49.8293	−51.3810	−49.6311
Break after 1982:12	−50.9337	−50.0563	−50.9958	−49.9039	−51.1078	−49.7982	−50.8802	−49.1253
1966:01 – 1979:09, 1983:01 – 1997:12:								
Fixed Coefficients	−51.5613	−50.9832	−51.5843	−50.8639	−51.6909	−50.8255	−51.6944	−50.5313
Separate Pre- and Post-Volcker periods	−52.2431	−51.2970	−52.2427	−51.0644	−52.3358	−50.9216	−52.1599	−50.2618

Notes: Akaike Information Criteria: $AIC = \log(|\Sigma|) + \frac{2}{T}\left(n(p+1) + \frac{n(n+1)}{2}\right)$, Schwarz Information Criteria: $SIC = \log(|\Sigma|) + \frac{\log(T)}{T} \cdot \left(n(p+1) + \frac{n(n+1)}{2}\right)$, where $|\Sigma|$ is the determinant of the reduced-form covariance matrix, n is the number of variables in the VAR, p is the lag length, and T is the sample size. Within a column, a more negative value of an information criteria suggests a better fit. Values for information criteria are not directly comparable across different lag lengths shown, as initial values for estimation are drawn from within the stated sample period.

1970:10 – 1982:12 period excluded. As cited above, some authors have suggested that, outside of the "Volcker disinflation," there has not been any statistically discernible change in the monetary policy regime.

The first row of table 2.1, labeled "Fixed Coefficients," lists the AIC and SIC values for the reduced-form model estimated over the 1966:01 –1997:12 period. The second row lists the same statistics for a model that allows for a break in all coefficient values (and the variances) after October 1979. Notice that at all reported lag lengths, the model that allows for a break is preferred (i.e. has a lower value for the information criteria) to the model with fixed coefficients. The third row allows for a break after 1982:12, and the results are the same: structural stability of the reduced-form VAR model is rejected in either case.[15]

A similar finding exists for the joint sample period. A model that imposes the same coefficient values across both sub-periods yields a larger value of either information criteria at almost all lag lengths: the single exception is the SIC criteria with 12 lags. Again, the evidence supports a different set of estimated parameters of the reduced-form model pre- and post-Volcker, even after allowing for changes in the distribution of the shocks themselves.

The information criteria also are consistently lower for samples that exclude the Volcker disinflation period than for those that include it. Although these statistics are not directly comparable, this finding suggests that exclusion of this period is warranted. Such intuition is consistent with the results of Bernanke and Mihov (1998), who illustrate that a model with nonborrowed reserves as the policy instrument provides a better fit of the data during 1979:10 to 1982:10 than does one with a fed

[15]Note that the results in table 2.1 are not directly comparable across different lag lengths: the initial values of the VAR are drawn from within the stated sample period, thereby making the number of observations used for estimation vary inversely with the number of lags.

funds instrument (their preferred specification overall).

But there also may have been a change in the preferences of the monetary authorities following the Volcker disinflation. Such a change would be reflected in different weights given to various macroeconomic aggregates within the reaction function. Several researchers, including Taylor (1997), Clarida et al. (1998) and Judd and Rudebusch (1998), find the coefficients of an estimated Taylor rule (in particular the relative emphasis on the inflation response) change at October 1979. In the VAR literature, Brunner (1994), Strongin (1995), Bernanke and Mihov (1998) and Bagliano and Favero (1998) also have found a statistically significant difference in the reaction function before and after this date, although Christiano et al. (1998) report that they do not. Disentangling this issue is the focus of the next several subsections. Not only must such changes have occurred, but they must be *quantitatively* sizable for monetary policy to be the main source of the different dynamics over time.

2.3.2 Equation-by-Equation Stability Tests

Table 2.1 points to possible instability in the VAR model, but cannot determine whether the reaction function is the primary culprit or, as suggested in section 2.1, the non-policy portion of the economy exhibited significant changes itself, independent of policy. In table 2.2 I test the individual equations of the estimated VAR for parameter stability. (Recall that the identifying restrictions imposed by this model do not suggest structural interpretations for any equations of the model other than the policy rule.) Based on the results of table 2.1, I investigate whether there exists evidence of instability for the joint sample, 1966:01 – 1979:09 and 1983:01 – 1997:12.

Part (a) of table 2.2 reports the p-values for Chow tests of parameter stability (a common test for sub-sample instability) for VAR models estimated with different

Table 2.2: Tests for Instability of Reduced-Form Model, 1966:01 – 1979:09 / 1983:01 – 1997:12

a. Chow Tests for Parameter Stability (Constant innovation variance)

Equation	*4 lags*	*6 lags*	*8 lags*	*12 lags*
IP	0.4034	0.1336	0.1056	0.0255
CPI	0.0000	0.0001	0.0011	0.0600
PCOM	0.0001	0.0049	0.0145	0.2764
FF	0.1052	0.1043	0.0378	0.3382
NBR	0.0007	0.0051	0.0067	0.0019
TR	0.0000	0.0000	0.0000	0.0009

Notes: Part (a) reports p-values for the null hypothesis of equal coefficient vectors across the two sub-sample periods. Test statistic has an F distribution.

b. Goldfeld–Quandt Tests for Heteroskedasticity of Innovation Variance

Equation	*4 lags*	*6 lags*	*8 lags*	*12 lags*
IP	0.0000	0.0000	0.0000	0.0002
CPI	0.0007	0.0011	0.0033	0.0066
PCOM	0.0001	0.0003	0.0003	0.0042
FF	0.0000	0.0000	0.0000	0.0000
NBR	0.0000	0.0001	0.0006	0.0466
TR	0.3368	0.3633	0.4830	0.5184

Notes: Part (b) reports p-values for the null hypothesis of equal residual variances across the two sub-sample periods. Test statistic has an F distribution.

c. Wald Tests for Parameter Stability (Varying innovation variance)

Equation	*4 lags*	*6 lags*	*8 lags*	*12 lags*
IP	0.2660	0.0686	0.0382	0.0088
CPI	0.0000	0.0000	0.0001	0.0179
PCOM	0.0000	0.0019	0.0077	0.2782
FF	0.0612	0.0718	0.0208	0.3414
NBR	0.0000	0.0003	0.0015	0.0002
TR	0.0000	0.0000	0.0000	0.0000

Notes: Part (c) reports p-values for the null hypothesis of equal coefficent vectors across the two sub-sample periods. Test statistic has a χ^2 distribution.

lag lengths. A low p-value implies a rejection of the null hypothesis of identical coefficients for the 1966:01 – 1979:09 and 1983:01 – 1997:12 periods. Surprisingly, at conventional significance levels, the null hypothesis of no change *cannot* be rejected for the federal funds rate equation (FF). This result would appear to contradict some of the research cited above.

But the Chow test is only valid if the variance of the residuals is constant across the sub-periods in question. Part (b) of table 2.2 reports the p-values of the Goldfeld-Quandt test of heteroskedasticity applied to the residuals of each equation. The test statistic for equation i is the ratio of the variance estimated for the 1966:01 – 1979:09 sample to the variance estimated for the 1983:01 – 1997:12 sample, scaled by their degrees of freedom:

$$\widehat{GQ}_i = \frac{\widehat{\sigma^2_{1i}} \,/\, (T_1 - np - 1)}{\widehat{\sigma^2_{2i}} \,/\, (T_2 - np - 1)} \sim F_{T_1 - np - 1, T_2 - np - 1} \ ,$$

where $\widehat{\sigma^2_{1i}} = \widehat{v}_{1it}{}' \, \widehat{v}_{1it}$ denotes the sample variance of equation i estimated over the pre-Volcker period (with $T_1 = 161$ observations), and $\widehat{\sigma^2_{1i}} = \widehat{v}_{2it}{}' \, \widehat{v}_{2it}$ is defined analogously for the post-Volcker period ($T_2 = 176$ observations).[16] n is the total number of variables in the VAR model (equal to 6 here) and p is the number of lags in the VAR (set to 4 in table 2.2). The low p-values in part (b) of table 2.2 imply rejection of the null hypothesis of equal variances for the two sub-samples in question, save for the total reserves equation.

In light of this finding, part (c) reports the p-values for Wald tests of parameter stability. Unlike the Chow test in part (a), the Wald test is valid even in the absence of homoskedasticity. For the VAR estimated over 1966:01 – 1979:09, let $\mathcal{A}_{1i}$ be the $(np+1)$ column vector corresponding to the coefficient estimates of equation i of this

[16] That is, $\widehat{\sigma^2_{1i}}$ is the i^{th} diagonal element of $\widehat{\Sigma}_1$, the covariance matrix defined in equation (2.7) estimated over the pre-Volcker sample.

VAR. If $\mathcal{X}_1$ denotes the $n \times (np+1)$ conformable matrix of lagged observations of X for this sample, and if $\mathcal{A}_{2i}$ and $\mathcal{X}_2$ are defined analogously for the 1983:01 – 1997:12 sample, then the Wald test statistic is

$$W = (\mathcal{A}_{1i} - \mathcal{A}_{2i})' \left(\widehat{\sigma^2_{1i}} (\mathcal{X}_1' \mathcal{X}_1)^{-1} + \widehat{\sigma^2_{2i}} (\mathcal{X}_2' \mathcal{X}_2)^{-1} \right)^{-1} (\mathcal{A}_{1i} - \mathcal{A}_{2i}) \,,$$

which has an asymptotic χ^2 distribution with $np+1$ degrees of freedom. In this case, the evidence against the null of equal coefficient vectors is still mediocre: at a conventional 5% level of significance, one could claim support for a structural break in the funds rate equation only for the model estimated with 8 lags.

On the other hand, the evidence against stability of most of the other equations of the model is much stronger, particular for the two price level series and for both measures of bank reserves. appear to be significant cross-regime changes in most of of the model. Consequently, it would be premature to attribute the difference in economic performance between the two regimes solely to changes in the form of monetary policy. Such a conclusion would require barely statistically discernible changes in the monetary reaction function to produce strongly significant instability in several other equations in the model. Implausible although not impossible, the quantitative significance of a change in the policy rule is investigated further in section 2.4.

These statistical tests are but one piece of the investigation of a regime change, and are not definitive. Next I explore methods to quantify the magnitude of the variation in monetary policy.

2.3.3 Exogenous Policy Shocks

Figure 2.1 plots the estimated funds rate residuals (in percentage points) from a semi-structural VAR (as described in section 2.2) estimated over the full sample period of 1966:01 to 1997:12. The most striking aspect of this figure is the extreme values between October 1979 and December 1982. This finding reflects the change in Fed operating procedures during the Volcker disinflation: as the Fed switched its instrument from the fed funds rate to nonborrowed reserves, the volatility of the funds rate (and its residual) increased dramatically. In light of the discussion at the beginning of this section, the funds rate residuals cannot be interpreted as policy innovations from 1979:10 – 1982:12. The apparent increase in volatility of the measured policy disturbances during this period is a consequence of misspecifying the policy instrument in the VAR model. Below I investigate the sensitivity of the results to this misspecification.

According to table 2.3, including the "Volcker disinflation" period results in estimated policy shocks that are nearly twice as large as those estimated for a sample that excludes the 1979:10 to 1982:12 period. (Compare the first and second rows of table 2.3.) The magnitude of this differential is robust to estimating the model with various lag lengths.[17] Notice that removing the "Volcker disinflation" period also generates a distribution of shocks that appears more symmetric and less leptokurtic (by an order of magnitude).

To get a fuller picture of how sensitive the moments of the estimated policy shocks are to the choice of sample period, figure 2.2 plots the empirical distribution of the policy shocks for the 1966:01 – 1979:09 period from models estimated over three samples: the 1966 to 1979 sub-sample itself (top panel), the joint 1966:01 – 1979:09 and 1983:01 – 1997:12 sample (middle panel) and the full sample (bottom panel).

[17]The results are reported with different lag lengths for several reasons: first, to facilitate comparison across other studies in this literature, which do not consistently use the same lag length (although one year is most commonly reported); second, the optimal lag length varies somewhat depending on the sample period in question (between 4 and 8 months for the samples reported herein); and third, the table illustrates that in general, most of these results are not particularly sensitive to the lag lengths considered.

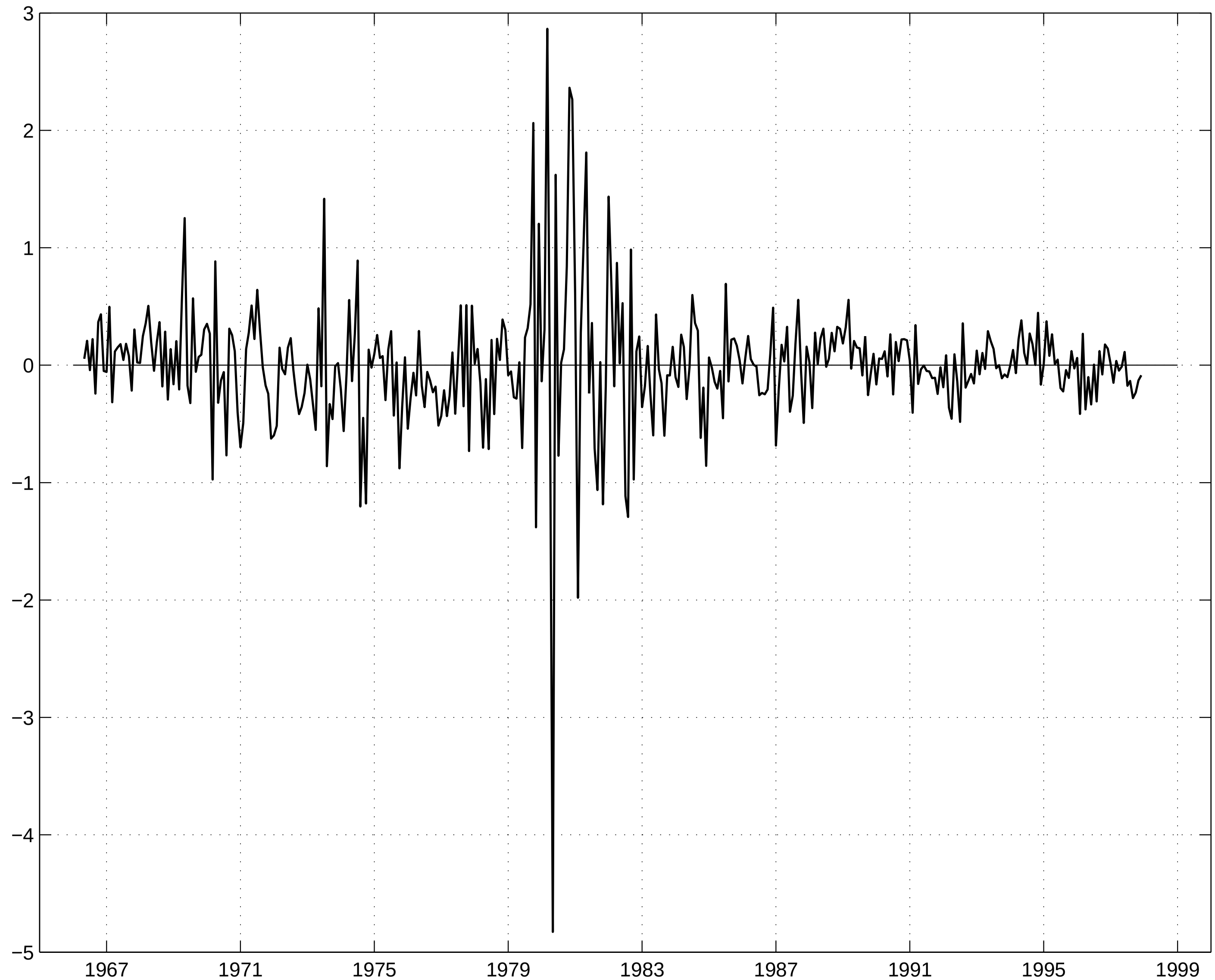


Figure 2.1: Estimated Monetary Policy Shocks, 1966:01 – 1997:12 [4 lags]

Table 2.3: Moments of Estimated Monetary Policy Shocks

	4 lags			*6 lags*		
Sample Period	Standard Deviation	Skewness	Kurtosis	Standard Deviation	Skewness	Kurtosis
1966:01 – 97:12	0.5431	−0.9554	20.4518	0.5189	−0.6871	16.7466
		(0.0000)	(0.0000)		(0.0000)	(0.0000)
1966:01 – 79:09, 1983:01 – 97:12	0.2837	0.0353	1.9834	0.2676	0.0350	1.6819
		(0.7925)	(0.0000)		(0.7954)	(0.0000)
1966:01 – 79:09	0.3213	0.3608	0.7186	0.2962	0.0274	0.5504
		(0.0641)	(0.0686)		(0.8891)	(0.1658)
1966:01 – 82:12	0.6874	−0.6483	9.8090	0.6372	−0.2697	7.0170
		(0.0002)	(0.0000)		(0.1241)	(0.0000)
1983:01 – 97:12	0.2044	−0.3922	9.4390	0.1822	−0.3033	1.5815
		(0.0000)	(0.0000)		(0.1054)	(0.0000)
1979:10 – 97:12	0.5537	−0.9276	6.4518	0.4821	−0.8467	9.5538
		(0.0003)	(0.0000)		(0.0000)	(0.0000)

	8 lags			*12 lags*		
Sample Period	Standard Deviation	Skewness	Kurtosis	Standard Deviation	Skewness	Kurtosis
1966:01 – 97:12	0.4927	−0.8611	16.0767	0.4586	−0.8796	16.4345
		(0.0000)	(0.0000)		(0.0000)	(0.0000)
1966:01 – 79:09, 1983:01 – 97:12	0.2592	0.1153	2.1197	0.2444	0.2979	2.3354
		(0.3952)	(0.0000)		(0.0301)	(0.0000)
1966:01 – 79:09	0.2694	0.0094	0.9351	0.2362	0.1659	1.0092
		(0.9618)	(0.0194)		(0.4070)	(0.0128)
1966:01 – 82:12	0.5919	−0.4913	7.0939	0.5103	−0.0703	6.2970
		(0.0053)	(0.0000)		(0.6931)	(0.0000)
1983:01 – 97:12	0.1719	−0.1516	0.9982	0.1522	−0.0345	0.0732
		(0.4210)	(0.0088)		(0.8563)	(0.8496)
1979:10 – 97:12	0.3350	0.1632	2.9003	0.2635	0.0914	0.9896
		(0.3367)	(0.0000)		(0.5939)	(0.0043)

Notes: Kurtosis is standardized so that a normal distribution has kurtosis of zero. Numbers in parentheses represent p-values under the null hypothesis that skewness or kurtosis is equal to zero.

Consistent with the results in table 2.3, the variance of the shocks increases as one looks down the figure. Figure 2.3 illustrates analogous relationships for the 1983:01 – 1997:12 period.

One interpretation of these findings is that the shocks estimated over the longer samples — especially any sample that includes 1979:10 to 1982:12 — confound purely exogenous innovations to policy with a change in the policy rule. If the monetary policy regime had changed during the period being examined, estimating a "one-size-fits-all" policy rule would be inappropriate. Failure to explicitly account for a change in regime coming about through a change in the reaction function therefore would yield mismeasured policy shocks.

Separating the sub-samples, the standard deviation of the policy shocks for 1966 to 1979 is nearly 1.5 times that of 1983 to 1997. (The earlier sample also appears slightly right-skewed, while the later one is left-skewed.) Adding the "Volcker disinflation" period to either the end of the 1966 – 1979 period (as in the fourth row) or to the beginning of the 1983 – 1997 period (as in the sixth row) yields standard deviations that generally exceed twice that of either sub-sample without this 1979:10 – 1982:12 period.

Figure 2.1 also indicates a greater variance of policy shocks during the 1966 to 1979 period as compared to the 1983 to 1997 one, consistent with the Goldfeld-Quandt test results reported in table 2.2. Table 2.3 quantifies the moments of the monetary policy shocks for models estimated across several (overlapping) time periods and with different lag lengths. These results confirm differences in the behavior of the exogenous policy component between the two sub-sample periods.

These differences deserve further comment. Regardless of the interpretation of the exogenous policy shocks (whether they be willful actions by the Fed or "mistaken"

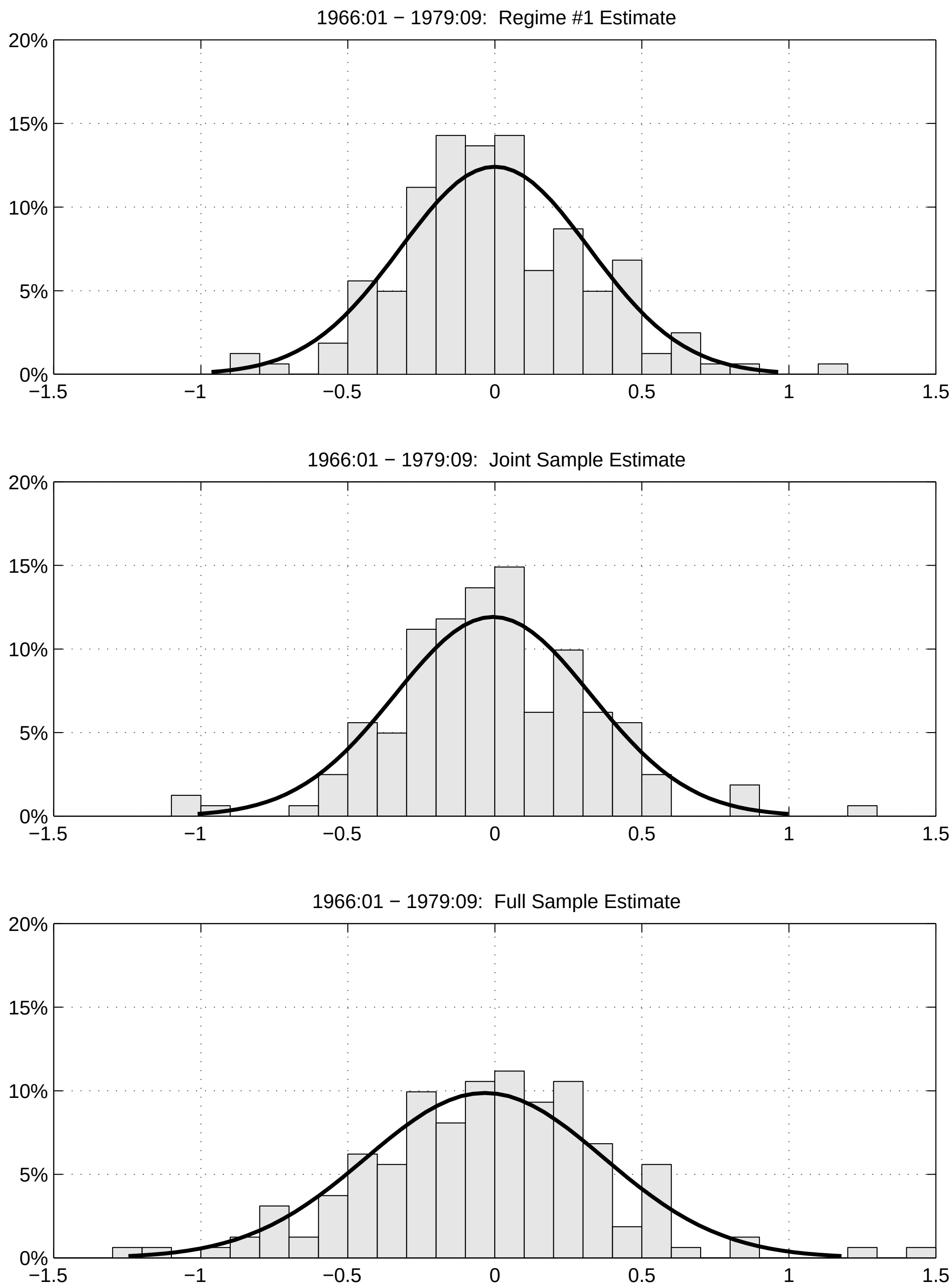


Figure 2.2: Distribution of Estimated Policy Shocks, 1966:01 – 1979:09 [4 lags]

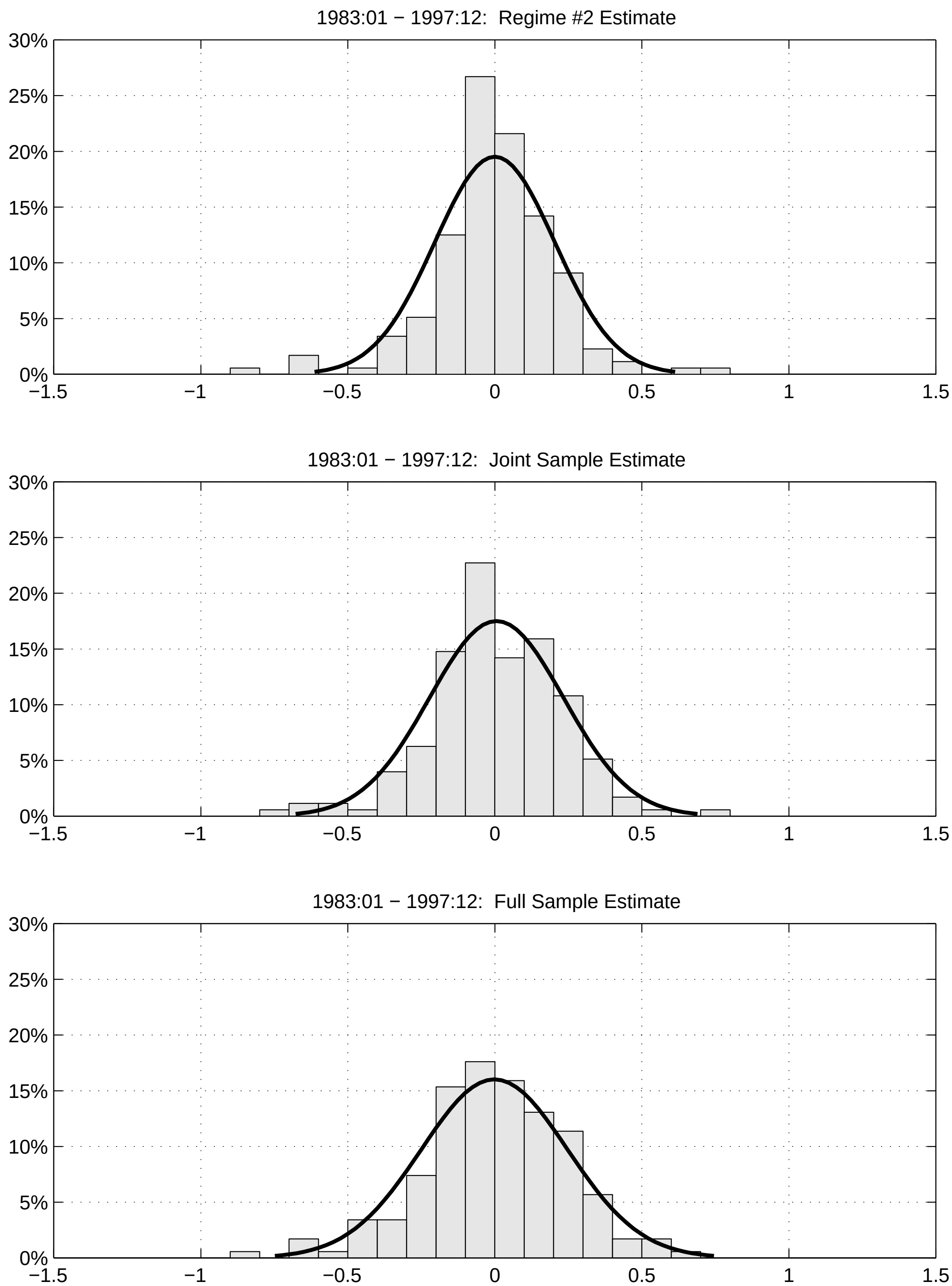


Figure 2.3: Distribution of Estimated Policy Shocks, 1983:01 – 1997:12 [4 lags]

deviations from the policy rule), one might conjecture that these shocks represent the primary source of economic fluctuations over time. In this view, a “regime change” could take the form of a change in the process that generates the policy shocks for a given reaction function and given dynamic relations among the activity variables. This is the preferred interpretation of Christiano et al. (1998), who note the unequal variances of their estimated policy shocks between 1965:01 – 1979:09 and 1979:10 – 1997:12 sub-samples. They further posit that any estimated differences in dynamic behavior across sample periods can be traced to this variance differential, rather than a change in the reaction function.[18]

But what explains *why* the process for the exogenous innovations to policy differs across sub-samples? One possible interpretation is that the policy rule for the 1966:01 – 1979:09 period was “inappropriate,” as some of the Taylor rules literature suggests. This could imply that substantial willful (yet unforecastable) corrections to monetary policy were necessary, or that significant policy “mistakes” were the result of such an inappropriate policy rule. In other words, the basic source of instability would be the policy rule, and the observed properties of the exogenous component of policy merely would be a consequence of it. This possibility is explored in the next subsection.

Another possibility is that the economy was much more difficult to control in the earlier sub-sample period. What might have been accomplished with an exogenous 25-basis point shock after 1982 may have taken a much larger shock, perhaps 50-basis points, prior to October 1979. Greater uncertainty about the economic environment also might have led to the systematic part of policy being a smaller proportion of the

[18]I am not aware of any other work in the VAR literature (or the Taylor rules literature) that has analyzed the data generating process for the exogenous monetary shocks.

total variation in the instrument.[19] Alternatively, larger policy shocks themselves may be responses to uncertain events upon which the Fed (and private agents) could not condition policy (i.e. they could not be incorporated into the reaction function). In each of these cases, greater instability in the non-policy portion of the model could generate the observed behavior of the monetary policy shocks.

While the policy rule may have been different under either of these circumstances — indeed, the optimal rule is unlikely to be invariant to changes in the dynamics of the activity variables — it need not have been "inappropriate." A different reaction function, in conjunction with greater instability inherent to the activity sector of the model, could produce a larger variance for the policy shocks in the early sample period.

Thus, rather than being the source of instability in the model, the variance differential simply could be a *symptom* of instability residing elsewhere: either the inappropriateness of the reaction function, or the inherent instability of the economy itself. The interpretation of Christiano et al. (1998) would hold only in the absence of any variation in the policy rule or in the dynamics of the broader economy. The properties of these two parts of the model are subsequently examined in section 2.4.

2.3.4 Endogenous Policy Rule

The VAR models studied herein allow the policy rule to change by varying the weights on each of the variables in the Fed's information set. This approach is not especially restrictive: the VAR contains a large number of lags that allow for fairly complicated dynamics; simpler models are nested by restricting certain lag coefficients to zero. On the other hand, because estimated VAR policy rules contain so many lagged observations it is cumbersome to analyze their form. I propose to

[19]Brainard (1967) discusses conditions under which this situation may occur.

measure the total net response of the monetary policy instrument (the federal funds rate) to each observed variable in the model by summing the coefficients of the lag polynomial for the variable in question.

Normalizing the diagonal elements of the leading matrix of the structural vector autoregressive form in equation (2.3) to unity (see footnote 6) yields a convenient expression for the structural form of the model:

$$X_t = \Phi_0 X_t + \Phi_1 X_{t-1} + \cdots + \Phi_p X_{t-p} + \varepsilon_t \,.$$

Here the diagonal elements of Φ_0 are zeroes. Let m^* be the index of the policy instrument in the vector X_t; with the activity block ordered before the policy block as in section 2.2, $(n - m + 1) \leq m^* \leq n$. Equation m^* then corresponds to the reaction function, which can be written as

$$X_{m^*t} = FF_t = \sum_{j=1}^{n} \sum_{k=0}^{p} \phi_{m^*jk} X_{jt-k} + \mu_t \,,$$

where FF_t is the funds rate, ϕ_{m^*jk} is the (m^*, j) element of Φ_k, and μ_t is the monetary policy shock (the m^*-th element of ε_t). The net response of the policy instrument (the Fed funds rate) to variable j can be estimated as

$$\text{sum}[\phi_j] = \sum_{k=0}^{p} \widehat{\phi}_{m^*jk} \,, \quad \forall\, j = 1, \ldots, n \,.$$

Table 2.4 reports these summations, along with their corresponding standard errors. In accordance with the identifying assumptions set forth in section 2.2, these statistics include the contemporaneous values of the activity variables in the reaction function. The sharp distinctions across sample periods suggest significant shifts in the relative importance the Fed assigned to different variables over time, which itself is evidence of a change in the monetary policy regime.[20] Several of the differences

[20]There are also some differences across the lag lengths used in estimation. The AIC selected 4 monthly lags as optimal for the sub-sample VARs, whereas 12 lags are common in the VAR literature.

between the pre-Volcker and post-Volcker periods are especially pronounced.

Specifically, the endogenous component of monetary policy is nearly twice as sensitive to output during the earlier sub-sample: a one percent increase in industrial production engenders an eventual response in the funds rate of 225 basis points for the 1966:01 – 1979:09 period, versus a little more than a percentage point (107 basis points) increase for the 1983:01 – 1997:12 period. (These results are estimated with 4 lags; the results with 12 lags are not statistically significant for either sub-sample period.) Notice that the constrained sum estimated over the joint sample (second row of table 2.4) implies only a 65 basis point response to output rising one percent, and is not estimated very precisely despite the much larger number of observations.

While those estimated coefficient sums on output appear plausible, consider the response of the policy instrument to the price level: it shifts from a positive value prior to October 1979 — (nominal) interest rates rise with higher prices — to a *negative* value for the 1983:01 – 1997:12 sample — higher prices actually lead the Fed to *reduce* (nominal) interest rates. In fact, the estimated response to a one percent increase in the price level is to reduce the fed funds rate by nearly 2 percentage points. This would hardly seem to be a stabilizing policy prescription! (Additionally, the cumulative response to the price level is not statistically discernible in the pre-Volcker period, while the cumulative responses to commodity price changes are never statistically significant for any sub-sample.)

These troubling results have not been brought to light in previous VAR research, as the endogenous policy rule is routinely ignored. Indeed, they appear to directly contradict sub-sample comparisons reported in the Taylor rules literature. (See Taylor, 1997; Clarida et al., 1998; Judd and Rudebusch, 1998.) While it is difficult to compare across the models employed in these two literatures — the Taylor rule spec-

Table 2.4: Coefficients of Estimated Monetary Reaction Functions: Summations of Lag Polynominals

	Regressors: 4 lags				*Regressors: 12 lags*			
Sample Period	IP	CPI	PCOM	FF	IP	CPI	PCOM	FF
1966:01 – 1997:12	1.1754	−0.6853	0.4309	0.9836	2.7929	−0.8163	−0.4966	1.0242
	[0.7719]	[0.5709]	[0.3186]	[0.0254]	[1.0781]	[0.7606]	[0.3843]	[0.0304]
1966:01 – 1979:09,	0.6554	−0.2654	0.1912	0.9768	0.8869	0.1820	−0.1974	0.9796
1983:01 – 1997:12	[0.4395]	[0.3421]	[0.1841]	[0.0173]	[0.6886]	[0.5251]	[0.2413]	[0.0227]
1966:01 – 1979:09	2.2514	0.1152	0.3311	0.9111	2.2045	4.1564	−0.1634	0.6909
	[1.0530]	[0.8234]	[0.3100]	[0.0478]	[2.2018]	[1.5685]	[0.4603]	[0.9764]
1966:01 – 1982:12	1.9021	1.5778	0.8452	0.9436	4.3138	2.5542	1.1039	1.0321
	[1.8165]	[1.2463]	[0.5600]	[0.0505]	[3.3266]	[2.0564]	[0.7622]	[0.0768]
1983:01 – 1997:12	1.0732	−1.8820	0.6641	0.9401	0.2845	−1.7836	1.2166	0.9271
	[0.6834]	[0.6205]	[0.6132]	[0.0309]	[1.2742]	[0.9623]	[1.0578]	[0.0520]
1979:10 – 1997:12	0.6500	−3.5579	2.6692	0.9510	3.3713	−4.0387	1.6392	0.9270
	[1.4371]	[1.1566]	[1.0853]	[0.0398]	[1.1806]	[1.1548]	[1.0557]	[0.0455]

Notes: Numbers in square brackets represent the standard error for the sum of the estimated coefficients of the lag polynomial, corresponding to the regressor listed at the top of the column.

ifies a response to the inflation rate, not the price level, and uses the output gap instead of the level of output — the implied reduction of interest rates in the face of rising prices is inconsistent with the inflation response of the Fed estimated in most Taylor rules. It also runs strongly counter to most economists' intuition.

One constant across all sub-samples is the high degree of persistence in the funds rate equation, even after conditioning on several lags of the other variables in the model. The cumulative sum of lagged funds rate coefficients is very large: nearly unity in several cases. While it seems implausible that the nominal interest rate is integrated, this finding highlights a weakness in basic specifications of the Taylor rule: the lack of explicit persistence terms in the funds rate equation.[21]

Summarizing the results of this section, two facets of the VAR approach stand out. First, most of the models estimated feature instability in both the residuals and the coefficients — arguably more so for the non-policy equations. The model appears to behave differently before and after the "Volcker disinflation" experiment with money targeting (1979:10 – 1982:12). Second, while the tests of parameter instability generated only weak evidence against stability of the monetary policy reaction function, the estimated sensitivity of this systematic component to particular activity variables differed markedly over time. Moreover, the variance of the exogenous monetary policy shocks changed measurably across the two sub-samples in question.

2.4 Quantifying the Effects of Regime Changes

The only way to ascertain how important variation in the formation of monetary policy has been for the U.S. economy — whether that variation is a change in the

[21] More recent specifications, such as Clarida et al. (1998), include one lag of the funds rate in their specification of the Taylor rule.

distribution of the policy shocks or a change in the parametrization of the policy reaction function — is to attempt to quantify the role played by such changes. In this section I investigate the quantitative significance of the policy shocks and the reaction function. First I compare the historical forecast error decomposition of the funds rate across different time periods, then I measure the contribution of the policy shocks to the historical dynamics of all the variables in the model.

2.4.1 Comparison of Historical Decompositions

Recall that the estimated model can be rewritten in a vector moving average form, as in equation (2.1). That is, each series can be expressed as an (infinite) lag process in each of the estimated innovations. Since an infinite history of data is not available, the figures below use a rolling window of two years ($K = 24$ months) to compute the contribution of the policy shock to the historical forecast error decompositions.[22]

At each date in the sample (after the initial K start-up values), the expression for X_{it} can be separated into the contribution of the policy shocks μ_t and the other (non-structural) innovations ε_{jt}, $j \neq m^*$ (plus a series-specific truncation error ξ_{it}):

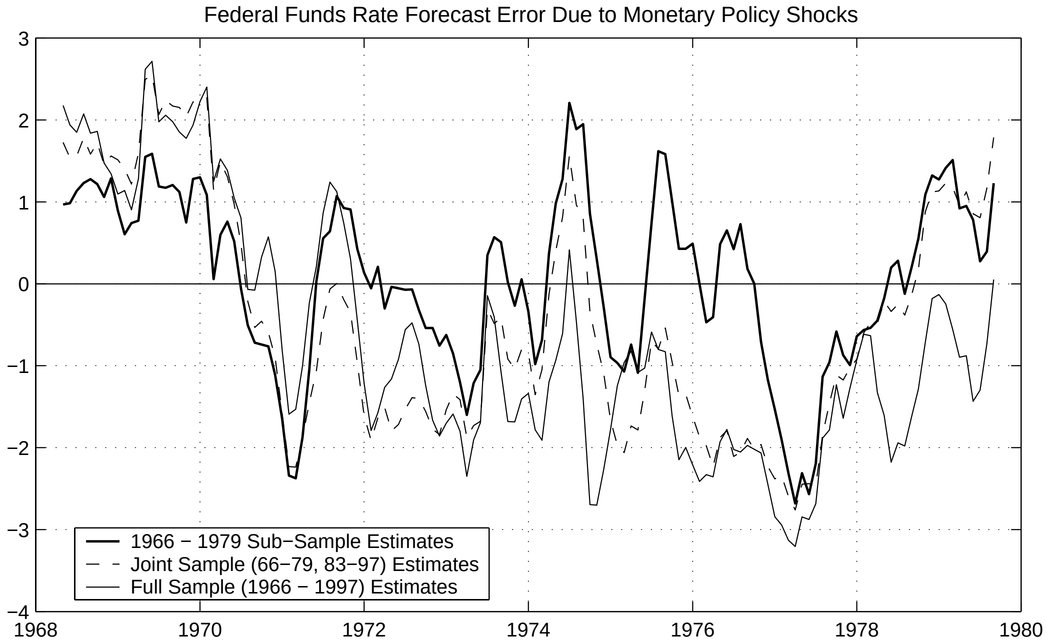
$$X_{it} = \sum_{j \neq m^*}^{n} \sum_{k=0}^{K} \widehat{\theta}_{ijk}\, \widehat{\varepsilon}_{jt-k} + \sum_{k=0}^{K} \widehat{\theta}_{im^*k}\, \mu_{t-k} + \xi_{it}$$

The upper panel of figure 2.4 illustrates the policy component of the federal funds rate forecast error,

$$\sum_{k=0}^{K} \widehat{\theta}_{m^*m^*k}\, \mu_{t-k} \; ,$$

estimated over the three sample periods in question. The thick solid line represents the forecast error for a VAR that is estimated over the 1966:01 – 1979:09 period only. The thin solid line gives the forecast error for a VAR estimated over the full sample,

[22]For each of the figures, the optimal lag length of four months was used in estimation. The results for rolling windows of width $K = 24$ and $K = 48$ (not reported here) are quite similar. For sufficiently large K, any truncation error will be minimal.



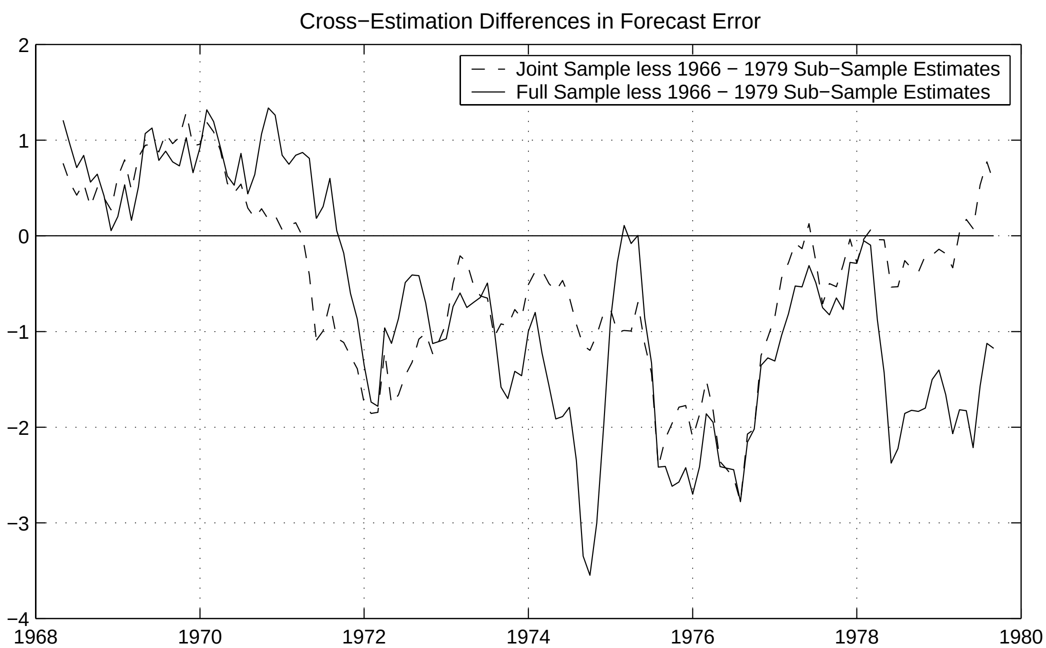


Figure 2.4: Policy Component of Federal Funds Rate Forecast Error, 1966:01 – 1979:09 (2-year forecast horizon, 4 lags)

through 1997:12, while the dashed line reflects the forecast error for a VAR estimated jointly over the 1966:01 – 1979:09 and 1983:01 – 1997:12 periods (that is, excluding the Volcker disinflation). Positive values indicate contractionary shocks to monetary policy — the actual values of the funds rate were higher than the estimated policy rule alone would indicate — while negative values imply expansionary policy shocks.

Under the null hypothesis of no change in regime, all three approaches should produce identical historical decompositions. If the Volcker disinflation was a unique policy experiment with monetary policy practiced in a similar manner before and after this experiment, then the joint sample and sub-sample results should coincide.

Notice in the lower panel of figure 2.4 that from 1968 until mid-1971, the full and joint sample estimates imply the exogenous component of policy was more contractionary (more positive) than the 1966 – 1979 sub-sample estimation does — by more than a full percentage point at times. Conversely, for the seven-year period from late 1971 through 1978, the full and joint sample estimates not only suggest an easier monetary policy than the sub-sample estimation (lower panel), but always (save a brief exception during 1974) imply the exogenous component of policy was responsible for a *lower* funds rate than the policy rule would otherwise suggest (upper panel). During this interval, the sub-sample results suggest the exogenous component of policy was leading to higher interest rates — tightening by the Fed — while the full and joint sample results imply the *opposite* exogenous policy action. The deviation between the sub-sample and the full or joint sample estimates is especially pronounced in 1976, averaging between 2 and 3 percentage points.

In light of the statistical results of section 2.3, this systematic expansionary bias is likely a symptom of a misspecified policy rule: namely, one that falsely remained fixed in estimation over a long sample period. Both the sign and magnitude of the path attributed to the exogenous component of policy are very different in the sub-sample estimation vis-à-vis the full or joint estimates. Interestingly, from the second half of 1978 until the end of this sub-sample in September 1979, the forecast error implied by the joint estimation and that implied by the sub-sample estimation correspond quite closely. However, while both of these measures imply a contractionary exogenous component of policy, the full sample estimates continue to suggest that policy was expansionary.

Figure 2.5 repeats this comparison for the 1983:01 – 1997:12 period. Here the differences across the three estimated models are more of degree than kind: all three imply the exogenous component of policy was contractionary from mid-1988 until early 1991, and again from the beginning of 1994 through the end of 1995. The main expansionary episode began in late 1991 and lasted into the first-half of 1993; exogenous policy also appears expansionary in 1996.

The largest deviations across these three measures occur at two points in time: late 1988 through early 1991, and the later months of 1996 through the end of the sample (at December 1997). In both of these cases, the difference between the sub-sample estimate and either other estimate runs between one and two percentage points — quite a sizable amount for the funds rate.

Unlike figure 2.4, figure 2.5 does not reveal a systematic bias towards easing or tightening. However, the results in both of these figures indicate a much larger forecast error for the federal funds rate — and therefore a larger implied effect of the exogenous component of policy — when the full or joint sample is used to estimate the model and its shocks, rather than when the estimation is restricted to the sub-sample period in question. And they are consistent the statistical findings in section 2.3.3, which demonstrated a larger variance for shocks that were estimated

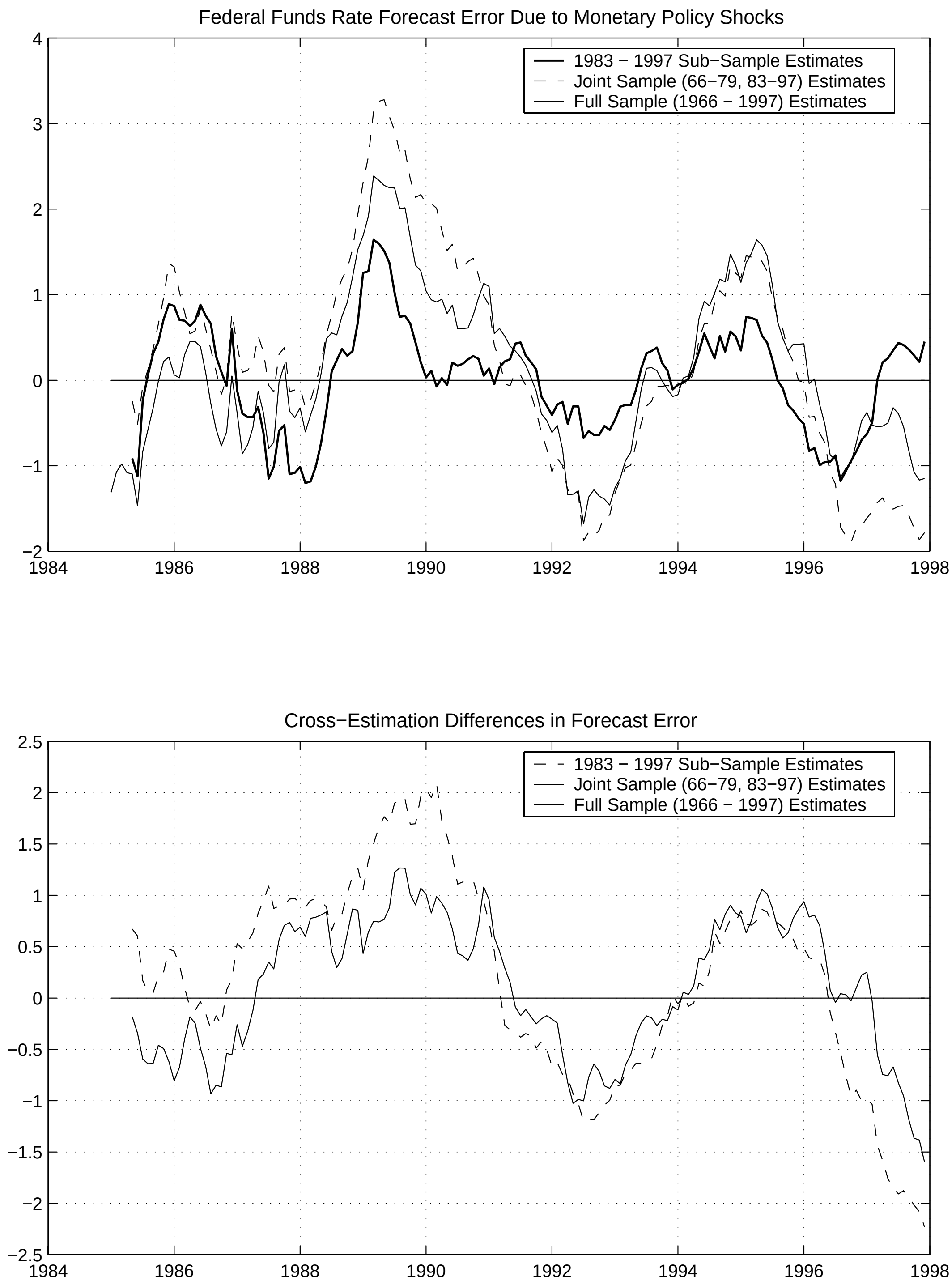


Figure 2.5: Policy Component of Federal Funds Rate Forecast Error, 1983:01 – 1997:12 (2-year forecast horizon, 4 lags)

under the assumption of no break in regime. The results of this section thus not only suggest a quantitatively discernible regime change, but also indirectly confirm that changes in the policy rule are an important source of fluctuations themselves. Estimation over long samples confounds shocks with variation in the policy rule. The reason other VAR studies may have found a significant contribution originating in the policy shocks may be due, in part, to a failure to account for this variation.

2.4.2 Historical Decompositions: Economic Significance

The estimated sub-sample forecast errors for the funds rate, graphed as the thick solid lines in the top panels of figures 2.4 and 2.5, are non-trivial in magnitude: they range between one and three percentage points at their peaks. The economic impact of these shocks is measured in figure 2.6 for the 1966:01 – 1979:09 period and figure 2.7 for the 1983:01 – 1997:12.[23] In these figures, the thick solid lines report the actual data for each series X_{it}, while the thin solid lines illustrate the path of the series once the contribution of the monetary policy shocks has been removed:

$$X_{it} - \sum_{k=0}^{K} \widehat{\theta}_{im^*k}\, \mu_{t-k}\ ,$$

where $i = 1, \ldots, n$. By removing this contribution from each series in the vector moving average representation, the role of exogenous monetary policy changes on each variable can be quantified. In a sense, these decompositions suggest what the historical path of each series would have been in the absence of any policy shocks.

Not surprisingly, the most pronounced effects of the policy shocks are revealed to be for the policy instrument itself, the federal funds rate. When the actual funds rate is greater than the funds rate less the policy shock, the exogenous component of Fed policy is contractionary. In other words, the funds rate would have been lower in

[23]Figures 2.6 and 2.7 report the (logarithmic) levels of each series except the federal funds rate, which is shown in percentage points.

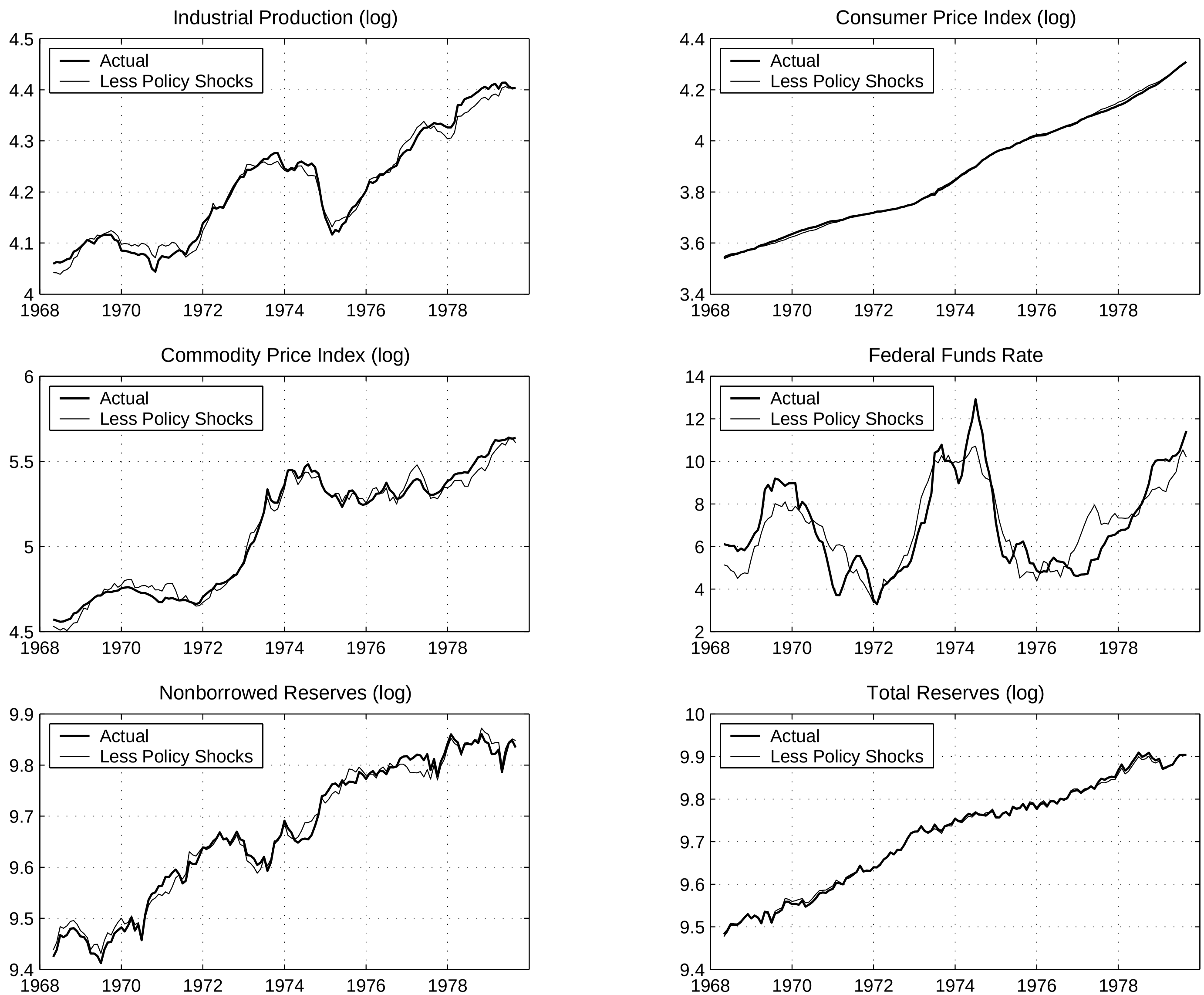


Figure 2.6: Historical Forecast Error Decompositions, 1966:01 – 1979:09: Log Levels of Variables (2-year forecast horizon, 4 lags)

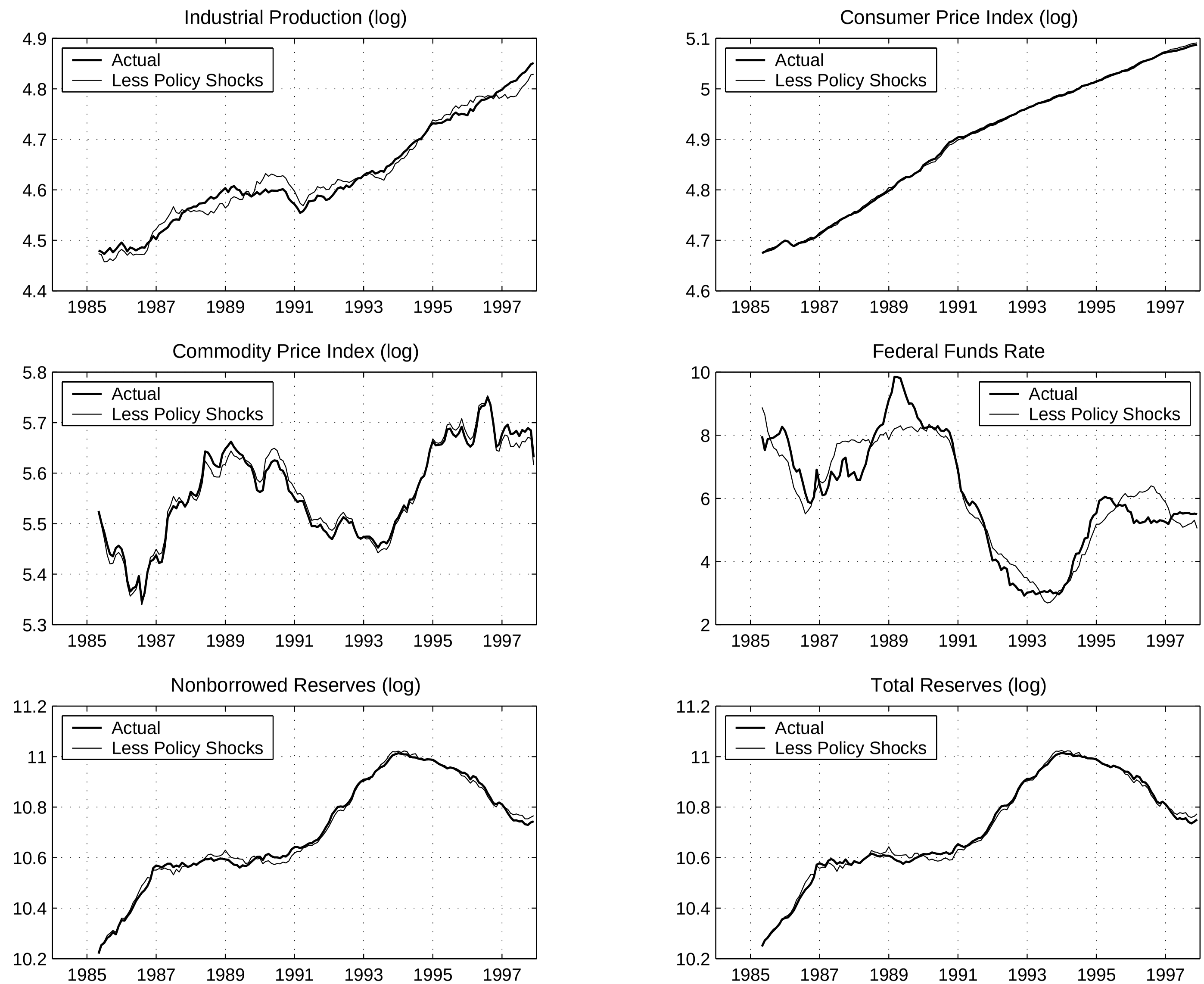


Figure 2.7: Historical Forecast Error Decompositions, 1983:01 – 1997:12: Log Levels of Variables (2-year forecast horizon, 4 lags)

the absence of these exogenous shocks.[24] This is the case, for example, from 1968:5 through 1970:6 in figure 2.6).

By this measure, a long period of expansionary policy shocks that began in July of 1970 — with a few brief periods of tightening, mainly in late 1971 — ran until the Fed started to clamp down following the first quarter of 1974. As was revealed in figure 2.4 above, the exogenous tightening from May to September 1974 averaged nearly two percentage points more than the policy rule had prescribed. Policy again eased during the first half of 1975, before an additional series of contractionary shocks occurred between 1975:7 and 1976:10. Yet again the Fed reversed course: during the second quarter of 1977, the funds rate averaged more than 250 basis points *below* the level given by the endogenous policy rule alone. In fact, cumulating the policy shocks since 1974:1 — approximately the beginning of the first oil price shock — policy was strongly expansionary from mid-1977 through mid-1979.[25]

In contrast with the pre-1979 period, exogenous shocks to the funds rate the post-1982 period are more contractionary on average. As shown in figure 2.7, exogenous policy tended to exert a contractionary influence during three main intervals: 1985:8 to 1986:12, 1988:7 to 1991:11, and 1993:6 to 1995:7. The period from July 1987 through October 1990 is rather interesting: as the thinner line indicates, the reaction function implies the funds rate should be at 8 percent, give or take about 25 basis points. However, from the beginning of this interval until April 1988, exogenous

[24]This treatment parallels figures 2.4 and 2.5. Notice that the thick line in the upper panel of figure 2.4, for example, equals the difference between the thick and thin lines of the funds rate panel in figure 2.6.

[25]By this metric, policy was contractionary in 1975 through early 1977. Cumulating the policy shocks since 1972:1 — roughly the onset of a sharp rise in overall commodity prices — yields even more evidence of an expansionary policy bias in the latter part of decade. Indeed, policy also appears quite expansionary from the start of 1973 into the third quarter of 1974 by this measure. This despite the inclusion of non-fuel commodity prices in the VAR, ostensibly because they help capture these "supply shocks." (See Sims, 1992; Balke and Emery, 1994.) Hanson (1998), however, demonstrates that the empirical support for this rationale is limited.

policy kept the funds rate more than 100 basis points below this level. By the beginning of 1989, however, this policy was sharply reversed: during the first half of that year the funds rate averaged nearly 150 points *higher* than the reaction function prescribed. However, the amplitude and frequency of swings in the contribution of the exogenous policy shocks is greatly attenuated here relative to the first sub-sample period.

Attempting to identify the effects of the exogenous policy shocks upon the rest of the variables in figures 2.6 and 2.7 is somewhat difficult. Macroeconomic intuition suggests that as the Fed increases interest rates, all of the activity variables in the model — output and the price levels — should fall. Thus, the portion of the forecast error of, say, output that is attributable to monetary policy should be negatively correlated with the forecast error for the funds rate (plotted in figure 2.4). However, the behavior shown in figure 2.6, for example, indicates *higher* forecast errors on average for the activity variables when exogenous policy is tight. Only three to five months (on average) after a contractionary policy episode do the output and commodity price variables decline. Figure 2.7 reveals a similar "phase shift" during the 1983:01 – 1997:12 sample. As it turns out, these responses are consistent with the policy shocks affecting the activity variables with "long and variable lags." Additional evidence for this attribute of the policy shocks is revealed in the impulse responses, discussed in the next subsection. (See figure 2.8.)

These issues notwithstanding, figures 2.6 and 2.7 reveal that a relatively small portion of the total variation of the other variables in the VAR model can be attributed to exogenous monetary policy shocks. To the extent that there is any effect, it appears somewhat larger in the 1966:01 – 1979:09 period. Thus, despite sometimes sizable forecast errors of the funds rate itself, exogenous monetary does not appear to have a quantitative large effect upon the broader economy.

2.4.3 Alternative Policy Rules

I now turn my attention to the endogenous component of monetary policy, the policy reaction function. Recall that the stability tests of table 2.2 provided inconclusive evidence of a change in the reaction function, the summations of the lag polynomials of the reaction function suggested more substantial cross-sample differences. (See table 2.4.) Figure 2.8 compares, across three different sample periods, the dynamic responses of several variables to a contractionary monetary policy shock.[26] In each panel, the solid line represents the point estimates of the impulse response function, while the dotted and dashed lines represent the 68% and 95% bootstrapped error bands, respectively.[27]

A few features appear common across all three estimation periods. First, there is some evidence of an "output puzzle" and a "price puzzle" in these results; only the perplexing rise in the CPI price level persists for any significant length of time following a monetary contraction.[28] Second, the initial 25 basis point increase in the funds rate is followed by further increases for several additional months, returning to its pre-shock value after roughly a year. Finally, the reduction of output in response to a policy shock appears stronger (and more statistically significant) than the sluggish responses of either price index. The joint sample estimates generate the greatest output response, falling by more than one-half of a percent at its nadir (about two years after the initial shock) and statistically remaining below its initial level for an additional year or two.

[26]To facilitate comparison, responses to a hypothetical 25 basis point increase in the funds rate are plotted in each column. However, as indicated in figure 2.1 and table 2.3, the size of a "typical" monetary policy shock varied substantially among these sample periods.

[27]See Kilian (1998) for a description of the bootstrap procedure.

[28]Hanson (1998) investigates this phenomenon in the VAR literature.

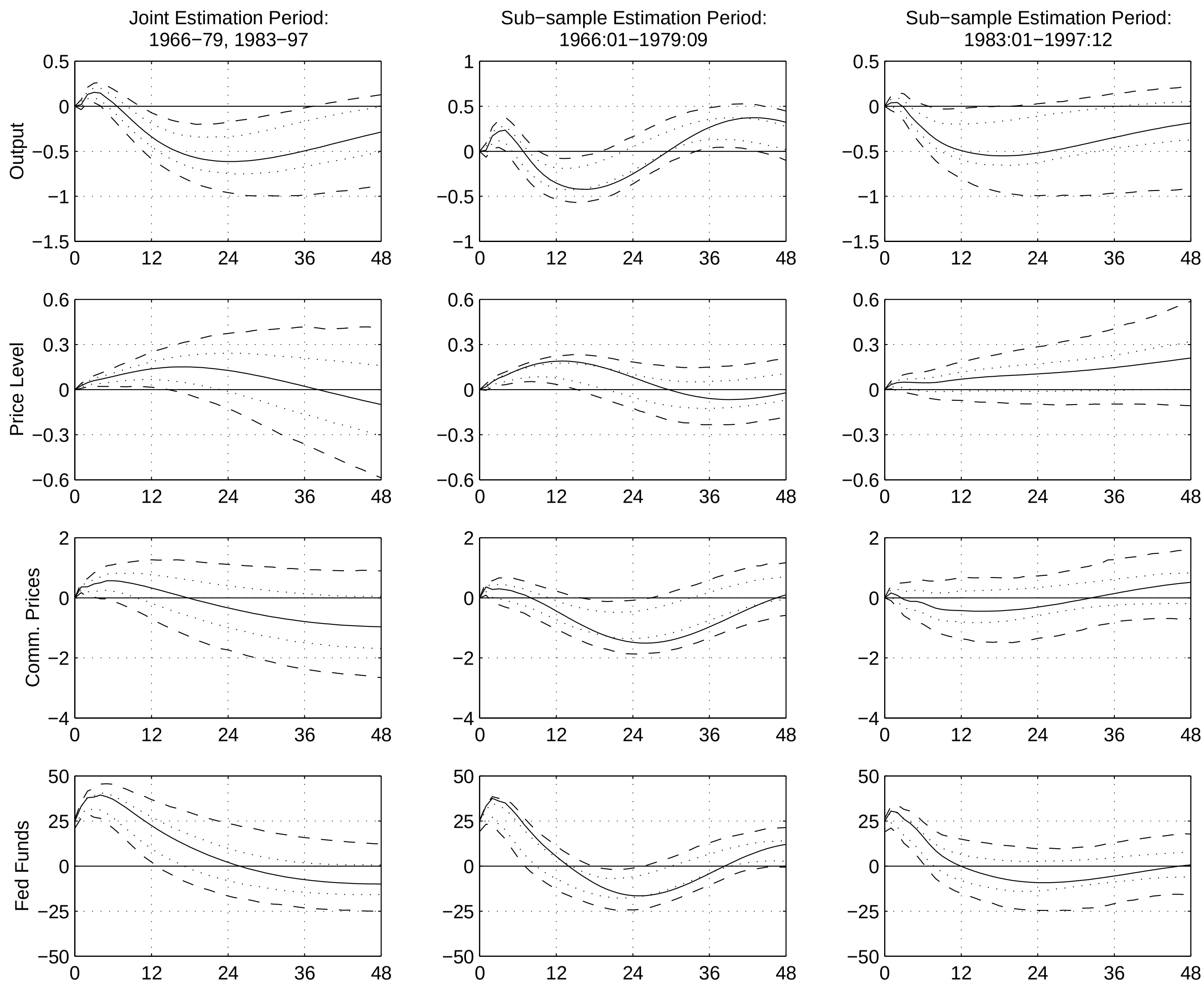


Figure 2.8: Comparison of Estimated Impulse Response Functions: Responses to a 25-basis point contractionary monetary shock

There are also a few important distinctions among the responses plotted across the rows of figure 2.8. During the 1966:01 – 1979:09 period, the increase in the fed funds rate is quickly reversed: within 18 to 24 months, there is a statistically discernible *reduction* in the funds rate of nearly the same magnitude as the initial exogenous shock. Such a rapid "stop-go" process for monetary policy is not evident during the joint or post-Volcker sample periods.[29] The responses of the activity variables reflect this path for the funds rate, with output in particular dropping a statistically significant amount within the first year, only to rise significantly around three years after the initial contractionary impulse.

On the other hand, a contractionary monetary policy shock has little statistically discernible effect on any of the activity variables during the 1983:01 – 1997:12 period. The funds rate is much less persistent than in the joint sample. Although the point estimates for the output response are similar to the joint sample results, the error bands are much wider — perhaps a reflection of the smaller amplitude of the exogenous shocks during this period, resulting in less precise estimates of the impulse responses. The responses of both price levels also are not statistically different from zero, although the point estimate for each series becomes positive within the four years plotted.

Figure 2.8 illustrates that the responses of the variables to a monetary policy shock do differ across the sample periods in question. What cannot be determined from this figure is the source of the different dynamics: it could be the policy rule, the underlying dynamics of the non-policy portion of the economy, or both. To quantify the role played by the reaction function, as well as to further test the hypothesis of a change in regime, I propose a series of counterfactual experiments.

[29] For an informative discussion of the policy environment during the latter years in this period, see De Long (1997).

In each, some portion of the joint sample VAR model is replaced with a corresponding portion of either of the sub-sample estimation periods. If the dynamics implied by the counterfactual monetary regime differed measurably from those generated by the actual regime, this would be evidence against the view of no change in regime. Further, the magnitude of the differences would help indicate the importance of the reaction function specification in generating economic fluctuations.

The first two experiments replace the estimated reaction function from the joint sample with that of the 1966:01 – 1979:09 and 1983:01 – 1997:12 sub-samples, respectively. Under the hypothesis of no change in regime, the reaction function estimated prior to 1979:10 should be economically indistinguishable from that estimated after 1982:12. Figure 2.9 compares the impulse responses of each counterfactual to those estimated for the joint sample. The left-hand column replicates the impulse responses for the joint sample model (also the left-most column of figure 2.8). The middle and right columns report the counterfactual responses when the equation of the monetary reaction function (the structural fed funds equation) is replaced by the monetary reaction function estimated for the pre-Volcker or post-Volcker period, respectively. Under the null of no variation in the policy rule, neither substitution should significantly change the dynamics of the impulse responses.

Comparing across the rows of figure 2.9, this conjecture largely appears to be supported. The counterfactual results almost always lie within the error bands of the joint model.[30] Notice the quantitative distinctions of the sub-sample estimates in figure 2.8 are still present for the corresponding counterfactuals: relative to the joint sample results, the funds rate is less persistent, the positive price response is

[30]The only exception to this statement is the output response under the 1966:01 – 1979:09 policy rule at the very end of the plotted response (roughly four years out). For comparison, the error bands for the actual estimated model are replicated in the center and right-hand columns.

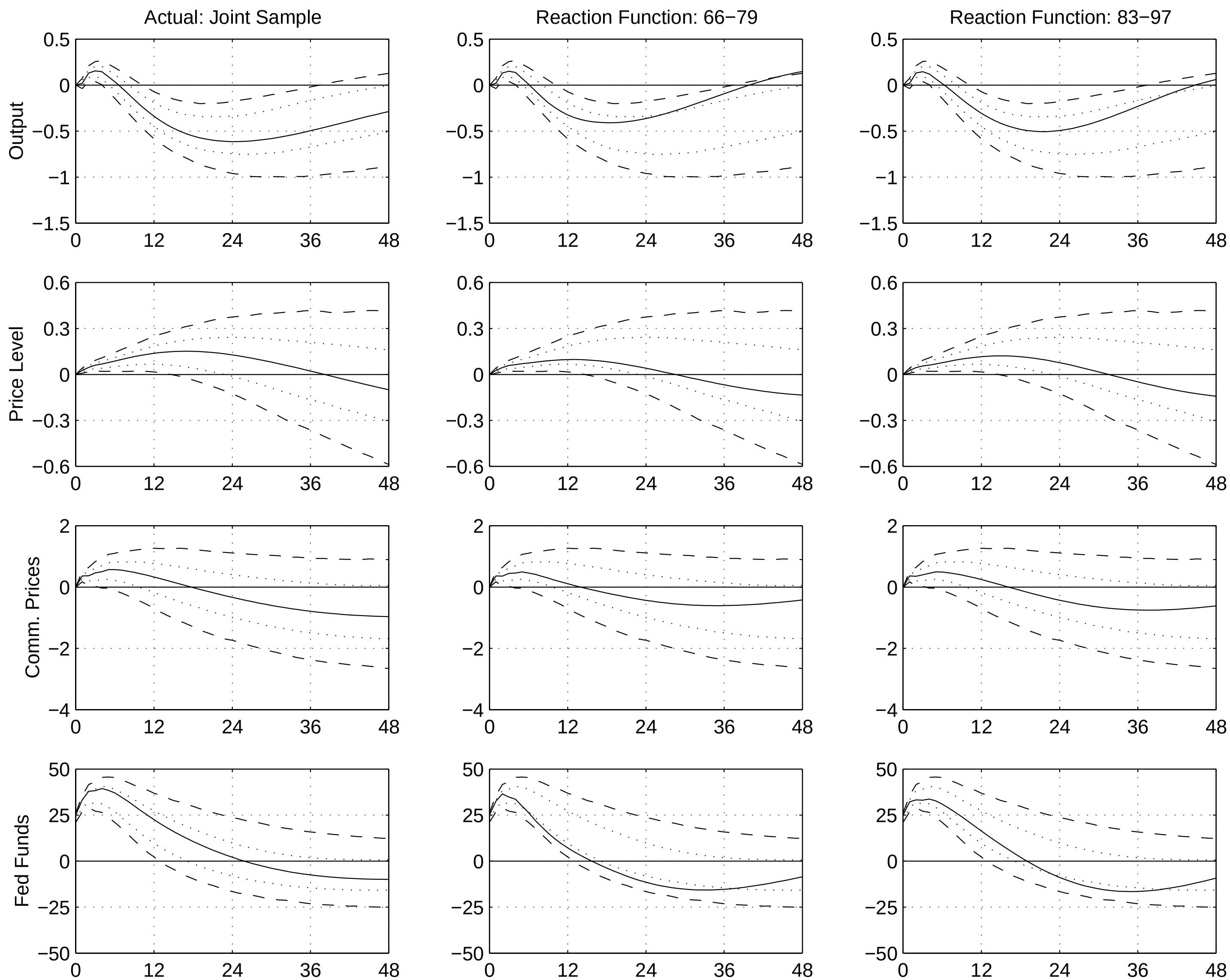


Figure 2.9: Counterfactual Impulse Response Functions with Varying Reaction Function: Responses to a 25-basis point contractionary monetary shock

shorter, and the maximum reduction in output is less. But these distinctions are not enough to negate the conclusion of no discernible difference between the actual estimates and the dynamics under either counterfactual.

It could be argued that the above counterfactuals do not completely test the significance of the change in the endogenous component of monetary policy, as they only vary the reaction function instead of the whole monetary block.[31] Figure 2.10 compares the actual impulse responses of the joint estimation period with counterfactual results in which the activity block remained as in the joint sample but the appropriate sub-sample estimate of the monetary block was used. The simulated impulse responses again are remarkably similar to the actual estimated responses of the joint sample.

Overall, the differences between corresponding counterfactuals of figures 2.9 and 2.10 are barely perceptible. Thus regardless of how it is measured in this context, the endogenous component of policy does not appear to be a significant source of the cross-sample variation illustrated in figure 2.8. Of course, all these experiments are conditioned upon the assumption that there is no change in the activity block (i.e. the rest of the model) when some aspect of policy is changed. But that assumption itself can be examined.

The other candidate explanation is a change in the dynamics of the private economy itself. In this case, counterfactual experiments can be conducted by replacing the activity block of the joint sample estimation with that of each of the sub-sample estimates. If the dynamics of the broader economy were not appreciably different across the sub-samples in question, then the impulse responses would appear largely

[31] But note that changes in the monetary block include any changes in the behavior of the banking sector as well, which arguably should be considered independently of any variation in monetary policy.

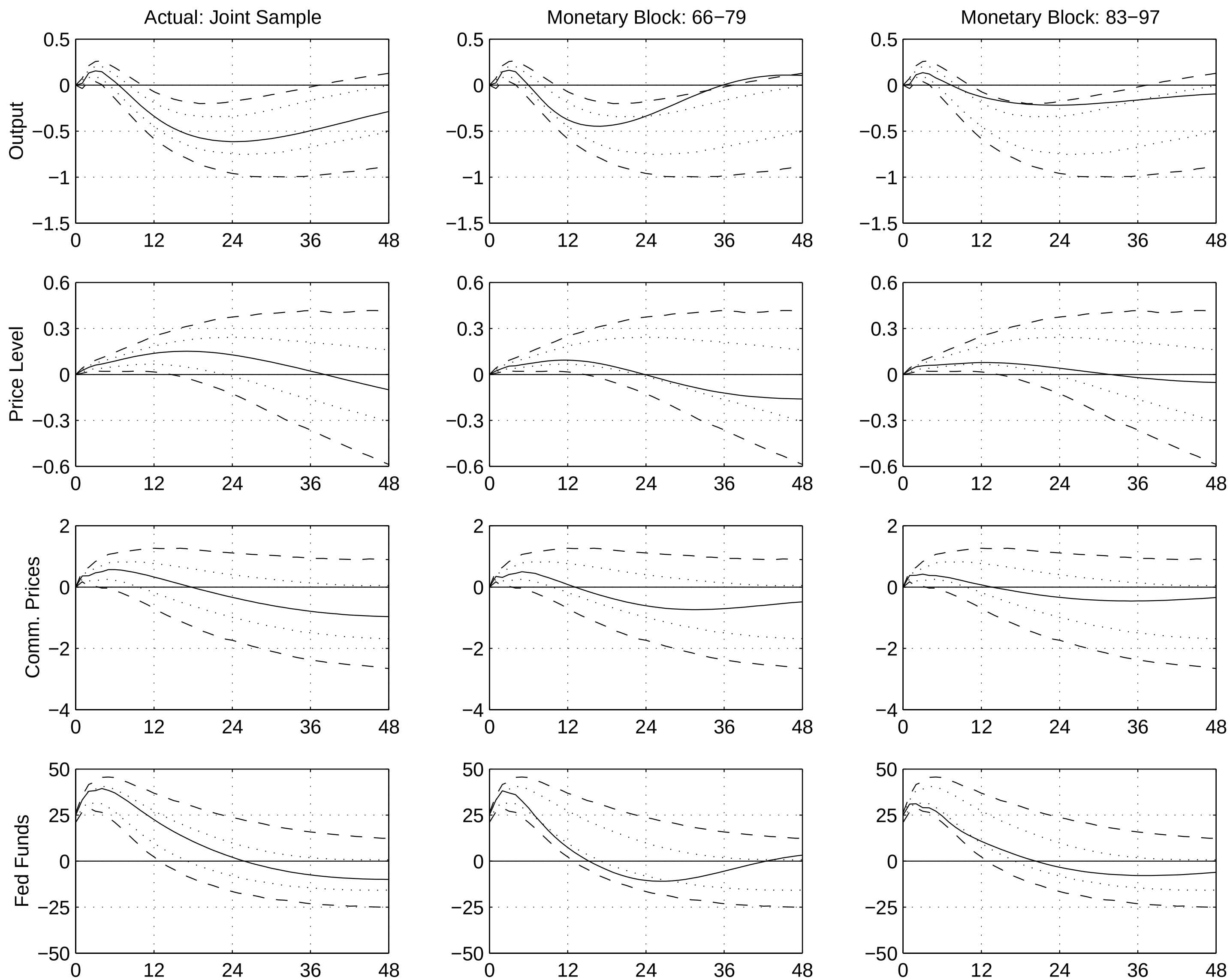


Figure 2.10: Counterfactual Impulse Response Functions with Varying Monetary Block: Responses to a 25-basis point contractionary monetary shock

the same. In effect, these experiments mirror those above: now the reaction function is held fixed at the values estimated over the joint sample period, and the remainder of the model is varied counterfactually.

Figure 2.11 compares these impulse response functions. As in the experiments above, relatively little change is found in the dynamics of the funds rate itself. However, the dynamics of the remaining variables change quite noticeably. The most dramatic is output, which shows a smaller (and earlier) decline for the 1966 – 1979 counterfactual, followed by an *increase* in point estimate within three years of the policy shock. For the 1983 – 1997 counterfactual, the converse occurs: the decline in output is much sharper — lying below even the 95% error band on the estimated model for most of the first year after the shock — and larger in magnitude. Notice also that the counterfactuals generate larger price puzzles and commodity price declines than the actual estimated model. In total, figure 2.11 provides compelling evidence of a economically significant change in the underlying structure of the non-policy portion of the U.S. economy during the sample period in question.

To mitigate concerns about the role of policy within the monetary block, an additional pair of counterfactual experiments were simulated. Graphed in figure 2.12, the activity block Y_t was replaced with the corresponding sub-sample values, while the monetary block M_t was held fixed at the joint sample estimates. The counterfactual impulse responses nearly match those of figure 2.11, with the price level responses slightly more pronounced in this case. These results even more strongly identify the activity variables as the source of the cross-sample variation. The VAR is not able to identify why the broader economy has changed over time, but only suggest that it cannot be attributed to monetary policy.

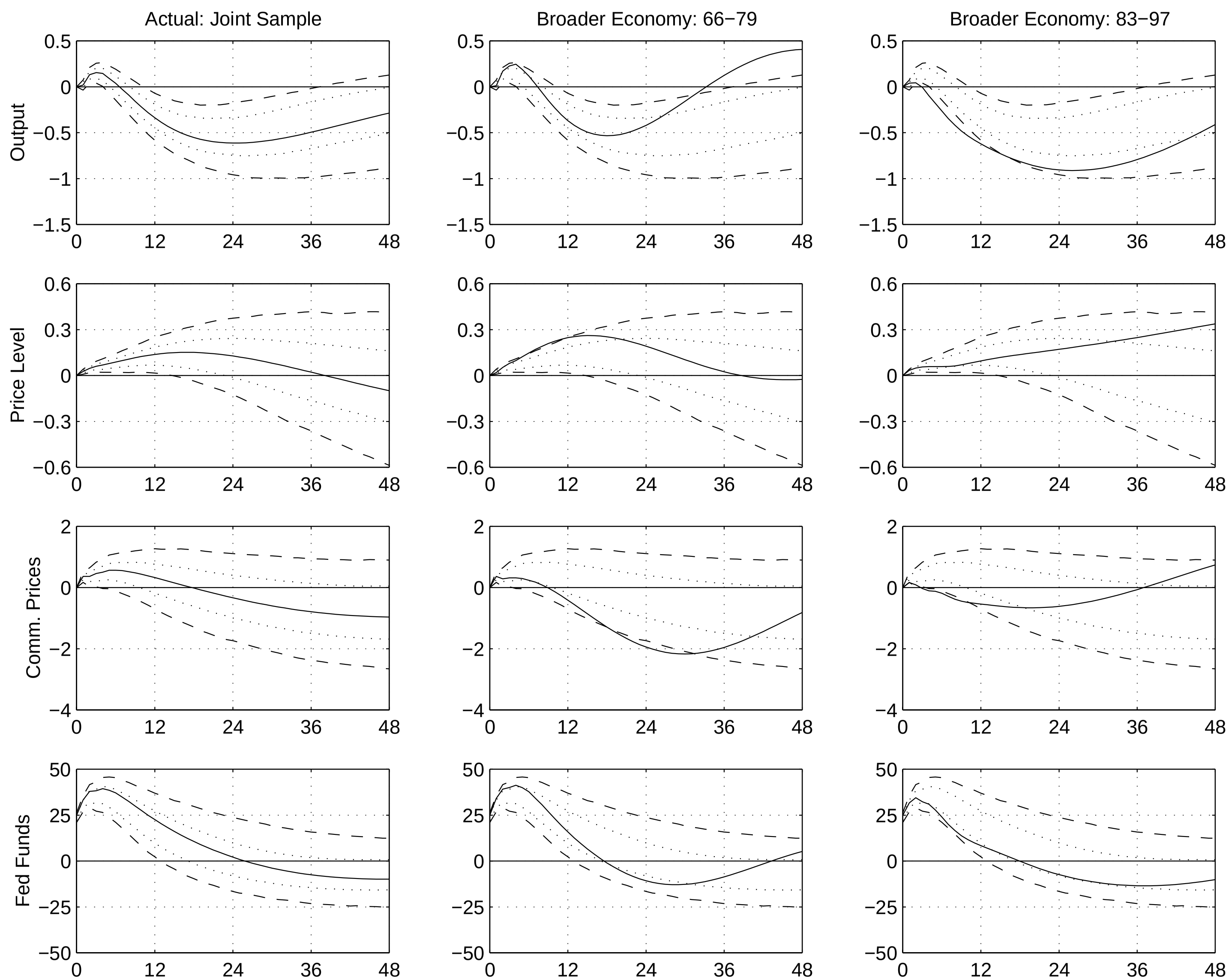


Figure 2.11: Counterfactual Impulse Response Functions with Fixed Reaction Function: Responses to a 25-basis point contractionary monetary shock

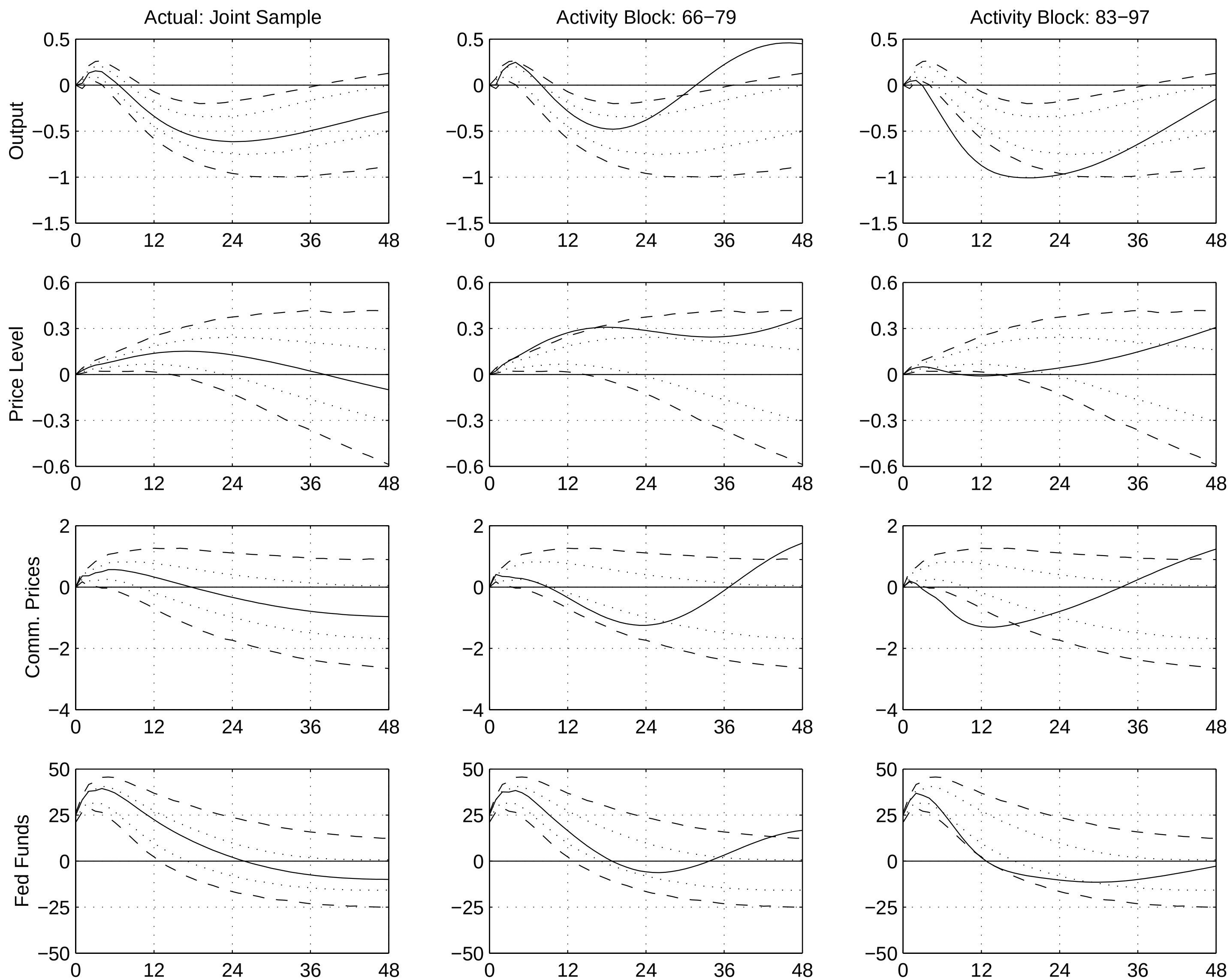


Figure 2.12: Counterfactual Impulse Response Functions with Fixed Monetary Block: Responses to a 25-basis point contractionary monetary shock

2.4.4 Further Counterfactual Experiments

Although the above findings reveal some statistically discernible differences in the form of the endogenous monetary policy rule, most of the cross-sample variation in the dynamics can be attributed to changes in the non-policy sector of the economy. With this view of the results, figure 2.11 has another interesting interpretation. It shows the impact of imposing the same reaction function — one estimated for the joint sample period — on the two distinct sub-samples. Under this interpretation, the form of the reaction function is revealed to be a much more significant source of fluctuations. That is, once the broader economy is assumed to differ between the two sub-sample periods, moderate modifications to the monetary reaction function generate substantially different dynamics.[32]

This interpretation suggests an additional set of counterfactual experiments, in which sub-sample variation of the broader economy is incorporated into the null hypothesis. That is, under the assumption that the non-policy portion of the model differs across the two sub-sample periods but the policy rule remained constant, one sub-sample policy rule could be replaced by the other — since these estimated rules would be equivalent if the null were correct. Figure 2.13 examines this experiment for the 1966:01 – 1979:09 period; figure 2.14 shows the analogue for the 1983:01 – 1997:12 period. For both of these figures, the left-hand column reports the actual responses to a 25 basis point contractionary policy shock for the sub-sample listed. The middle column reports the impulse responses when the actual reaction function is replaced with the one estimated for the other sub-sample, while the right-hand column does the same for the full monetary block. Again, the error bands from the

[32]Recall that the statistical tests of table 2.2 suggested only small variations in the specification of the policy reaction function over time, although table 2.4 suggests greater cross-sample differences in the policy rule specification.

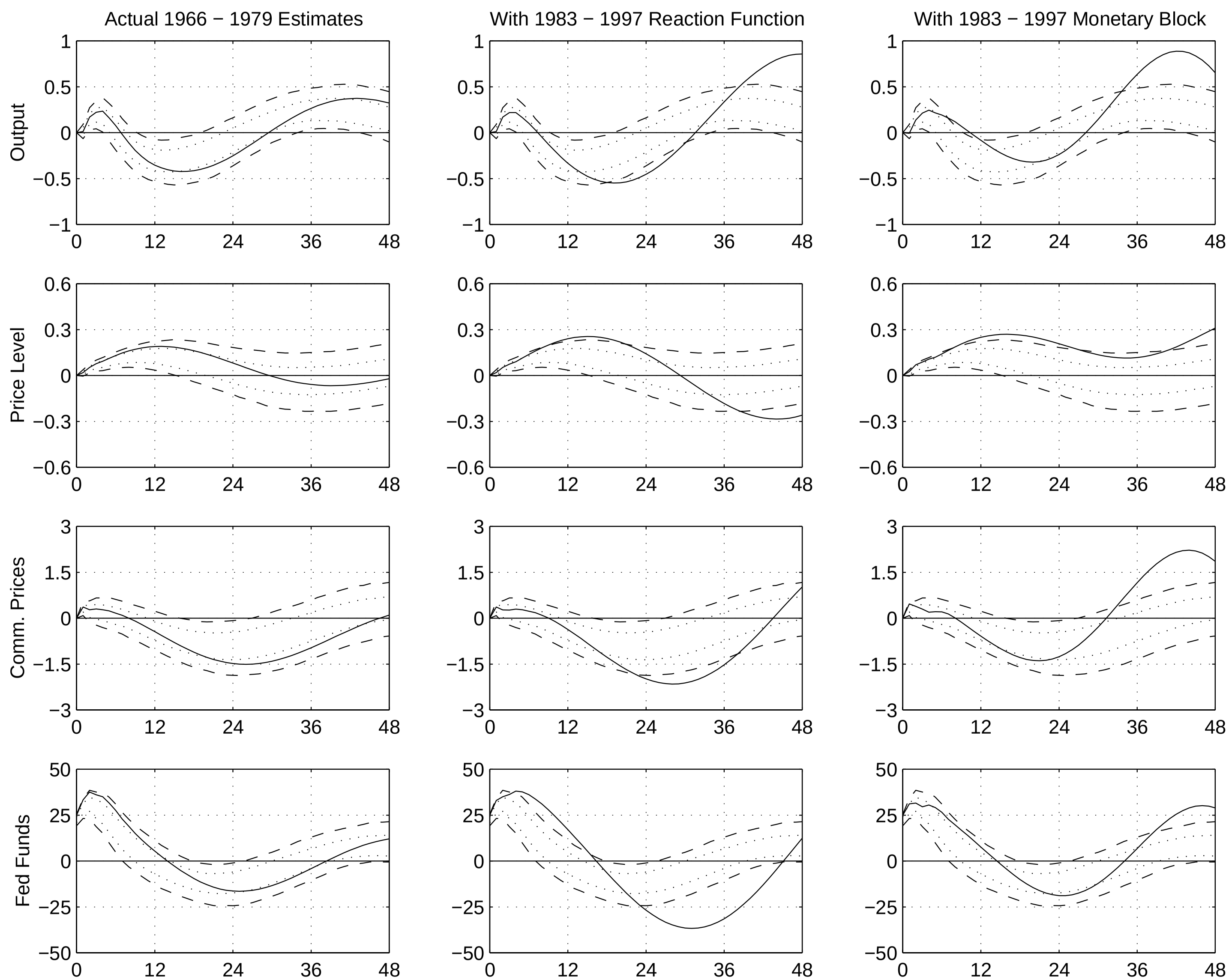


Figure 2.13: 1966 – 1979 Impulse Response Comparison:
Actual vs. counterfactual responses to a 25-basis point contractionary monetary shock

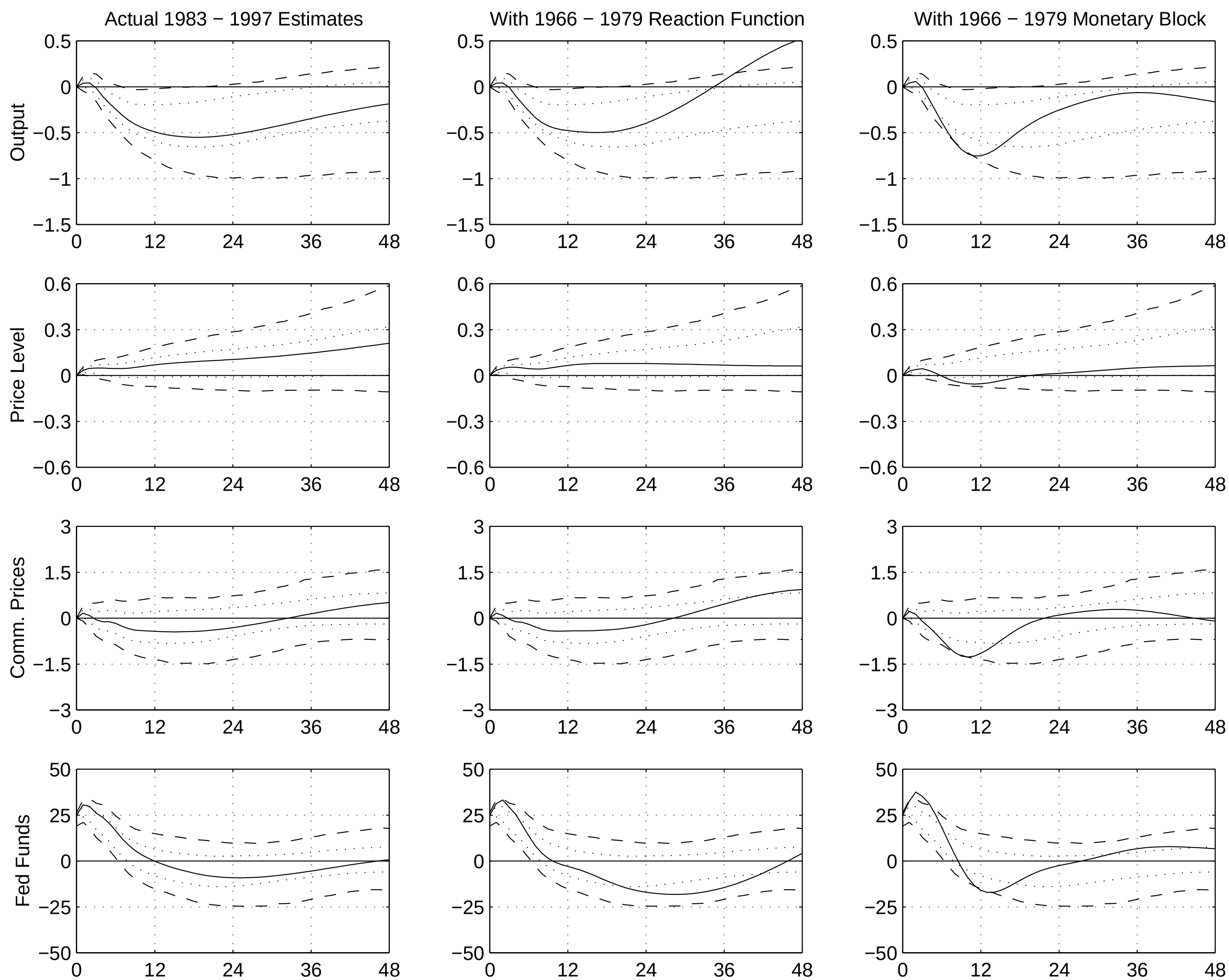


Figure 2.14: 1983 – 1997 Impulse Response Comparison: Actual vs. counterfactual responses to a 25-basis point contractionary monetary shock

actual model are replicated in the center and right-hand columns.

Figure 2.13 reveals greater instability in the impulse responses of the counterfactual experiments vis-à-vis the in-sample estimated responses. The funds rate response, in particular, appears much more volatile in the presence of the post-Volcker reaction function. While this difference is not as strong when the full 1983:01 – 1997:12 monetary block is employed, the activity variables nonetheless exhibit greater volatility in response to contractionary policy shocks of identical size.

Figure 2.14 also illustrates some differences between the actual and counterfactual results for the 1983:01 –1997:12 sub-sample, but these are less extreme than those of figure 2.13: in general, the counterfactual responses lay inside the error bands of the estimated in-sample responses. (The main exception appears to be the response of output under the counterfactual reaction function, about three years after the initial shock.) The counterfactual monetary block does yield a pronounced dip in all the activity variables during the first year after the policy shock, but these responses are not inconsistent with the sampling variation of the original estimates.

These findings, especially those of figure 2.13, provide stronger evidence against the null of no change in the reaction function specification over time. More significantly, they also are inconsistent with the view of the pre-Volcker reaction function as a main source of macroeconomic instability. Substituting the Volcker-Greenspan era reaction function (or monetary block) into the 1966:01 – 1979:09 sub-sample actually generates *additional* instability in response to exogenous policy innovations. Conversely, replacing the 1983:01 – 1997:12 reaction function (or monetary block) with the "sub-optimal" one from the Martin-Burns-Mitchell era does not appear to dramatically worsen the nature of the responses for this more recent period. By implication, the non-policy aspects of the model appear to be the determining factors for degree of instability realized.

Given the analysis and rhetoric of the Taylor rules literature, this conclusion is quite surprising. Clarida et al. (1998) illustrate how an insufficiently strong response to inflation in a Taylor rule specification can lead to instability through self-fulfilling expectational bubbles — and then estimate a pre-Volcker Taylor rule that exhibits this deficiency. In light of the above results, that finding is, if anything, reversed here. Also recall the net response coefficients reported in table 2.4 were shown to be inconsistent with the restrictions suggested by the Taylor rules literature. This disparity between the VAR and the Taylor rules results deserves further investigation.

2.4.5 Interpretations

It is important to recognize that the counterfactual experiments do not need to have a policy interpretation for the above conclusions to hold. Indeed, in the face of the Lucas critique, it is difficult to conclude that these experiments represent valid policy simulations. That has not prevented other authors in the VAR literature from conducting similar experiments and drawing inference about alternative policy regimes; recent examples include Bernanke et al. (1997), Sims and Zha (1998) and Sims (1998b). These papers assume that the Lucas critique has limited relevance for their simulations.[33]

For counterfactual experiments of this kind to admit a policy interpretation, two conditions must be satisfied. First, the estimated reaction function must be a structural equation, so that replacing the coefficients reflects changing "deep parameters" of the Fed's objective function. Second, the reduced-form equations of the rest of

[33]Leeper and Zha (1999) propose a clever alternative: they perturb the path of policy shocks with counterfactual innovations that are within the average MSE of the actual shocks. Similarly, a "plausible" perturbation to the policy rule could be defined as one in which the two rules are statistically indiscernible. I briefly develop this concept below.

the model must be largely invariant to the hypothesized change in the policy rule.

Obviously the closer the similarity between the actual and the counterfactual policy rules, the less likely this second condition will be violated. But the macroeconomic literature has not yet determined the "elasticity" of private sector behavior to the kind of changes in the policy rule explored in the VAR papers just cited. Blanchard (1984) reports no evidence of a change in a Phillips curve relationship as a consequence of the Volcker disinflation. More generally, in a meta-analysis that surveyed the empirical literature since Lucas, Jr (1976), Ericsson and Irons (1993) found no support for the empirical relevance of the Lucas critique. Thus, the range of alternative policy rules that will cause — or not cause — the Lucas critique to "kick in" is still an open question.[34]

As for the first condition, Rudebusch (1998a, p. 908) has argued that the reaction functions estimated with semi-structural VAR models are themselves structural equations, and that VAR practitioners should treat them as such. In this case, it would be legitimate to consider replacing the policy rule of one period with that of another. However, if the estimated policy rule changes for reasons other than changes in the Fed's preferences (the "deep parameters" of the Fed's optimization problem), then Rudebusch's interpretation may not be appropriate. For example, if the policy rule were revised in response to changes in the broader economy only — which would be likely if the rule were set "optimally" — then the estimated VAR policy rule (or Taylor rule) would vary *despite no change in the preferences of the Fed.* This is

[34]Although it is rarely recognized explicitly, this issue also will arise in stylized macroeconomic models that include a Taylor rules equation. Most of these models feature an aggregate price level or inflation equation (such as a Phillips Curve), and an aggregate output equation that themselves have no *a priori* claim to policy invariance. For example, a Phillips curve based on a Calvo-Rotemberg price setting model treats the frequency of price changes as a fixed parameter, when it is likely to vary with the monetary policy regime. The same can be said of "optimizing" IS (or AD) relationships based on Euler equations, which depend on the path of expected interest rates.

simply an application of the Lucas critique "in reverse," from the economy to an optimizing monetary authority.

To address when these conditions will hold, Sims (1997, 1998b) has proposed a standard based upon the *ex post* "plausibility" of the results. If variation in the policy rule results in "drastically different behavior of the economy" (Sims, 1998b, p. 153) as judged by the resultant impulse responses, the experiment is judged invalid. This standard has a bias toward "plausible" results of no measurable impact of the policy variation. An *ex ante* justification for the legitimacy of a counterfactual experiment would be preferable. That is, the specification of the counterfactual rule itself should not appear too dissimilar from the actual rule. Arguably the counterfactual simulations above could satisfy this criteria: the policy instruments and rule specifications are the same across the sub-sample periods; only the weights in the reaction functions vary. The statistical tests of table 2.2 suggests it is difficult to distinguish among the estimated sub-sample reaction functions, thereby providing an *ex ante* metric for "similarity." Furthermore, these counterfactuals are based on estimated reaction functions; they are not products of *ad hoc* fiddling with the rule coefficients.

Given the general failure of structural models to provide reasonable and robust estimates of "deep parameters" (see Heckman, 1999), counterfactual simulations may be the best way to estimate bounds on alternative policy regimes that are not in the data. Were the reaction function the only thing that varied across the sub-samples, observing different policy regimes would provide information about the role of endogenous policy directly. But the results of this section highlight the likelihood of economically significant instability residing as much in the non-policy portion of the model as in the policy rule.

None of the results of section 2.4.3 depend on the counterfactual experiments having a policy interpretation. Rather, I propose these counterfactual experiments only as a way to complement the statistical tests with quantifiable measures of economic significance. The results attribute most of the cross-sample variation to instability in the broader economy. This finding does not imply that the form of the policy rule is irrelevant, only that the historical behavior of the periods in question have a much larger non-policy component.

2.5 Conclusion

The introduction of this paper put forth two questions. One asked how the endogenous and exogenous components of U.S. monetary policy have changed over the post-war period. The statistical evidence of section 2.3 revealed a much greater dispersion of the exogenous policy shocks before the "Volcker disinflation" (1979:10 – 1982:12) than after, but the historical decompositions of section 2.4 indicated a relatively minor role for the policy shocks in either sub-sample investigated. However, the results of both sections highlighted the need to treat the pre- and post-Volcker samples as distinct policy regimes.

The converse was evident in the analysis of the endogenous reaction function (policy rule): statistical tests of a change in the parameterization of the rule were inconclusive, even after excluding the monetary targeting experiment of the Volcker disinflation period. The counterfactual simulations developed in section 2.4 reinforced the lack of any statistically significant evidence of cross-sample variation when the rest of the model was specified as unchanged over the full sample period, 1966:01 – 1997:12. But once the activity block, representing the broader economy, was allowed to vary across the sub-samples in question — as strongly evidenced by the statistical tests of section 2.3 — the counterfactual simulations indicated a much more sizable impact of varying the coefficients of the reaction function.

The counterfactual simulations in this paper are more modest than those of either Bernanke et al. (1997) or Sims (e.g. Sims, 1998a,b; Sims and Zha, 1998): the experiments only are used to quantify economic significance rather than to simulate alternative policies. Nonetheless, conditions under which these simulations may permit a policy interpretation are presented at the end of section 2.4. In the face of continued disagreement among macroeconomists as to the proper specification of a structural model of the economy, limited counterfactual simulations from semi-structural VAR models may become more popular. Criteria for such experiments are developed in section 2.4 as well, where the claim is made that the simulations investigated herein would satisfy these criteria.

The initial question of the introduction asked whether policy was the most important contributor to the change in the nature of economic fluctuations observed over the 1966:01 – 1997:12 sample. Recent work in the Taylor rules literature emphasizes how different monetary policy rules may lead to sizable differences in the dynamics of the economy. As already mentioned above, the results of this paper suggest that the cross-sample instability in the broader economy cannot be attributed primarily to variation in the monetary policy regime. By not explicitly modeling the dynamics of the rest of the economy, the methodology of much of the current Taylor rules literature assumes away any change in the structure of the rest of the model. Note that my results do not question the normative contributions of the Taylor rules literature. But the data cannot support the positive conclusion that monetary policy is responsible for the observed historical patterns.

Moreover, the additional counterfactual implications (shown in figures 2.13 and

2.14) do not lend credence to the theory of strong sub-sample differences in the nature of the policy rule. Indeed, they appear to contradict this proposed explanation. In order to better understand this issue, it would be desirable to recast the semi-structural VAR model in a form that more directly nests the Taylor rule specification. Reconciling the approaches — and ultimately the conclusions — of these two literatures would be an important step in improving our understanding of the role of monetary policy in economic fluctuations.

In the end, the results of this investigation suggest that over the past 15 years or so, the Fed has benefitted as much from good fortune as from wise policy decisions. The macroeconomy appears to have been inherently more unstable during the 1966:01 – 1979:09 period. An investigation of what changes in the non-policy structure of the economy occurred during the sample period examined, and what were the macroeconomic consequences of this variation, is the next logical step in this research program.

CHAPTER III

Monetary Factors in the Long-Run Co-movement of Consumer and Commodity Prices

3.1 Introduction

Commodity prices play a peculiar yet important role in modern monetary economics. Many researchers have proposed using commodity prices to forecast general consumer prices. A number of empirical papers have uncovered varying degrees of success for commodity prices in this role, although most recent evidence suggests this relationship — if it had ever existed — has broken down (Blomberg and Harris, 1995; Cecchetti, 1995; Garner, 1995; Webb and Rowe, 1995; Furlong and Ingenito, 1996). Some authors have moved beyond simple forecasting relationships to investigate whether commodity prices can serve as an indicator of inflationary pressures for monetary policymakers.[1] Others (implicitly) posit such a role: following seminal work by Sims (1992), it has become commonplace for monetary vector autoregressive (VAR) studies to include an index of commodity prices in their empirical specifications,[2] despite evidence that commodity prices do not serve this particular role well (Hanson, 1998). Non-VAR investigations of the form of the Federal Reserve's "reac-

[1] Woodford (1994) offers a critique of this approach.

[2] Recent work includes Cochrane (1994), Christiano et al. (1996a), Leeper et al. (1996), Bernanke et al. (1997), Bernanke and Mihov (1998) and Christiano et al. (1998).

tion function" recently have included commodity prices for similar reasons. (See, e.g., Clarida et al., 1998; Sims, 1999.) Interest in a potential policy role for commodity prices resurfaces every few years in the financial press as well.

These macroeconomic studies generally treat the data-generating process for commodity prices as given with respect to current monetary policy actions. In contrast, most of the literature on commodity price determination has tended to focus on microeconomic factors and emphasized how the speculative nature of certain storable commodities, interacting with occasional stock-outs, may lead to highly volatile price paths.[3] Straddling these two approaches, Pindyck and Rotemberg (1990) argue that a large portion of the volatility of individual commodity prices represents "excess co-movement" that cannot be attributed to aggregate or macroeconomic factors, including money supplies.

All the studies cited above examine the short-term relationship between consumer and commodity prices (or their rates of growth). Horizons beyond a few years are rarely considered. In contrast, this paper investigates the long-run relationship of consumer and commodity prices, with particular focus on the role of monetary policy. Neoclassical theory suggests that in response to a nominal innovation in the economy, both variables ultimately should respond in the same direction and proportion. If the growth rate of any aggregate price index is ultimately a monetary phenomenon, then it may be possible to exploit the long-run co-movement of distinct aggregate price indices to uncover information about monetary pressures in the economy.

The historical antecedent for this avenue of investigation begins with the literature on sluggish price adjustment to disequilibrium conditions (see, e.g., Mussa, 1981). A modern interpretation, cast in a dynamic general equilibrium framework,

[3] See e.g. Williams and Wright (1991) and Deaton and Laroque (1992).

is offered by Ohanian and Stockman (1994).[4] The most well-known applications of this literature are models of the monetary determinants of exchange rates, namely Dornbusch (1976) and Mussa (1982). In these models aggregate prices adjust only slowly to monetary shocks, whereas flexible exchange rates "overshoot" their eventual equilibrium response to such shocks. In the long-run, nominal exchange rates and aggregate prices move proportionally.

Note that commodity prices share a key attribute of exchange rates: both are flexible prices that can adjust to shocks more quickly than sluggish consumer prices.[5] Frankel (1986) extended the overshooting model to commodity prices, and a few authors (e.g. Hook and Walton, 1989; Boughton and Branson, 1991) have extended this variation further. An arbitrage condition for storable commodities links their rate of price change to the rate of interest (the Hotelling model). In the presence of sticky consumer prices, a nominal shock to money results in higher real balances and a reduction in interest rates. Since both consumer and commodity prices eventually will rise in response to this nominal innovation, commodity prices must overshoot in order to satisfy the arbitrage condition.

Since commodity prices are more flexible than consumer prices, they can provide information about current realized nominal innovations that eventually will affect consumer prices, albeit with a lag. Much of the empirical work that examines the forecasting power of commodity prices for consumer prices exploits this idea. If commodities are inputs into final (consumer) goods, then any shock to commodity

[4] Ohanian and Stockman (1994) present a two-sector model in which the prices of one sector are fixed for a single period, then adjust completely. This approach contrasts with the New Keynesian literature, which models the sluggish adjustment of aggregate prices more explicitly. (See Rotemberg, 1987)

[5] Ohanian and Stockman (1994) anticipate this aspect when they briefly note in their conclusion that commodity prices may be a close approximation for the flexible price sector of their model.

prices could eventually be passed through to consumer prices.[6]

Commodity prices tend to be much more flexible than consumer prices because the former are determined in (nearly) continuously-clearing auction markets, whereas the latter are set in "customer markets" characterized by some degree of market power and by relationships between buyers and sellers that preclude continuously allocative pricing.[7] Since commodities are traded as financial assets in these auction markets, their prices should reflect the inflationary expectations of forward-looking investors. Thus, commodity prices can also convey information about expected future nominal innovations.

As flexible asset prices, commodity prices reflect both current realized and future expected nominal innovations. These innovations eventually affect consumer prices as well. This is the key ingredient for the empirical model developed and tested in this paper. Commodity prices do not need to be the channel by which these shocks are transferred to consumer prices; rather, they simply can be an indicator by virtue of the market structure in which they are determined. In particular, my empirical model does not depend in any way on the degree of pass-through being complete or time-invariant. Rather, my model rests upon the premise that consumer and commodity prices should move equiproportionally in response to nominal shocks in the long-run.

In section 3.2 of this paper, I review the univariate and bivariate statistical properties of the price series. I section 3.3 I propose a vector autoregressive model for consumer and commodity prices, in which I exploit the long-run properties of a nom-

[6]Indeed, this "chain-of-production" structure can account for some of the lags in consumer price changes, if multi-stage production simultaneously takes multiple time periods to complete. Fischer (1981) discusses this approach; see Clark (1999) for more recent empirical work on this topic.

[7]Blinder (1994) reported the prices of surveyed U.S. firms were fixed for slightly less than one year.

inal shock to identify the component attributable to monetary policy. The results, reported in section 3.4, imply monetary policy accounted for a large portion of the historical fluctuations of the two price series. In particular, much of the inflation commonly attributed to so-called "supply shocks" appears to have a monetary root. My conclusions are stated in section 3.5.

A similar idea is the basis for recent empirical studies of the role of monetary shocks in the behavior of real exchange rates. Several studies employing VAR techniques have found some evidence of both overshooting dynamics at short horizons and co-movement between exchange rates and the price level at long horizons.[8] What limited empirical work exists on commodity price determination in response to monetary disturbances has focused on the overshooting behavior (or lack thereof) and not the long-run behavior of the two series. One of the benefits of the empirical model I develop below is the ability to test for both short-run dynamics and long-run co-movements implied by theory. In so doing, I am able to provide a firmer empirical foundation for the role of commodity prices in macroeconomic research, particularly in monetary policymaking.

3.2 Preliminary Statistical Analysis

The aggregate price measure used in this study is the U.S. consumer price index (CPI) less shelter for all urban consumers.[9] A number of different commodity price indices are commonly cited in both the financial press and academic work; here I focus on several indices from the International Monetary Fund (IMF), primarily an

[8]See. for example, Lastrapes (1992), De Arcangelis (1993), Clarida and Galí (1994), Enders and Lee (1997), and Rogers (1998).

[9]This measure is used to avoid problems with the accounting of mortgage interest prior to 1982. The GDP deflator and the all-items CPI (not shown) yield very similar results.

overall world commodities index.[10] More information about the composition of the IMF indices is available in appendix A.

The final month of a quarter constitutes the quarterly observation for each price index series. The data begin with the first quarter of 1959 and run through the second quarter of 1998. Figure 3.1 plots the logarithmic levels of each of the series: the CPI in the top panel, the IMF spot index in the middle panel, and the implied real commodity price (log IMF less log CPI) in the lower panel. Figure 3.2 plots the annualized growth rates.

The general upward trend in the levels of both nominal series is demonstrated in figure 3.1, where the consumer price index increases much more smoothly than the commodity price index. In contrast, the implied real commodity price has declined steadily over time, interrupted by a few sizable jumps (the largest in the early 1970s). The observed trend behavior of real commodity prices is consistent with the Prebisch-Singer hypothesis.

Several differences between the growth rates of the nominal series stand out in figure 3.2: CPI inflation is almost exclusively positive, whereas commodity price growth is often quite negative; CPI inflation appears to be more serially correlated; the magnitude of the annualized changes in commodity prices dwarfs those for CPI inflation. However, both series exhibit more volatile growth rates between the early 1970s and the mid-1980s.

Later I will examine some narrower price indices that reflect certain commodity groups. My empirical model is general enough to allow consideration of specific commodities, such as gold or petroleum. However, a broad-based commodity price

[10] I have also examined the Commodity Research Bureau (CRB) spot index, the CRB futures index, and an index reported by the Journal of Commerce. The various indices, despite containing somewhat different mixtures of individual commodities, all behave in a qualitatively similar manner. Results are available upon request.

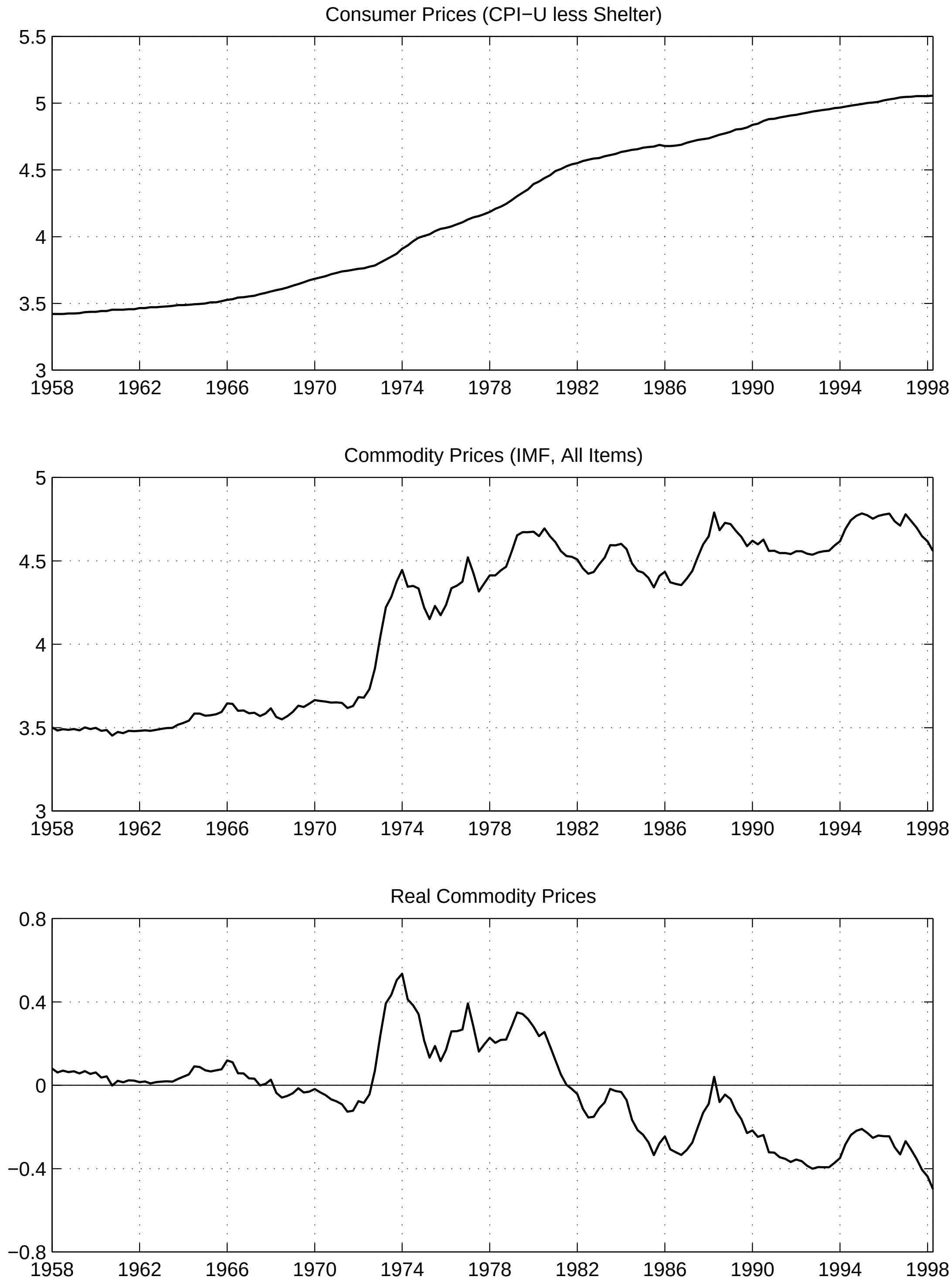


Figure 3.1: Logarithmic Levels of Consumer Prices and Commodity Prices, All Items Index

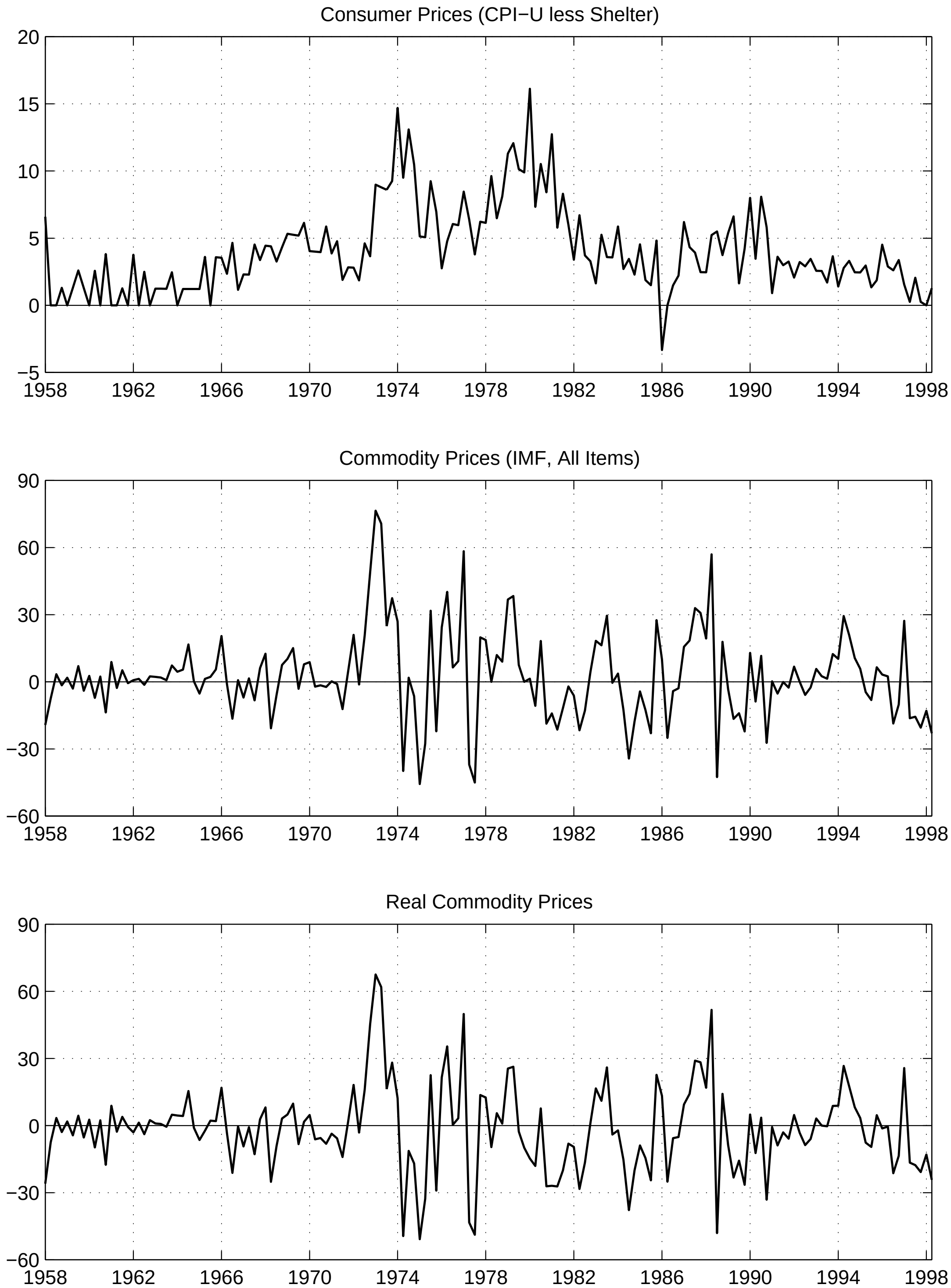


Figure 3.2: Annualized Percentage Changes of Consumer Prices and Commodity Prices, All Items Index

index should be preferable to any single commodity for the purpose of isolating expectations about inflation and monetary policy, as "commodity-specific" shocks are more likely to average out with a diversified basket of commodities. Even so, figure 3.2 clearly indicates that the commodity price index is substantially more volatile than the aggregate consumer price index. The estimation technique can be viewed as a signal-extraction procedure: an attempt to separate real, market-specific shocks from nominal shocks to both prices.

3.2.1 Univariate Statistical Properties

Table 3.1 provides statistical confirmation for several of the observations made above. In addition to the IMF overall world commodity price index, summary statistics for several component indices of this overall price index are reported. Statistics for the full sample period are given first. These broadly confirm the attributes noted above in figures 3.1 and 3.2: the variance of each commodity price index is substantially greater than that of the CPI, and the implied real commodity price growth rates are negative for all series. Notice also that the contemporaneous correlation of any of the commodity price indices with CPI inflation tends to be only minimally positive; the correlation coefficients range between 0.1 and 0.2.

The rest of table 3.1 separates the full sample into three distinct sub-periods: 1959 – 1971, 1972 – 1982, and 1983 – 1998. Poor harvests worldwide in late 1972 and early 1973 — which preceded the first OPEC oil price hike by about a year — led to sharp increases in the IMF overall index that motivated the timing of the first division. As figure 3.2 illustrates, CPI inflation climbed dramatically in 1972, averaging nearly 9%, before jumping into the double-digit range in 1973. The second break point is motivated by the Volcker disinflation, which by the end of 1982

Table 3.1: Comparison of Nominal Commodity Price Indices

Annualized Quarterly Growth Rate over Specified Period

	CPI Inflation	IMF Commodity Price Indices: Overall	Food-stuffs	Agri. Raw Mat'ls	Metals
Full Sample: 1959 Q1 – 1998 Q2					
Mean	4.13	2.71	2.21	3.94	2.57
Standard Deviation	3.19	19.67	26.20	23.73	30.31
Correlation w/ Inflation	1.0	0.198	0.177	0.103	0.107
Sub-sample: 1959 Q1 – 1971 Q4					
Mean	2.52	1.10	1.14	1.36	1.88
Standard Deviation	1.82	7.72	10.35	10.28	22.82
Correlation w/ Inflation	1.0	−0.038	−0.010	−0.021	−0.036
Sub-sample: 1972 Q1 – 1982 Q4					
Mean	7.57	7.30	6.61	9.03	6.11
Standard Deviation	3.31	28.84	39.23	35.55	31.61
Correlation w/ Inflation	1.0	0.160	0.155	−0.059	0.217
Sub-sample: 1983 Q1 – 1998 Q2					
Mean	3.04	0.81	−0.01	2.49	0.64
Standard Deviation	1.95	18.39	23.93	21.06	34.80
Correlation w/ Inflation	1.0	0.167	0.195	0.182	0.000

Table 3.2: Tests of Stationarity of Nominal Price Series
Augmented Dickey-Fuller test results for 1959:1 – 1998:2 sample

	Consumer Prices		Commodity Prices	
Transformation	ADF Statistic	Largest Autoregressive Root	ADF Statistic	Largest Autoregressive Root
Log level	−0.892	0.999	−1.153	0.991
with trend	−1.769	0.992	−1.580	0.967
First log difference	−2.210	0.859	−5.629	0.162
with trend	−2.088	0.856	−5.635	0.159
Second log difference	−4.488	−0.734	−6.517	−1.572

Notes: A constant and 6 lags were included in each test. Sample size is 156 observations for all tests. Critical values: without trend, −2.88 at the 5% level, −2.57 at the 10% level; with trend, −3.44 at the 5% level, −3.14 at the 10% level. Reject null of a unit root if ADF statistic is less than critical value.

arguably had wrung inflation out of the economy: the average quarterly inflation rate for 1983 of 3.5% is scarcely above the 3.3% for 1971.

The differences across these sub-periods are dramatic: for each series, the mean and standard deviation of price growth is far greater between 1972 and 1982 than either sample before or after. The correlation between the commodity price indices and inflation also varies substantially across these sub-samples.

Table 3.1 also confirms the pronounced upward trends of both the CPI and the overall IMF index that were exhibited in figure 3.1. Tests for stationarity are shown in table 3.2, which reports augmented Dickey-Fuller statistics for each series and the largest autoregressive root of the test. Both consumer and commodity prices appear to contain a unit root. Examining first differences of the two series suggests that commodity price growth is stationary, while inflation is on the borderline between being stationary or integrated. In the estimation below I will maintain that consumer prices are $I(1)$ — i.e., that inflation is stationary.

3.2.2 Cointegration Tests

I begin the multivariate analysis by testing whether consumer and commodity prices are cointegrated; that is, whether they share a common stochastic trend and tend towards a long-run equilibrium relationship. Notice that a statistical precondition for cointegration is for each series to be integrated of the same order. The above univariate results confirm that each series can be considered $I(1)$. The assumptions of my model imply that consumer and commodity prices will move together in the long-run in response to monetary shocks, which should cause each nominal price series to eventually rise equiproportionally. However, I also allow for the possibility of real shocks (most significantly, market-specific shocks to the supply and demand for commodities) that can drive a wedge between the long-run paths of these two series. If this real component itself is integrated (i.e. if some portion of these real shocks is permanent), then one should not expect to find evidence of cointegration between these two series.

I employ two conceptually different tests of the null hypothesis of no cointegration. The first test, due to Phillips and Ouliaris (1990), posits an unrestricted equilibrium relationship and tests whether the residual is stationary, as cointegration requires.[11] The second test, due to Horvath and Watson (1995), is a Wald-type test which requires the imposition of a particular cointegrating vector for the alternative hypothesis. A monetary shock should move both price series in proportion in the long-run; it should not affect the *real* commodity price. With both series measured in logs, this relationship implies the cointegrating vector $(1, -1)$. The test is implemented by estimating a VAR in vector-error correction form under the alternative.

[11] That is, the cointegrating vector is not imposed, but freely estimated by OLS.

Table 3.3: Cointegration Tests of Consumer and Commodity Prices

Test results for 1959:1 – 1998:2 sample

Cointegration Test	Test Statistic	Sample Size	Parameters
Log levels of both series:			
Phillips - Ouliaris (1990)	$Z_\rho = -7.728$	158	$n-1=1$
Horvath - Watson (1995)	$W = 3.966$	158	4 lags
Log level of commodity prices, first difference of consumer prices:			
Phillips - Ouliaris (1990)	$Z_\rho = -2.386$	158	$n-1=1$
Horvath - Watson (1995)	$W = 3.217$	158	4 lags

Notes: For Phillip-Ouliaris test, null hypothesis is no cointegration; alternative is a spurious cointegrating relationship. Reject null at 5% level if $Z_\rho < -21.5$; reject null at 10% level if $Z_\rho < -18.1$. For Horvath-Watson test, null hypothesis is no cointegration; alternative has cointegating vector $[1, -1]$. Reject null at 5% level if $W > 10.18$; reject null at 10% level if $W > 8.30$.

The results of each test are reported in table 3.3, along with the appropriate critical values. Neither test appears to provide evidence of cointegration between the levels of the two series. This empirical finding lends support to the identification scheme employed below which jointly decomposes the price series into real and nominal shocks, each with a permanent component.

Some researchers have suggested a cointegrating relationship between the *level* of commodity prices and the *rate* of CPI inflation. For completeness I examined whether such a relationship exists in the data, but as shown in table 3.3, I could not find any supporting evidence within the sample period examined.

3.2.3 Granger Causality

Many studies of commodity prices as an indicator variable examine whether commodity prices "Granger-cause" inflation. Granger causality tests measure the incre-

Table 3.4: Granger-Causality Tests of Consumer and Commodity Prices

Test results for 1959:1 – 1998:2 sample

Null Hypothesis	F–Statistic	Significance Level
C does not Granger-cause P	4.391	0.0004
P does not Granger-cause C	0.429	0.8584
ΔC does not Granger-cause ΔP	4.251	0.0006
ΔP does not Granger-cause ΔC	0.233	0.9651
$\Delta^2 C$ does not Granger-cause $\Delta^2 P$	1.580	0.1571
$\Delta^2 P$ does not Granger-cause $\Delta^2 C$	1.796	0.1038
ΔC does not Granger-cause $\Delta^2 P$	3.975	0.0010
$\Delta^2 P$ does not Granger-cause ΔC	0.348	0.9101

Notes: All tests are constructed using six lags of each variable. Critical value is F(6,156). 'C' represents nominal commodity price index, 'P' represents consumer price index.

mental forecasting power of a candidate indicator for a particular series (conditional on lagged values of the series being forecasted). Table 3.4 summarizes the results of tests of Granger causality between the overall IMF index and the CPI. Commodity prices Granger-cause the CPI whether each is measured in levels, first or second differences, whereas consumer prices do not Granger-cause commodity prices. From the standpoint of forecasting future inflation, commodity prices do appear to contain useful statistical information.[12]

One cannot conclude from Granger-causality tests that commodity price changes "cause" consumer prices changes in any economically meaningful (i.e. structural) sense. Rather, these results suggest only that movements in commodity prices are temporally prior to, and correlated with, movements in the CPI. These results are

[12]These results are robust to the inclusion of trend and squared-trend terms.

consistent with the view that both series share a common factor, and that this common factor is realized in the flexible commodity prices before it affects the sluggishly-adjusting CPI. The results of section 3.4 attribute this common factor to monetary policy.

Interestingly, the first difference of the commodity price series also Granger-causes the second difference of consumer prices, as evidenced by the final pair of results in table 3.4. In other words, an increase in commodity prices helps forecast an acceleration of inflation. I do not further investigate this finding here, but note that it simply may be a statistical artifact driven by a single large episode in the 1970s. It is difficult to motivate such a relationship from a theoretical standpoint.

3.3 Structural VAR Estimation and Results

As stated in the introduction, this paper proposes a structural econometric model of the relationship between an index of commodity prices and the U.S. consumer price level. The model is "structural" in the sense that the primitive shocks driving the two series are endowed with economic meaning via restrictions on the dynamic responses of the series to each of these primitive innovations. In particular, I propose that the growth rates of consumer and commodity prices can be decomposed into two distinct innovations: a monetary or nominal component, ν_t, and a commodity-specific or real component, ψ_t.

The bivariate vector moving average (VMA) representation of this structural model can be written as

$$\Delta c_t - \Delta p_t = A_{11}(L)\,\psi_t + A_{12}(L)\,\nu_t \tag{3.1}$$

$$\Delta p_t = A_{21}(L)\,\psi_t + A_{22}(L)\,\nu_t \tag{3.2}$$

or, more compactly,

$$\begin{bmatrix} \Delta c_t - \Delta p_t \\ \Delta p_t \end{bmatrix} = A\left(L\right) \begin{bmatrix} \psi_t \\ \nu_t \end{bmatrix} \tag{3.3}$$

where Δp_t represents the rate of consumer price inflation and $\Delta c_t - \Delta p_t$ represents the growth rate of real commodity prices. According to the results in section 3.2, each of these transformed variables should be stationary. $A\left(L\right)$ is a (2×2) matrix of the lag polynomials $A_{ij}(L)$ for $i, j = 1, 2$. Each polynomial has lag order q.[13] The structural innovations are mutually and serially uncorrelated by assumption.[14]

The model as written above is underidentified. Identification is achieved by supposing that a monetary innovation is reflected one-for-one in the (log) level of any nominal price series when measured over some suitably long time period. In particular, a nominal shock will result in equivalent long-run movements in consumer prices and nominal commodity prices. Thus, real commodity prices should be unaffected by a monetary innovation in the long-run: $A_{12}(1) = 0$ in equation (3.1). The matrix of long-run responses, $A(1)$, is therefore lower triangular in equation (3.3).

Notice that the restriction on the co-movement of consumer and commodity prices to the unobserved nominal shock holds for the long-run only; the nominal shock is otherwise allowed to affect each series differentially in the short run. Hence real commodity prices can (and generally will) move over time due to a nominal innovation. But ultimately all nominal prices must rise together in response to a positive nominal shock, resulting in no long-run effect of a monetary shock on real commodity prices. An important implication is that permanent changes in real commodity prices (due to ψ_t shocks) preclude any cointegrating relationship between consumer

[13]For most of the subsequent estimation, results for a lag length of $q = 4$ are reported. Akaike and Schwarz Information Criteria suggested lag lengths of 5 and 3, respectively.

[14]Although included in the estimation, constant terms are suppressed for ease of exposition.

and commodity prices. In the absence of permanent shocks to real commodity prices, the common nominal component would make the two price series cointegrated. This implication is consistent with the statistical evidence cited in the previous section.

The commodity-specific shock (ψ_t), then, is assumed to reflect "real" or "relative" movements in the supply of or demand for commodities.[15] No restriction is placed on the nature of the commodity-specific shock: it can have both permanent and transitory components. Thus, this SVAR model can match the suggestions by some authors (e.g. Deaton and Laroque, 1992) that commodity prices have a temporary (mean-reverting) component, and yet not require the nominal shock to be the only source of temporary movements in real commodity prices. Note that the long-run identifying restriction implies the nominal shock cannot be the source of any permanent changes in real commodity prices.

3.3.1 Estimation Technique

The structural vector autoregression that gives rise to the bivariate VMA representation is identified by the imposition of the long-run restriction on $A_{12}(1)$ discussed above. Numerous authors have utilized these types of restrictions for the purposes of identification of structural innovations; see, for example, Blanchard and Quah (1989) and Shapiro and Watson (1988). The estimation technique employed in this paper utilizes the instrumental-variables approach of Shapiro and Watson.

The bivariate VMA representation above can be inverted and written as a struc-

[15]Since my focus is on U.S. CPI data, changes in international monetary policy that do not impact U.S. inflation will also be reflected in the ψ_t shock. To a first-order approximation, monetary shocks abroad only affect real commodity prices from the U.S. vantage point (in the absence of any domestic monetary accommodation), so this classification satisfies the above definition of a "commodity-specific" shock.

tural vector autoregression (VAR):

$$B(L)\begin{bmatrix}\Delta c_t - \Delta p_t \\ \Delta p_t\end{bmatrix} = \begin{bmatrix}\psi_t \\ \nu_t\end{bmatrix} \tag{3.4}$$

where $B(L) = A(L)^{-1}$ is the 2×2 matrix of lag polynomials $B_{i,j}(L)$ for $i, j = 1, 2$. Writing out the equations from system (3.4) yields (after normalization):

$$\Delta c_t - \Delta p_t = \sum_{k=1}^{q} \beta_{11,k}\,(\Delta c_{t-k} - \Delta p_{t-k}) + \sum_{k=0}^{q} \beta_{12,k}\,\Delta p_{t-k} + \psi_t \tag{3.5}$$

$$\Delta p_t = \sum_{k=0}^{q} \beta_{21,k}\,(\Delta c_{t-k} - \Delta p_{t-k}) + \sum_{k=1}^{q} \beta_{22,k}\,\Delta p_{t-k} + \nu_t\,, \tag{3.6}$$

To impose the long-run restriction, notice first that the condition $A_{12}(1) = 0$ on the bivariate VMA representation of equation (3.3) is equivalent to $B_{12}(1) = 0$ in equation (3.4).[16] Since $B_{12}(1) = \sum_{k=0}^{q} \beta_{12,k}\,L^k$, this restriction can be imposed directly and equation (3.5) can be estimated by ordinary least squares (OLS) as

$$\Delta c_t - \Delta p_t = \sum_{k=1}^{q} \beta_{11,k}\,(\Delta c_{t-k} - \Delta p_{t-k}) + \sum_{k=0}^{q-1} \gamma_{12,k}\,\Delta^2 p_{t-k} + u_{1t}\,, \tag{3.7}$$

where

$$\gamma_{12,k} = \sum_{\ell=0}^{k} \beta_{12,\ell}$$

and $\Delta^2 p_t$ is the second difference of the price level (i.e. the first difference of inflation, Δp_t). Equation (3.5) for the real commodity price must be estimated by instrumental variables since the contemporaneous observation $\Delta^2 p_t$ that appears on the right-hand side is correlated with the error term. The instruments are a constant and lags 1 through q of both Δp_t and $\Delta c_t - \Delta p_t$.

The price equation (3.6) can be estimated by OLS as

$$\Delta p_t = \sum_{k=1}^{q} \gamma_{21,k}\,(\Delta c_{t-k} - \Delta p_{t-k}) + \sum_{k=1}^{q} \gamma_{22,k}\,\Delta p_{t-k} + u_{2t}\,, \tag{3.8}$$

[16]Since $A(1)$ is lower triangular, it follows that $B(1) = \left[A(1)\right]^{-1}$ is also lower triangular.

where

$$\gamma_{2j,k} = \frac{\beta_{21,0}\,\beta_{1j,k} + \beta_{2j,k}}{1 - \beta_{21,0}\,\beta_{12,1}}$$

for $j = 1, 2$ and for each $k = 1, \ldots, q$. u_{2t} is a linear combination of the structural shocks:

$$u_{2t} = \frac{\nu_t + \beta_{21,0}\,\psi_t}{1 - \beta_{21,0}\,\beta_{12,1}}\,.$$

Equation (3.8) follows from substituting for the contemporaneous real commodity price term (equation 3.5) on the right-hand side of equation (3.6) and renormalizing.

After equations (3.7) and (3.8) are estimated, the orthogonalized structural innovations, ν_t and ψ_t, can be recovered from the estimated (reduced-form) VAR residuals, u_{1t} and u_{2t}, with a Choleski decomposition.

3.4 Empirical Results

All the results below were estimated using quarterly data from 1959:1 to 1998:2 (unless noted otherwise).[17] Consistent definitions of the commodity price indices are not available prior to 1957. The estimation procedure is undertaken with the assumption that a 40-year span of data contains sufficient information about the long-run behavior of the series. Four lags are used in the estimation. The findings are fairly robust across different measured commodity price indices; results for the International Monetary Fund World Price index are discussed below. Results are also reported for several sub-indices that comprise this overall IMF index: food (32.9% of the broad index), agricultural raw materials (32.3%), and metals (26.7%).[18]

[17]The third month of each quarter is used as the quarterly observation.

[18]The remainder is accounted for by beverages (6.8%) and fertilizers (1.3%). See Appendix A for the composition of the IMF indices. The results for the last two sub-indices are not reported here, both in the interest of space and because these are essentially one-commodity indices: over sixty percent of the beverages sub-index is coffee (the remainder split almost equally between tea and cocoa beans) and the fertilizers sub-index is essentially phosphate. Otherwise, the results for these

For each commodity price series, three measures of the statistical importance of the structural shocks are reported. I start by examining the estimated impulse response functions (IRFs) to the identified shocks. Then I consider forecast variance decompositions, and finally historical decompositions of the series. The impulse responses demonstrate the effect on each series of a hypothetical one-period shock to each of the innovation terms,[19] while the variance decompositions report what fraction of the k–period ahead forecast error of a series is due to each of the identified innovations. The historical decompositions, on the other hand, estimate the contributions of the two shocks at each point in time during the sample period. The historical decompositions are more informative than the structural shocks themselves, which are white noise by nature of the estimation procedure. By utilizing the information contained in the vector moving average representation, the historical decompositions capture the dynamic effects of these shocks and thus allow an assessment of the relative importance of each shock over time.[20]

3.4.1 Results for the Broad Commodity Price Index

The impulse responses for a bivariate system comprised of the IMF overall index and consumer prices are shown in figure 3.3. The point estimates of each impulse response are given by the solid line, with the 68% and 95% bootstrapped error-bands drawn as light and heavy dashed lines, respectively.[21]

two are not very dissimilar from those presented below. As might be expected in the absence of many offsetting individual commodity shocks, commodity-specific real shocks account for a much greater fraction of the movements in these sub-indices than in the other above.

[19]The estimation procedure produces structural shocks which are serially uncorrelated. The coefficients of the moving average representation reflect the degree of permanent or transitory movements of the price series due to the underlying shocks.

[20]The first eight lags of the VMA representation are used to construct the historical decompositions.

[21]See Kilian (1998) for a discussion of these bootstrapped error bands.

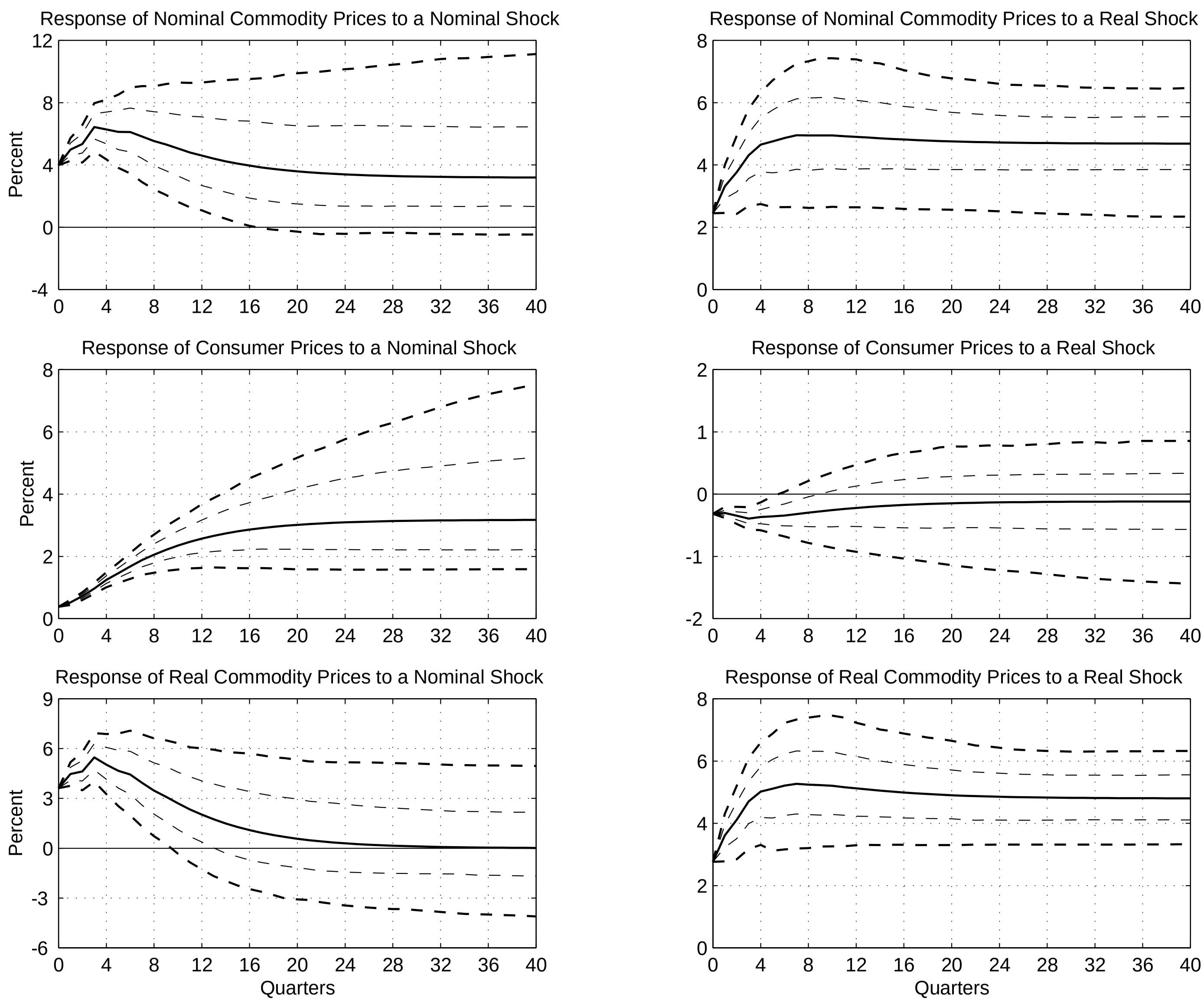


Figure 3.3: Impulse Responses of Consumer and Commodity Prices(All Items Index), 1959Q1 – 1998Q2

The left column of figure 3.3 illustrates the effect of a nominal shock. Not surprisingly, both nominal price series (consumer and commodity prices) increase in response. Consistent with their relative sluggishness, consumer prices react only mildly in the period the nominal shock occurs (rising less than one-half of one percent) then continue to rise for some time. Not until three to four years after the initial impulse is the full effect of this shock felt. Ultimately the consumer price level increases by 3.2%.

By contrast, nominal commodity prices immediately jump 4% in response to a nominal shock. Moreover, they continue to rise slightly for about three more quarters before declining to their long-run increase of 3.2%. The identifying restriction ensures that both nominal price series move together in the long-run; that is, the real commodity price eventually returns to its initial level following a nominal shock (as can be seen in the bottom row of figure 3.3). This "delayed overshooting" of real commodity prices is puzzling in the context of a forward-looking, rational-expectations view of the world. Yet numerous studies have uncovered similar empirical irregularities in other real asset prices, such as the real exchange rate. The error bands on the impulse response function confirm that this overshooting pattern is a statistically significant property of the data. Notice that after about four years the effect of the nominal shock on nominal commodity prices is not *statistically* discernible (at the 95% level) at long horizons.

Given the small initial response of consumer prices to a nominal shock, real commodity prices track nominal commodity prices fairly closely over the first year or so. Following a nominal shock, real commodity prices jump up significantly (by 3.6%) and continue to rise, peaking after three quarters (at 5.4%), then reverting back to zero as the identifying restriction requires. (The point estimate reaches zero at eight years; the response is statistically indistinguishable from zero after 10 to 14 quarters.)

Turning to the right column of figure 3.3, an impulse to the real (commodity-specific) component does not affect consumer prices in an economically significant manner. Initially the point estimate falls about 0.4 percent; the long-run response (achieved within 2 years) is not statistically different from zero.

In response to a real shock, nominal commodity prices initially jump about 2.5%. They continue to rise for slightly more than a year before reaching their long-run level, 4.7% above their initial pre-shock value. Although the point estimate exhibits some minor "overshooting," these dynamics are not statistically discernible. Indeed, the pattern of the error bands suggest that nominal commodity prices react immediately to real shocks with minimal transitory fluctuations — unlike their response to nominal shocks. Since the real shock has almost no effect on the consumer price index, the response of the real commodity price index mirrors that of the nominal index. In response to real shocks, commodity prices arguably behave much like a random walk.

The forecast variance decompositions of table 3.5 indicate that real shocks do impart some high-frequency movement to consumer prices. Real shocks are responsible for nearly 37% of the forecast variance of consumer prices initially, and over 12% at a horizon of four quarters. (By 16 quarters, less than 2% of the forecast variance is attributable to the real innovation.) Asymptotically, the real innovation accounts for about 14 percent of the forecast error variance for inflation, but effectively none of the forecast error variance for the price level. Consistent with other results reported here, nominal shocks are responsible for a majority of the consumer price fluctuations and virtually all the movement in the long-run; real shocks impart only some

Table 3.5: Variance Decompositions for Consumer and Commodity Price Indices

Percent, 1959:1 – 1998:2 sample

Forecast Horizon	Consumer Prices		Nominal IMF Overall Index		Real IMF Overall Index	
(*Quarters*)	Nominal	Real	Nominal	Real	Nominal	Real
0	63.32	36.68	68.85	31.15	58.92	41.08
1	72.19	27.81	66.70	33.30	57.24	42.76
2	78.31	21.69	64.91	35.09	54.74	45.26
3	82.99	17.01	64.90	35.10	54.12	45.88
4	87.59	12.41	63.68	36.32	51.98	48.02
6	92.75	7.25	61.58	38.42	47.96	52.04
8	95.50	4.50	59.45	40.55	43.93	56.07
12	97.86	2.14	55.65	44.38	36.92	63.08
16	98.75	1.25	52.46	47.54	31.48	68.52
20	99.16	0.84	49.84	50.16	27.36	72.64
32	99.60	0.40	44.47	55.53	19.69	80.31
40	99.70	0.30	42.21	57.79	16.62	83.38

high-frequency variation.

Meanwhile, the nominal shock is even more significant for variation in the real commodity price index. Initially, nearly 60% of the forecast error variance is due to the nominal shock. After four quarters, more than half (52%) still can be attributed to the nominal shock, and better than one-third after three years. Even ten years out, one-sixth of the forecast error variance of real commodity prices is due to the nominal innovation. Of course, the long-run restriction implies that the nominal contribution to the forecast error variance approaches zero as the horizon increases. However, this restriction does not impose any particular pattern upon the forecast error variance for consumer prices.[22]

Figures 3.4 and 3.5 show the estimated historical decompositions for consumer and real commodity prices, respectively. Consistent with the above discussion, almost all of the movement in (detrended) consumer prices over the sample period is due to nominal price pressures, rather than the real component. According to these results, inflation ultimately is a monetary phenomenon. Even more striking is the historical decomposition for real commodity prices. Although the real component now plays a more sizable role in the path of this price index over the sample, the nominal component is still the dominant factor underlying the real commodity price movements. In particular, the periods preceding the two oil price shocks (1972–1974 and 1978–1980) have very sizable nominal price components.

3.4.2 Results for IMF Commodity Groups

It is interesting to see which of the above results hold with more disaggregated commodity indices, and how these findings differ from those for the aggregate in-

[22]That is, the data — and not the restriction — cause the nominal shock to account for 100% of the long-run forecast error variance of the consumer price level.

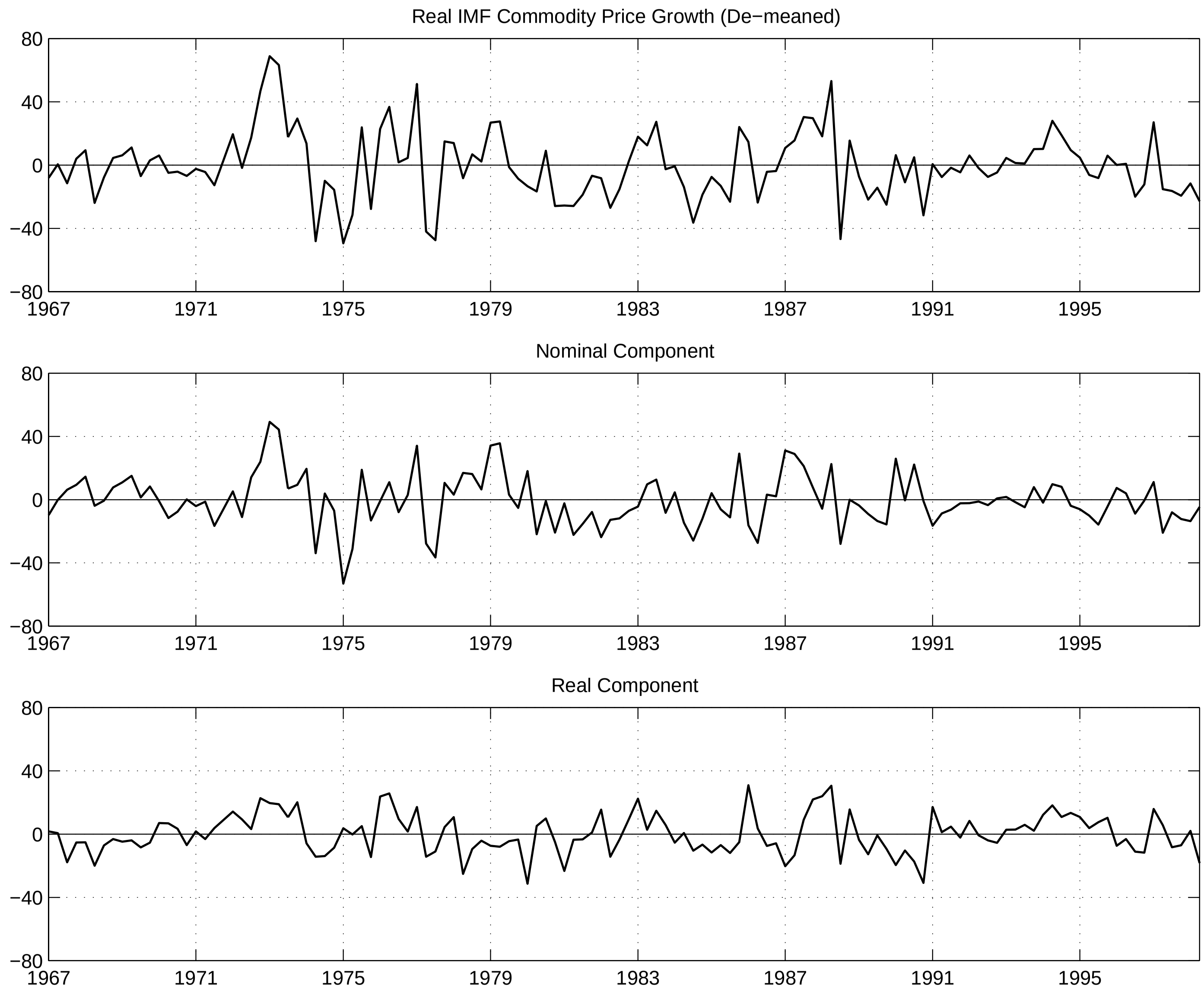


Figure 3.4: Historical Decomposition of Real Commodity Price Growth(All Items Index), 1967Q1 − 1998Q2

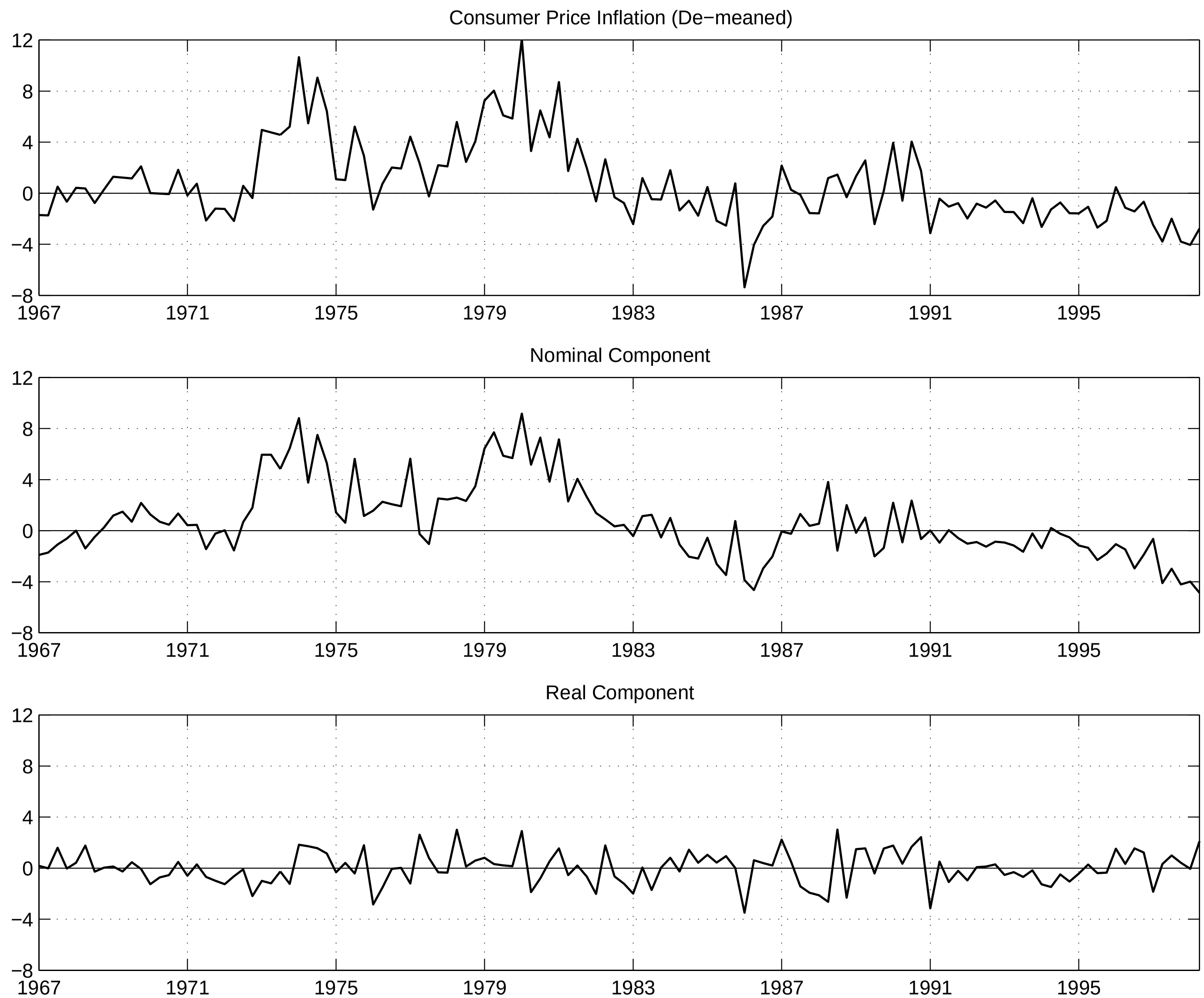


Figure 3.5: Historical Decomposition of Consumer Price Inflation,1967Q1 − 1998Q2

dex.[23] Parallel with the previous sub-section, the analysis below is based on a bivariate system with one disaggregated commodity price index and the CPI index used previously.

In terms of their share in the overall IMF commodity price index, the three largest sub-indices are foodstuffs, agricultural raw materials, and non-precious metals, respectively. These three comprise nearly 92% of the IMF index studied above. These more narrowly defined sub-indices are less likely to be subjected to unrelated commodity-specific real shocks that tend to offset each other. The broader the index the more likely idiosyncratic shocks will average out, yielding information primarily about the nominal component. Thus, *a priori* one should expect more variation in these sub-indices to be attributed to the real shock. Second, one might expect these idiosyncratic shocks to be primarily transitory in nature for agricultural products, which should imply more of the fluctuations could be attributed to the real shock. To the extent that agricultural goods are less storable than, say, primary non-precious metals, they should also exhibit less overshooting behavior.

Despite these issues, the impulse responses appear quite similar across the various sub-indices and resemble those for the overall index. Figure 3.6 plots the impulse responses for foodstuffs, figure 3.7 for agricultural raw materials, and figure 3.8 for (non-precious) metals. In each case, the response of the commodity price index to a nominal shock exhibits some limited evidence of delayed overshooting. The magnitude of the response is somewhat larger for metals, but this index does not exhibit more substantial overshooting than the others. For all the indexes studied, the effect of the nominal shock on real commodity prices persists for two to three years before it is statistically indistinguishable from zero. In each case, the sluggish

[23] This technique can be applied to further levels of disaggregation; subsequent research may examine individual commodities.

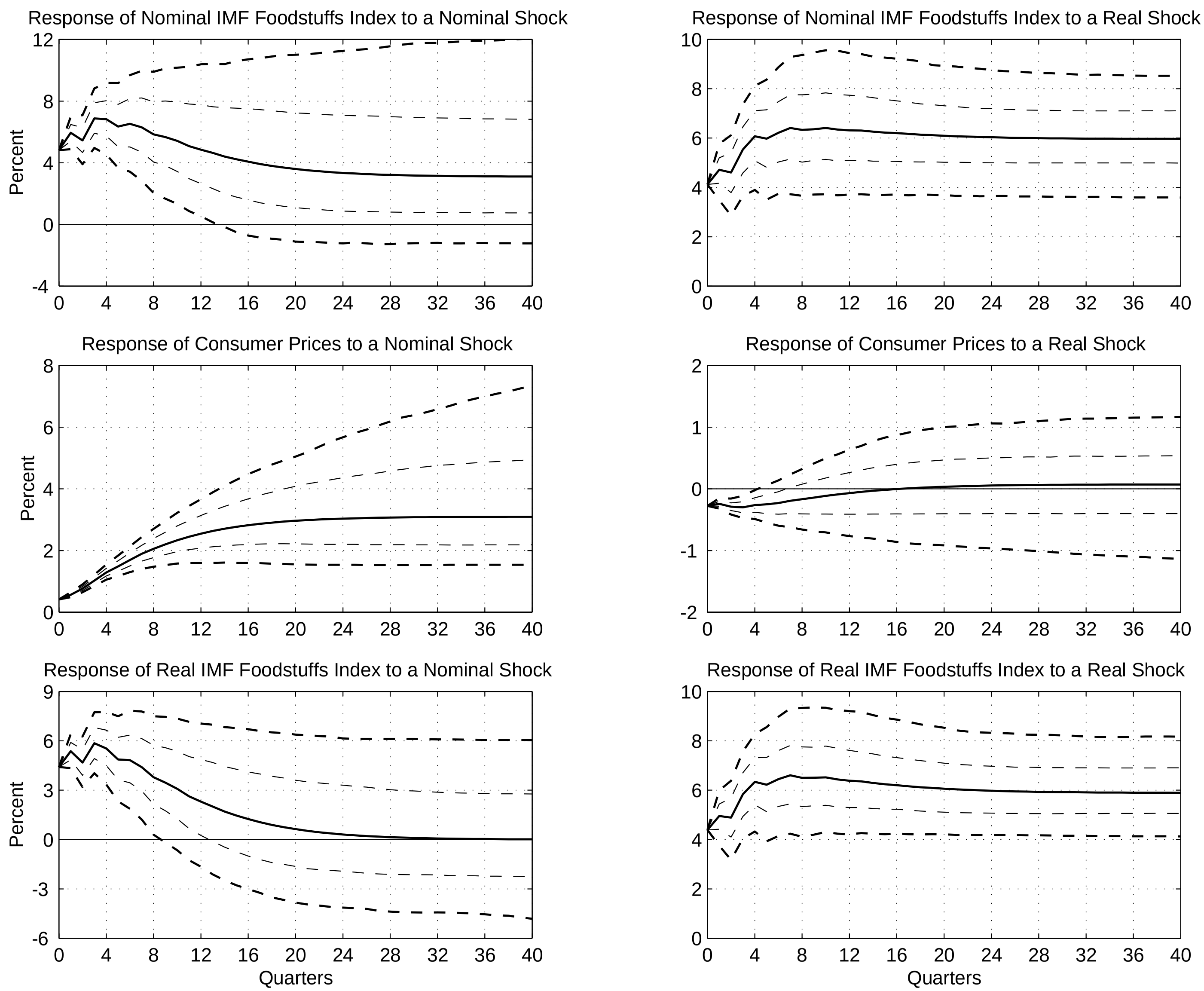


Figure 3.6: Impulse Responses of Consumer and Commodity Prices(Foodstuffs Index), 1959Q1 – 1998Q2

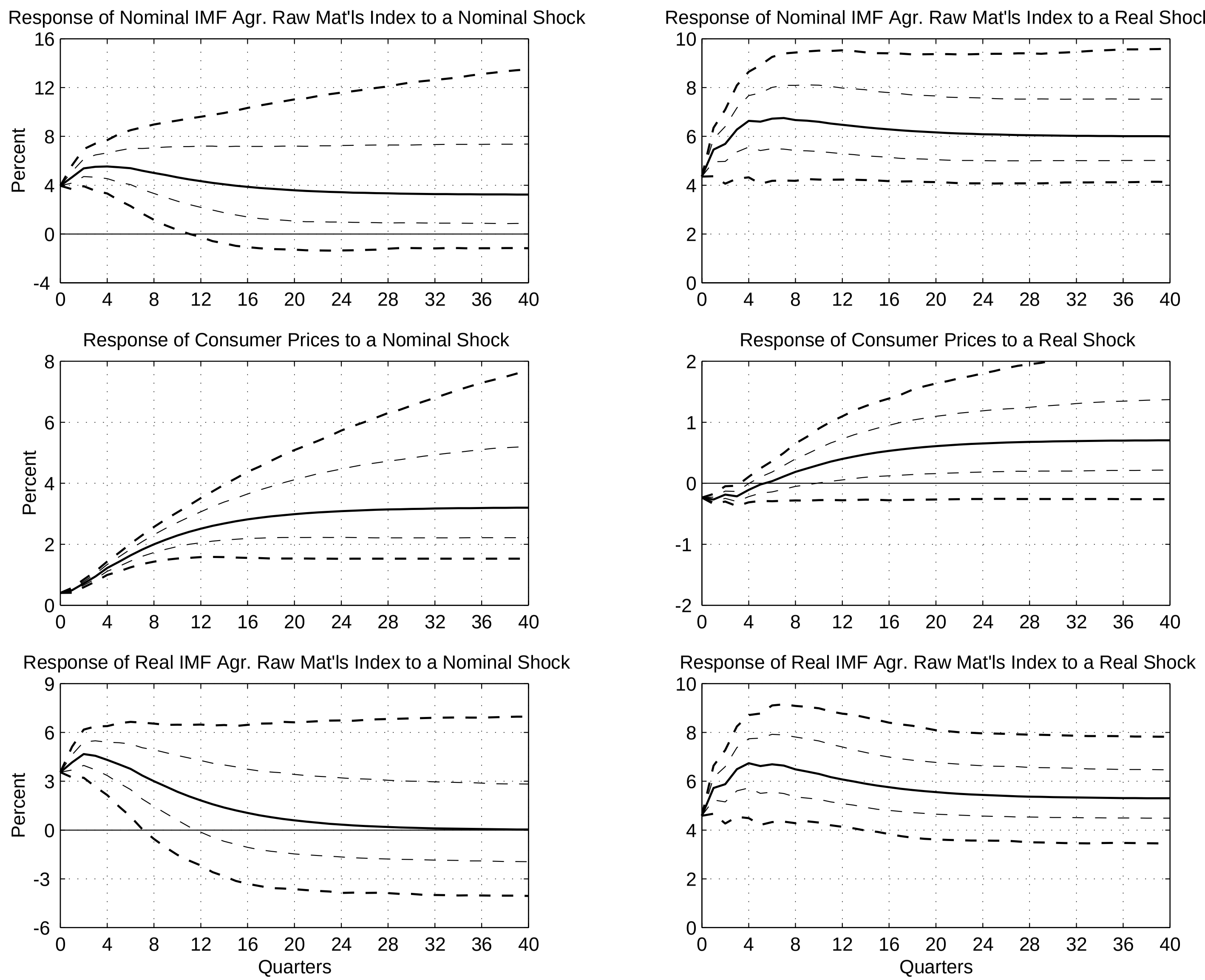


Figure 3.7: Impulse Responses of Consumer and Commodity Prices (Agricultural Raw Materials Index), 1959Q1 – 1998Q2

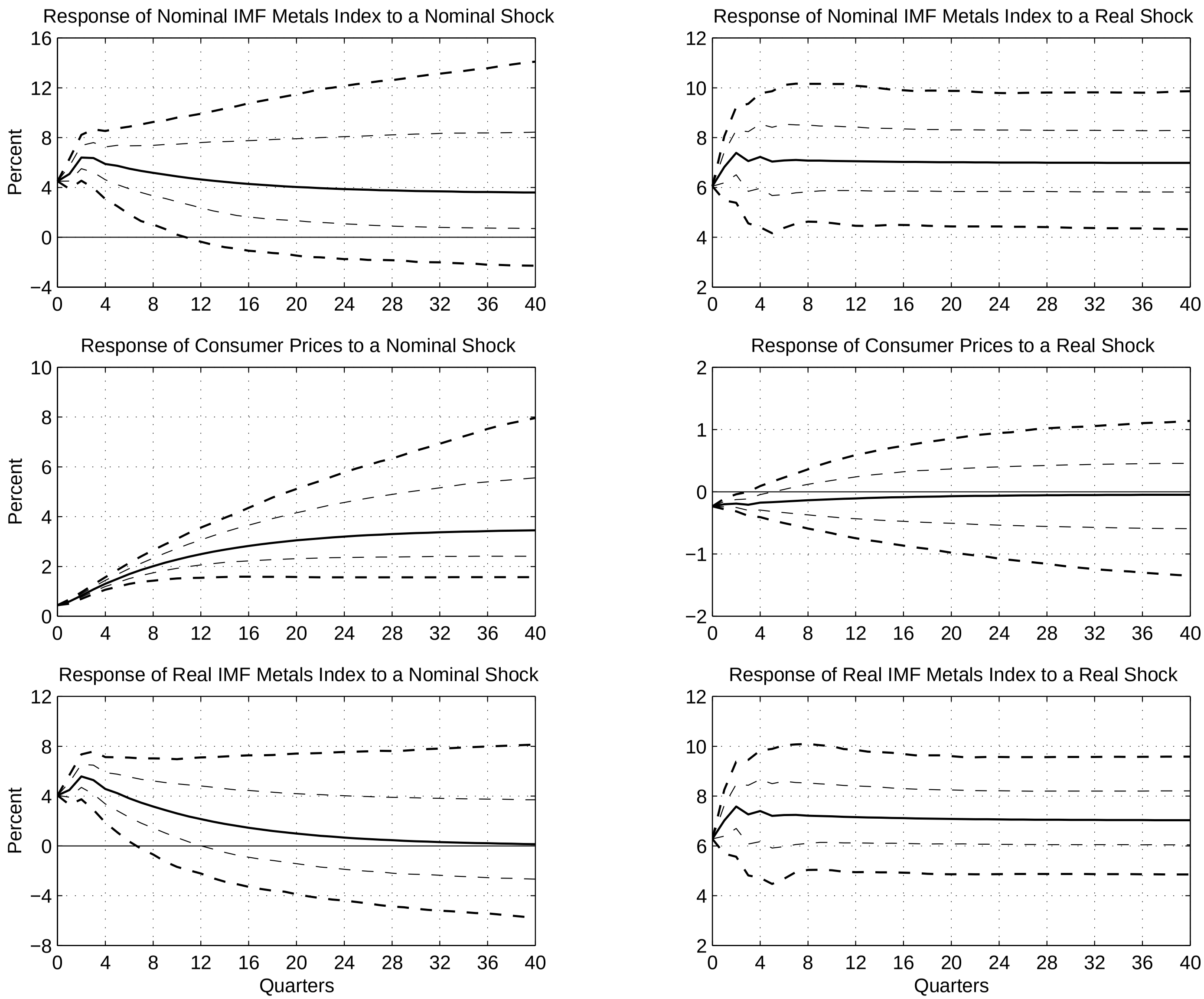


Figure 3.8: Impulse Responses of Consumer and Commodity Prices(Metals Index), 1959Q1 – 1998Q2

Table 3.6: Nominal Contribution to Variance Decompositions for Various IMF Sub-Indices

Percent, 1959:1 – 1998:2 sample

Forecast Horizon (*Quarters*)	Overall	Foodstuffs	Agricultural Raw Mat'ls	Metals
0	58.92	45.39	39.71	29.36
1	57.24	47.52	37.67	29.17
2	54.74	45.81	38.56	31.64
3	54.12	45.47	36.98	32.44
4	51.98	43.73	35.24	31.46
6	47.96	40.19	32.34	29.39
8	43.93	36.64	29.48	27.00
12	36.92	30.50	24.72	22.70
16	31.48	25.76	21.15	19.26
20	27.36	22.23	18.48	16.59
32	19.69	15.82	13.50	11.52
40	16.62	13.30	11.48	9.54

response of consumer prices to the nominal shock is virtually the same at all horizons as well.

In figure 3.8, the real shock does seem to have a slightly more noticeable transitory component for the metals index than for either foodstuffs or agricultural raw materials, in that it reaches its long-run response more quickly, but this effect is swamped by the estimated error bands. In all three cases the response of consumer prices to a real shock is not statistically (nor economically) discernible from zero.

Table 3.6 reveals that a larger fraction of the forecast error variance is due to the real innovation for the sub-indices than for the overall index, as hypothesized above. Interestingly, the contribution of the nominal shock is lower at all horizons for metals than for either agricultural index. In conjunction with the impulse response analysis above, this finding suggests that metals prices are buffeted more frequently by permanent, real (i.e. commodity-specific) shocks than are the other indices. When the various idiosyncratic shocks are allowed to offset one another, as in the overall index, the contribution of the nominal component is measurably higher.

Much like the impulse response functions, the historical decompositions for each sub-index exhibited behavior nearly identical behavior to the overall index, and so in the interest of space are not included here. These historical decompositions are consistent with the conclusion that the nominal component is responsible for a large fraction of the changes in real commodity prices and virtually all of the movement in consumer prices.

3.4.3 Summary of Results

The results for the various sub-indices reported above generally confirm the relative importance of the nominal shock in understanding the behavior of consumer prices and real commodity prices over the sample in question. Three related metrics — impulse responses, forecast error variance decompositions, and historical decompositions — yield broadly similar findings across the various indices.

The impulse responses indicate that, on average, a positive nominal shock is responsible for a significant increase in real commodity prices that lasts for two to three years. Although the long-run assumption for identification imposes that *eventually* the nominal shock will have no effect on real commodity prices, it does not impose any restrictions on the specific length of the response.[24] The nominal shock also results in a permanent increase in consumer prices; these prices increase slowly and reach their final peak after about three to four years. Such results are consistent

[24] The point estimates are indistinguishable from zero at four significant digits between eight and ten years after the shocks.

with the view that these prices are set in slowly adjusting "customer markets." Note that the identification scheme imposes no restrictions on the shape of the impulse response function for consumer prices.[25] Interestingly, the length of time for which significant dynamics are observed for these two series is consistent with the broad consensus in the literature on the time frame for the effects of monetary policy shocks on most macroeconomic variables.

On the other hand, a real (commodity-specific) shock seems to have no significant economic impact upon consumer prices in the various systems estimated here. The real and nominal commodity price indices respond fairly quickly to the real shock, although metals have a noticeably shorter period of transitory fluctuations; their response most closely resembles a random walk.

For the overall IMF index, the variance decompositions reveal that a large fraction of the variation in real commodity prices is due to the nominal shock at short horizons. Indeed, it takes over a year for the real shock to account for a greater share of the forecast variance of the real commodity price indices than the nominal shock. This is not so with the narrower indices: only for foodstuffs is anything close to one-half of the forecast variance due to the nominal shock. That these sub-indices show at least as much sensitivity to real shocks as the overall indices should not be surprizing, since the idiosyncratic shocks across different commodities are more likely to cancel out each other for the broader index.

According to the historical decompositions, the nominal component appears to be the dominant source of movements in the real commodity prices analyzed here. The real component is not unimportant, as it imparts high frequency fluctuations to commodity prices. However, as indicated by the impulse responses, the real compo-

[25]Nor, for that matter, does it impose any restrictions on the shape of the IRF inherited by the transition dynamics for the real commodity price. The only restriction is at frequency zero.

nent generally contributes very little to the movements in consumer prices over the sample period.

3.5 Conclusions

The underlying motivation of this research begins with the fact that as assets, commodity prices are sensitive to the policy stance of the monetary authority. Despite the large fluctuations observed for the commodity price series (the standard deviation is about six times as large as that for consumer prices), I am nonetheless able to identify a significant portion of commodity price volatility with nominal shocks to the economy. This nominal component accounts for most of the movements in consumer prices as well. A positive impulse to this nominal shock, as identified herein, results in significant increases in real commodity prices which persist for several years. Moreover, much of the forecast error variance of real commodity prices — for horizons up to several years — is attributable to the nominal shock. Historical decompositions demonstrate the relative importance of the nominal component for both consumer and (real) commodity prices, particularly preceding periods commonly associated with the two major oil price “shocks” of the 1970s. These results are robust to different indices measuring commodity prices, and to several sub-indices by commodity type.

The main advantage of this approach is that it avoids placing “incredible” restrictions upon the model. An equiproportional long-run response of nominal prices to a nominal (monetary) shock is well established in neo-classical macroeconomics, and is an even weaker assumption than the classical dichotomy. It also follows directly from the theoretical models cited in the introduction.

One could argue that the effect of monetary policy has been only indirectly identified by this study. Yet this is a benefit of the approach, in that it avoids the need to specify a reaction function (or money supply function) for the monetary authority. In particular, I can infer the effects of monetary policy without imposing a particular structure on the nature of monetary policy making, or modeling a particular channel for monetary policy. Several authors (for example, Bernanke and Mihov, 1995) have noted that the nature of monetary policy has varied over my sample period. Moreover, I do not arbitrarily "order" one variable causally prior to another, but instead allow both to react within the quarter to both types of shocks. It seems unlikely that either of these series would be unresponsive to either type of shock for an entire quarter.

The results also are consistent with other studies in this literature that emphasize the contribution of monetary shocks. In particular they lend credence to the findings of Bernanke et al. (1997), who argue that monetary accommodation was a main portion of the observed inflation following the oil price shocks. Whereas the Bernanke et al. (1997) paper utilized only contemporaneous restrictions, these results suggest it may be fruitful to mix short- and long-run restrictions in a a more detailed empirical macro-model of the economy which includes, amongst other things, measures of real activity and interest rates.[26]

The main disadvantage of my approach is the possible convolution of shocks. This problem arises with any VAR analysis, but it is especially acute in a bivariate model: at most only two independent structural innovations can be identified. I have interpreted the shock that has only temporary effects on real commodity prices as a nominal, and hence monetary, innovation. Since the identifying restriction of the

[26]A second extension would be to examine individual commodity prices for additional information on the relative importance of nominal disturbances, as well as re-examine the "excess co-movement" in commodity prices that other authors (Pindyck and Rotemberg, 1990) have identified.

SVAR analysis is consistent with long-run monetary neutrality, this interpretation is the most natural. However, it is possible that some of the structural disturbance which I claim to be monetary in nature may in fact be due to other factors. Any structural shock which has only a temporary effect on real commodity prices — i.e. any shock that moves nominal commodity prices and nominal consumer prices proportionally — will be reflected in my "monetary" disturbance. It is challenging to suggest any systematic non-monetary source of shocks that would satisfy this condition, but it could happen. To some extent, then, these results may suggest an upper-bound for the contribution of nominal shocks to the variability of commodity prices. Nonetheless, they are likely non-trivial.[27]

Additionally, while some non-monetary interpretations of the shock cannot be completely excluded, it is important to recognize that the real innovation cannot account for any of what I have called the nominal shock. The estimation procedure places no restrictions on the real shock; in particular it allows for the real shock to have a temporary component. I have not restricted the real shock to have only permanent effects, but have only restricted the nominal shock not to have any permanent effect on real commodity prices.

[27]See Faust and Leeper (1997) for further discussion of this issue.

APPENDIX

APPENDIX

Data Description

The International Monetary Fund index is a weighted average of non-fuel commodities, with the weights based on the 1987–89 average world export earnings for 175 countries. Although this index incorporates a larger number of commodities than some other commonly examined commodity price indices, it is composed primarily of agricultural products and non-precious metals. Supply shocks from weather or pestilence may add noise to the agricultural components of this series and therefore increase the overall variance of the index, potentially reducing its effectiveness for forecasting inflation. Moreover, several commodity prices are quoted in oversees markets, which may render the index less useful as a potential indicator of future U.S. inflation. However, active daily trading in the markets for these goods suggests that this index may serve my purposes well. At issue is the correlation of the commodity-specific shocks across those goods in the index. Given the different growing seasons and areas of the world from which these commodities are shipped, this issue may not be as severe as it initially seems, although the additional noise may mask the inflationary expectations signal.

The table below lists the individual commodities that comprise the sub-indices examined in section 3.4.2.

International Monetary Fund Index (33 commodities):

Foodstuffs: bananas, beef, coconut oil, fish meal, groundnut oil, lamb, maize, palm oil, rice, soybeans, soybean meal, soybean oil, sugar, wheat

Beverages: cocoa beans, coffee, tea

Raw Agricultural Materials: cotton, hides, rubber, timber, tobacco, wool

Metals: aluminum, copper, iron ore, lead, nickel, tin, zinc

Fertilizers: phosphate rock, triple phosphate

BIBLIOGRAPHY

Bibliography

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

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

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

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

_ and Ilian Mihov, "Measuring Monetary Policy," *National Bureau of Economic Research Working Paper 5145*, June 1995.

_ and _ , "Measuring Monetary Policy," *Quarterly Journal of Economics*, 1998, *113* (3), 869–902.

_ , Mark Gertler, and Mark Watson, "Systematic Monetary Policy and the Effects of Oil Price Shocks," in William C. Brainard and George L. Perry, eds., *Brookings Papers on Economic Activity*, Vol. 1 1997, pp. 91–157.

Blanchard, Olivier and Danny Quah, "The Dynamic Effects of Aggregate Supply and Demand Disturbances," *American Economic Review*, 1989, *79* (4), 655–73.

Blanchard, Olivier J., "The Lucas Critique and the Volcker Deflation," *American Economic Review, Papers and Proceedings*, May 1984, *74* (2), 211–215.

Blinder, Alan S., "On Sticky Prices: Academic Theories Meet the Real World," in N. Gregory Mankiw, ed., *Monetary Policy*, Chicago: University of Chicago Press, 1994, pp. 117–150.

Blomberg, S. Brock and Ethan S. Harris, "The Commodity-Consumer Price Connection: Fact or Fable?," *Federal Reserve Bank of New York Economic Policy Review*, October 1995, pp. 21–38.

Boughton, James M. and William H. Branson, "Commodity Prices as a Leading Indicator of Inflation," in Kajal Lahiri and Geoffrey H. Moore, eds., *Leading Economic Indicators*, New York: Cambridge University Press, 1991, pp. 305–338.

Brainard, William C., "Uncertainty and the Effectiveness of Policy," *American Economic Review, Papers and Proceedings*, May 1967, *57* (2), 411–425.

Brunner, Allan, "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 1994.

Bryant, Ralph C., Peter Hooper, and Catherine L. Mann, eds, *Evaluating Policy Regimes: New Research in Empirical Macroeconomics*, Washington, DC: The Brookings Institute, 1993.

Cecchetti, Stephen, "Inflation Indicators and Inflation Policy," in B. Bernanke and J. Rotemberg, eds., *NBER Macroeconomics Annual 1995*, MIT Press Cambridge, MA 1995, pp. 189–219.

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

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

_ , _ , and _ , "Monetary Policy Shocks: What Have We Learned and to What End?," *National Bureau of Economic Research Working Paper 6400*, February 1998.

Clarida, Richard and Jordi Galí, "Sources of Real Exchange-Rate Fluctuations: How Important are Nominal Shocks?," *Carnegie-Rochester Conference Series on Public Policy*, 1994, *41*, 1–56.

_ , Jordi Galí and Mark Gertler, "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *National Bureau of Economic Research Working Paper 6442*, March 1998.

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

_ , "The Responses of Prices at Different Stages of Production to Monetary Policy Shocks," *Review of Economics and Statistics*, August 1999, *81* (3), 420–433.

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

—, "What Do the VARs Mean? Measuring the Output Effects of Monetary Policy," *Journal of Monetary Economics*, April 1998, *41* (2), 277–300.

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

De Arcangelis, Giuseppe, "The Influence of Real and Nominal Shocks on the Exchange Rates," *Journal of International and Comparative Economics*, 1993, *2*, 247–277.

De Long, J. Bradford, "America's Peacetime Inflation: The 1970's," in Christina D. Romer and David H. Romer, eds., *Reducing Inflation: Motivation and Strategy*, Chicago: University of Chicago Press, 1997, pp. 247–276.

Deaton, Angus and Guy Laroque, "On the Behaviour of Commodity Prices," *Review of Economic Studies*, 1992, *59* (1), 1–23.

Dornbusch, Rudiger, "Expectations and Exchange Rate Dynamics," *Journal of Political Economy*, 1976, *84* (6), 1161–1176.

Eichenbaum, Martin, "Comment on 'Interpreting the Macroeconomic Time Series Facts: The Effects of Monetary Policy.' ," *European Economic Review*, 1992, *36* (5), 1001–1011.

Enders, Walter and Bong-Soo Lee, "Accounting for Real and Nominal Exchange Rate Movements in the Post-Bretton Woods Period," *Journal of International Money and Finance*, 1997, *16* (2), 233–254.

Ericsson, Neil and John Irons, "The Lucas Critique in Practice: Theory without Measurement," in Kevin D. Hoover, ed., *Macroeconometrics: Developments, Tensions and Prospects*, Boston: Kluwer Academic Publishers, 1993, pp. 263–312.

Fair, Ray C., "Actual Fed Behavior and Interest Rate Rules," December 1999. Yale University.

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

_ and Eric Leeper, "When Do Long-Run Identifying Restrictions Give Reliable Results?," *Journal of Business and Economic Statistics*, 1997, *15* (3), 345–353.

Fischer, Stanley, "Relative Shocks, Relative Price Variability, and Inflation," in "Brookings Papers on Economic Activity," Vol. 2 The Brookings Institute Washington, DC 1981, pp. 381–431.

Frankel, Jeffery A., "Expectations and Commodity Price Dynamics: The Overshooting Model," *American Journal of Agricultural Economics*, 1986, *68* (2), 344–348.

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

Furlong, Fred and Robert Ingenito, "Commodity Prices and Inflation," *Federal Reserve Bank of San Fransisco Economic Review*, 1996, (2), 27–47.

Galí, Jordi, "How Well Does the IS-LM Model Fit Post-War U.S. Data?," *Quarterly Journal of Economics*, May 1992, *92* (2), 709–738.

Garner, C. Alan, "How Useful are Leading Indicators of Inflation?," *Federal Reserve Bank of Kansas City Economic Review*, 1995, *Second Quarter*, 5–18.

Gerlach, Stefan and Frank Smets, "The Monetary Transmission Mechanism: Evidence from the G-7 Countries," *Centre for Economic Policy Research Discussion Paper 1219*, February 1995.

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

Hall, Robert, "Comment on Leeper, Sims and Zha, "What Does Monetary Policy Do?"," in William C. Brainard and George L. Perry, eds., *Brookings Papers on Economic Activity*, Vol. 2 1996, pp. 64–68.

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

Hanson, Michael S., "On the Identification of Monetary Policy: the 'Price Puzzle' Reconsidered," November 1998. University of Michigan.

_ , "Varying Monetary Policy Rules: a Vector Autoregressive Investigation," November 2000. Wesleyan University.

Heckman, James L., "Causal Parameters and Policy Analysis in Economics: A Twentieth Century Retrospective," *National Bureau of Economic Research Working Paper 7333*, September 1999.

Hook, Alison and David Walton, "Commodity Price Indicators: A Case for Discretion Rather than Rules," in Barry Eichengreen Marcus Miller and Richard Portes, eds., *Blueprints for Exchange Rate Management*, San Diego, CA: Academic Press, Inc., 1989, pp. 99–119.

Horvath, Michael T. K. and Mark W. Watson, "Testing for Cointegration When Some of the Cointegrating Vectors are Prespecified," *Econometric Theory*, 1995, *11* (5), 984–1014.

III, Marvin J. Barth and Valerie A. Ramey, "The Cost Channel of Monetary Transmission," August 1999. University of California, San Diego.

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

Judd, John P. and Glenn D. Rudebusch, "Taylor's Rule and the Fed: 1970–1997," *Federal Reserve Bank of San Fransisco Economic Review*, 1998, (3), 3–16.

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

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

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

Kozicki, Sharon, "Predicting Real Growth and Inflation with the Yield Spread," *Federal Reserve Bank of Kansas City Economic Review*, 4th quarter 1997, pp. 39–57.

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

Lastrapes, William D., "Sources of Fluctuations in Real and Nominal Exchange Rates," *Review of Economics and Statistics*, 1992, *74* (3), 530–539.

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

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

_ **and Tao Zha**, "Modest Policy Interventions," November 1999. Presented at the January 2000 AEA Meetings.

_ **, Christopher A. Sims, and Tao Zha**, "What Does Monetary Policy Do?," in William C. Brainard and George L. Perry, eds., *Brookings Papers on Economic Activity*, Vol. 2 The Brookings Institute Washington, DC 1996, pp. 1–63.

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

Lucas, Jr, Robert E., "Econometric Policy Evaluation: A Critique," in Karl Brunner and Allan H. Meltzer, eds., *The Philips Curve and Labor Markets, Carnegie-Rochester Conference Series on Public Policy*, Vol. 1 1976, pp. 19–46. Supplementary issue.

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

McCallum, Bennett T., "Analysis of the Monetary Transmission Mechanism: Methodological Issues," *National Bureau of Economic Research Working Paper 7395*, October 1999.

Meulendyke, Ann-Marie, *U.S. Monetary Policy and Financial Markets*, Federal Reserve Bank of New York, 1998.

Mussa, Michael, "Sticky Prices and Disequilibrium Adjustment in a Rational Model of the Inflationary Process," *American Economic Review*, 1981, *71* (5), 1020–1027.

_ , "A Model of Exchange Rate Dynamics," *Journal of Political Economy*, 1982, *90* (1), 74–104.

Ohanian, Lee E. and Alan C. Stockman, "Short–Run Effects of Money When Some Prices Are Sticky," *Federal Reserve Bank of Richmond Economic Quarterly*, 1994, *80* (3), 1–23.

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

Phillips, Peter C. B. and Sam Ouliaris, "Asymptotic Properties of Residual Based Tests for Cointegration," *Econometrica*, 1990, *58* (1), 165–193.

Pindyck, Robert S. and Julio J. Rotemberg, "The Excess Co-Movement of Commodity Prices," *Economic Journal*, 1990, *100* (403), 1173–1189.

Rogers, John H., "Monetary Shocks and Real Exchange Rates," International Finance Discussion Paper no. 612, Board of Governors of the Federal Reserve System May 1998.

Rotemberg, Julio J., "The New Keynesian Microfoundations," in Olivier Jean Blanchard and Stanley Fischer, eds., *NBER Macroeconomics Annual 1987*, MIT Press Cambridge, MA 1987, pp. 69–104.

Rudebusch, Glenn D., "Do Measures of Monetary Policy in a VAR Make Sense?," *International Economic Review*, November 1998, *39* (4), 907–931.

_ , "Do Measures of Monetary Policy in a VAR Make Sense? A Reply to Christopher A. Sims," *International Economic Review*, November 1998, *39* (4), 943–947.

Sargent, Thomas, "The Conquest of American Inflation," 1998. Manuscript: Hoover Institution.

Shapiro, Matthew and Mark Watson, "Sources of Business Cycle Fluctuations," in Olivier Jean Blanchard and Stanley Fischer, eds., *NBER Macroeconomics Annual 1988*, MIT Press Cambridge, MA 1988, pp. 111–48.

Sims, Christopher A., "Macroeconomics and Reality," *Econometrica*, January 1980, *48* (1), 1–48.

_ , "Interpreting the Macroeconomic Time Series Facts: The Effects of Monetary Policy," *European Economic Review*, 1992, *36* (5), 975–1000.

_ , "Comment on Bernanke et al. 'Systematic Monetary Policy and the Effects of Oil Price Shocks'," in William C. Brainard and George L. Perry, eds., *Brookings Papers on Economic Activity*, Vol. 1 1997, pp. 143–148.

_ , "Comment on Glenn Rudebusch's 'Do Measures of Monetary Policy in a VAR Make Sense?'," *International Economic Review*, November 1998, *39* (4), 933–941.

_ , "The Role of Interest Rate Policy in the Generation and Propagation of Business Cycles: What Has Changed since the '30s?," in Jeffrey C. Fuhrer and Scott Schuh, eds., *Beyond Shocks: What Causes Business Cycles?*, number 42 Federal Reserve Bank of Boston Conference Proceedings June 1998, pp. 121–160.

_ , "Drifts and Breaks in Monetary Policy," July 1999. Yale University.

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

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

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

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

—, "An Historical Analysis of Monetary Policy Rules," December 1997. Stanford University.

Uhlig, Harald, "What are the Effects of Monetary Policy on Output? Results from an Agnostic Identification Procedure," January 1997. CentER, Tilburg University.

Webb, Roy H. and Tazewell S. Rowe, "An Index of Leading Indicators of Inflation," *Federal Reserve Bank of Richmond Economic Quarterly*, 1995, *81* (2), 75–96.

Williams, Jeffery and Brian Wright, *Storage and Commodity Markets*, Cambridge, England: Cambridge University Price, 1991.

Woodford, Michael, "Nonstandard Indicators for Monetary Policy: Can Their Usefulness Be Judged from Forecasting Regressions?," in N. Gregory Mankiw, ed., *Monetary Policy*, Chicago, IL: The University of Chicago Press, 1994, pp. 95–115.

Zha, Tao, "Identifying Monetary Policy: A Primer," *Federal Reserve Bank of Atlanta Economic Review*, 2nd quarter 1997, *82* (2), 26–43.

ABSTRACT

VECTOR AUTOREGRESSIVE MEASURES OF MONETARY POLICY: ISSUES AND CRITIQUES

by

Michael S. Hanson

Chairs: Robert B. Barsky and Matthew D. Shapiro

This dissertation explores the vector autoregressive methodology, arguably the predominant paradigm in modern monetary economics, through three distinct but related essays on the empirical measurement of monetary policy. These essays investigate the measured role of Federal Reserve policy in the post-war United States, and question several results cited in the current academic literature.

The first examines the "price puzzle:" a rise in the aggregate price level in response to a contractionary innovation to monetary policy, common in models that do not include commodity prices. Conventional wisdom maintains that commodity prices resolve the price puzzle because they contain information that helps the Federal Reserve forecast inflation. However, an investigation of several dozen plausible alternative indicator variables finds 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. One implication of these results is that commodity prices are unlikely to be serving as proxies for "supply shocks" in these models.

The sub-sample results of the first chapter suggest a role for changes in the reaction function of the monetary authorities. Indeed, recent academic and popular discussions credit Federal Reserve policy for the favorable economic performance of the past fifteen years while blaming it for the stagflation of the 1970s. For example, the "Taylor rules" literature typically estimates the monetary policy rule in a context that cannot separate changes in the reaction function from changes originating elsewhere in the economy. On the other hand, the vector autoregressive literature tends to ignore the endogenous policy rule to focus upon the exogenous monetary policy shocks. Utilizing a common SVAR specification in the second essay, I find evidence of a shift in the policy reaction function around the 1979 – 1982 Volcker disinflation. However, most of the changes in the economic dynamics can be attributed to parameter shifts in the *non-policy* portion of the estimated model. Furthermore, the implied SVAR policy rule differs substantially from that reported in the "Taylor rules" literature.

The third essay returns to an examination of commodity prices, which usually are treated as an exogenous information variable for monetary policy, at least with respect to contemporaneous policy actions. Explicitly recognizing that commodity prices, as forward-looking flexible prices, react to current and expected future monetary policy, I estimate a bivariate structural VAR model of commodity price and U.S. consumer price indices. An equiproportional long-run response of nominal price levels to a monetary (i.e. nominal) shock yields identifying restrictions. The results suggest that a significant share of the co-movement of these series is due to (exogenous innovations in) monetary policy, including episodes more commonly attributed to "supply shocks."