Oil Shocks and Aggregate Macroeconomic Behavior: The Role of Monetary Policy*

June 2000
Revised: July 2001

James D. Hamilton
Department of Economics, 0508
University of California, San Diego
La Jolla, CA 92093-0508
jhamilton@ucsd.edu

Ana Maria Herrera
Michigan State University
Department of Economics
101 Marshall Hall
East Lansing, MI 48244
herrer20@msu.edu

ABSTRACT

A recent paper by Bernanke, Gertler and Watson (1997) suggests that monetary policy could be used to eliminate any recessionary consequences of an oil price shock. This paper challenges that conclusion on two grounds. First, we question whether the Federal Reserve actually has the power to implement such a policy; for example, we consider it unlikely that additional money creation would have succeeded in reducing the Fed funds rate by 900 basis points relative to the values seen in 1974. Second, we point out that the size of the effect that Bernanke, Gertler and Watson attribute to oil shocks is substantially smaller than that reported by other researchers, primarily due to their choice of a shorter lag length than used by other researchers. We offer evidence in favor of the longer lag length employed by previous research, and show that, under this specification, even the aggressive Federal Reserve policies proposed would not have succeeded in averting a downturn.

1 Introduction

A large number of studies have reported a correlation between increases in oil prices and subsequent economic downturns. Examples include Rasche and Tatom (1977, 1981), Hamilton (1983), Burbidge and Harrison (1984), Santini (1985, 1994), Gisser and Goodwin (1986), Loungani (1986), Davis (1987), Mork (1989), Lee, Ni and Ratti (1995), Davis, Loungani, and Mahidhara (1996), Keane and Prasad (1996), Rotemberg and Woodford (1996), Daniel (1997), Davis and Haltiwanger (2001), Raymond and Rich (1997), Carruth, Hooker, and Oswald (1998), Brown and Yücel (1999), Lee and Ni (1999), Balke, Brown and Yücel (1999), and Hamilton (2000).[1] Although the existence of this correlation is well established, there is substantial disagreement as to what it means.

Bohi (1989) has argued that the recessions that followed the big oil shocks were caused not by the oil shocks themselves, but rather by the Federal Reserve's contractionary response to inflationary concerns attributable in part to the oil shocks. Bernanke, Gertler, and Watson (1997) presented key evidence supporting this view, demonstrating that, if one shuts down the tendency of the Federal funds rate to rise following an oil shock, or simulates a VAR subsequent to the big oil shocks under the condition that the Federal funds rate could not rise, it appears that the economic downturns might be largely avoided.

This paper challenges that conclusion on two grounds. First, we adapt methods proposed by Sims (1982) and Leeper and Zha (1999) to evaluate the plausibility of some of the coun-

---

[1] It is clear that a linear regression to describe this relation is not stable over time, and researchers differ in their interpretation of this instability. See Mork (1989), Lee, Ni, and Ratti (1995), Hooker (1996, 1997), Hamilton (1996, 2000), and Balke, Brown and Yücel (1999) for discussion.

terfactual policy simulations employed by Bernanke, Gertler, and Watson, and conclude that both the nature and magnitude of the actions suggested for the Fed are sufficiently inconsistent with the historical correlations as to call into question the feasibility of such a policy. Second, we demonstrate that the initial effect attributed by Bernanke, Gertler, and Watson to an oil shock without any policy intervention is substantially smaller than that found by other researchers due to the authors' specification of the lag length of the VAR. We argue that the econometric evidence favors the longer lags used by other researchers, and, when these are allowed, the simulations support the conclusion that even the policies considered by Bernanke, Gertler, and Watson would not have succeeded in averting a downturn.

We first briefly summarize Bernanke, Gertler, and Watson's methodology and principal findings.

2 Review of previous results

Bernanke, Gertler, and Watson estimated a monthly structural vector autoregression describing $\mathbf{y}_t$, which contains a monthly interpolated series for the rate of growth of real GDP ($y_{GDP,t}$), the log of the GDP deflator ($y_{DEF,t}$), log of the commodity price index ($y_{COM,t}$), a measure of oil prices ($y_{OIL,t}$), the Fed funds rate ($y_{FED,t}$), the 3-month Treasury bill rate ($y_{TB3,t}$), and the 10-year Treasury bond rate ($y_{T10,t}$) of the form

$$\mathbf{B}_0 \mathbf{y}_t = \mathbf{k}_0 + \mathbf{B}_1 \mathbf{y}_{t-1} + \mathbf{B}_2 \mathbf{y}_{t-2} + ... + \mathbf{B}_p \mathbf{y}_{t-p} + \mathbf{v}_t. \tag{1}$$

The matrix $\mathbf{B}_0$ is taken to be lower triangular with ones along the principal diagonal, and lagged Fed funds rates are assumed to affect the first four variables only through their effects

on the other interest rates, so the row i, column 5 element of $\mathbf{B}_s$ is zero for $i = 1, 2, 3, 4$ and $s = 1, 2, ..., p$. The system is estimated by OLS, equation by equation. The lag length p was set to 7, the consequences of which we discuss in Section 4. All estimates reported in this paper are based on exactly the same 1965:1-1995:12 data set used by Bernanke, Gertler, and Watson.

The authors explored a variety of alternative measures for the oil shock variable $y_{OIL,t}$. They concentrated primarily on the series proposed by Hamilton (1996), which describes the amount by which the log oil price in month t exceeds its maximum value over the previous 12 months; if oil prices are lower than they have been at some point during the past year, no oil shock is said to have occurred ($y_{OIL,t} = 0$).

By simulating (1) recursively we can calculate the (7×1)-vector-valued impulse-response function $\mathbf{c}_s = \partial \mathbf{y}_{t+s}/\partial v_{OIL,t}$ to determine the effect of a 10% increase in the net oil price on the value of each element of $\mathbf{y}_{t+s}$. One way to find this is to set $\mathbf{k}_0 = \mathbf{y}_{t-1} = \mathbf{y}_{t-2} = ... = \mathbf{y}_{t-p} = \mathbf{0}$, $y_{GDP,t} = y_{DEF,t} = y_{COM,t} = 0$, and $y_{OIL,t} = 0.1$. Then calculate $y_{FED,t}$, use this to calculate $y_{TB3,t}$, then $y_{T10,t}, y_{GDP,t+1}, y_{DEF,t+1}$,and so on. The value for $\mathbf{c}_s$ is the magnitude of $\mathbf{y}_{t+s}$ from this simulation. These values are plotted as the solid lines in the first seven panels of Figure 1, which reproduce Bernanke, Gertler, and Watson's Figure 4. A 10% increase in oil prices would result in 0.25% slower real GDP growth and 0.2% higher prices after two years, with the Fed funds rate rising 80 basis points within the first year.

This pattern suggests asking whether it is the rise in interest rates rather than the oil shock itself that causes the slowdown. To sort this out, ideally we would study historical

episodes in which oil prices rose but interest rates did not. Unfortunately, there is only one instance in which this occurred: oil prices rose 65% (logarithmically) following Iraq's invasion of Kuwait in 1990, while interest rates changed very little. The U.S. did experience a recession at this time, but it is difficult to draw a conclusion from a single observation. Instead, one would like to use all the data, as summarized by a VAR, to say something about what would have happened in other episodes or typical episodes under a more expansionary monetary policy.

Asking such a counterfactual question raises a host of issues, such as the Lucas (1976) critique– if the process followed by monetary policy differed from the historical pattern, other equations of the system might have behaved differently as well. Bernanke, Gertler, and Watson devoted much of their paper to dealing with this problem, proposing a modification of the term structure relations that might have resulted from the different monetary policy, reported as "anticipated policy" simulations in some of the figures from their paper. In this paper, we do not comment on the seriousness of the Lucas critique or the adequacy of Bernanke, Gertler, and Watson's solution to it. Instead, we focus primarily on the base case simulations of Bernanke, Gertler, and Watson, which they attributed methodologically to Sims and Zha (1996). One way to implement the Sims-Zha calculation of the consequences of a policy of holding the Fed funds rate constant in the face of the shock is to calculate the value of $v_{FED,t+s}$ that would reset $y_{FED,t+s}$ to zero, and add this shock in before calculating $y_{TB3,t+s}$ at each step $s = 0, 1, 2, ...$ of the simulation described above.[2]

---

[2] This is numerically identical to (though computationally more trouble than) Bernanke, Gertler and Watson's method, which was to set all the coefficients on the lagged Fed funds rate to zero and simply drop

The authors also used this Sims-Zha method to find the effects attributable to specific historical shocks. For example, by setting $\mathbf{y}_{t-1}, \mathbf{y}_{t-2}, ..., \mathbf{y}_{t-p}$ equal to the values actually observed in 1971:12, 1971:11, ..., $1971:12 - p + 1$, equation (1) can be used to calculate $\hat{y}_{GDP,t}$, $\hat{y}_{DEF,t}$, and $\hat{y}_{COM,t}$. A value can then be imputed to $\hat{v}_{OIL,t}$ to match the observed value for the oil price variable $y_{OIL,t}$ in 1972:1. From these one goes on to calculate $\hat{y}_{FED,t}, \hat{y}_{TB3,t}, ..., \hat{y}_{COM,t+1}$, again shocking oil prices so as to match the observed $y_{OIL,t+1}$. The thin solid lines in Figure 2 plot the results of these simulations, which Bernanke, Gertler and Watson interpret as the portion of historical movements in $\mathbf{y}_{t+s}$ that might be attributed to the 1973-74 oil shock. The bold lines in Figure 2 display the actual behavior of each series. The simulation predicts a modest downturn in GDP growth, though most of the recession of 1974-75 would seem to require other explanations. Repeating the same simulation with policy shocks $\hat{v}_{FED,t+s}$ added so as to hold the simulated Fed funds rate constant at 4% produces the dashed lines in Figure 2. Again, our Figure 2 simply replicates Bernanke, Gertler, and Watson's Figure 6.

3 On the feasibility of Federal Reserve Policy

An increase in the rate of growth of the money supply has two effects. Added liquidity should reduce nominal interest rates, whereas more rapid inflation will increase nominal interest rates. Most economists agree that, for a modest and unanticipated monetary expansion, the liquidity effect would dominate. There is considerable disagreement, however, as to

---

it from the simulation. The reason for calculating this multiplier as described in the text is to obtain the implicit series for $v_{FED,t+s}$.

how large a decrease in the nominal interest rate the Fed could achieve through a monetary expansion, or how long it might be able to keep interest rates down, before the second effect would start to overwhelm its efforts.

We follow Sims (1982) and Leeper and Zha (1996) in exploring this question on the basis of the policy residuals $v_{FED,t+s}$ that are necessary to sustain the policy experiments discussed above.[3] The last panel of Figure 1 plots the values of $v_{FED,t+s}$ that are implied by the dashed lines (constant Fed funds rate policy) of the other panels in Figure 1. There is a troubling pattern to the hypothesized innovations: every innovation from $t+11$ on is negative. This means that the Fed would have to surprise the public (if the public used the historical VAR for forecasting) by setting the funds rate lower than they would have predicted on the basis of $\mathbf{y}_{t+s-1}$ for 36 months in succession in order to achieve the dashed lines shown in Figure 1. Under rational expectations, true forecast errors should be white noise, positive as often as negative. The probability that a fair coin would come up tails 36 times in a row is on the order of 1 in 100 billion. On the basis of the criterion suggested by Sims (1982, p. 144), one might thus entertain doubts about the plausibility of the Fed achieving the dashed time paths in Figure 1.

Another way to evaluate the overall plausibility of a particular proposed path for the innovations can be adapted from a suggestion of Leeper and Zha (1999). First, we calculate impulse-response coefficients $\tilde{\mathbf{c}}_s$ as in Figure 1 with the proviso that feedback from the shocks $v_{OIL,t}$ and $v_{FED,t}$ to oil prices is shut down; we did this by adding shocks $v_{OIL,t+s}$ so that

_____

[3] Although Leeper and Zha suggest a Bayesian approach, here we adopt a classical statistical perspective throughout.

$y_{OIL,t+s}$ remains at zero for $s = 1, 2, ...$ in the impulse-response simulations of Figure 1. The contribution of the policy shocks $\hat{v}_{FED,t+s}$ implicit in Figure 2 to the simulated paths for $\hat{\mathbf{y}}_{t+s}$ is then found from

$$\hat{\mathbf{h}}_{t+s} = \tilde{\mathbf{c}}_0 \hat{v}_{FED,t+s} + \tilde{\mathbf{c}}_1 \hat{v}_{FED,t+s-1} + ... + \tilde{\mathbf{c}}_s \hat{v}_{FED,t}. \tag{2}$$

The solid lines in Figure 3 plot the elements of $\hat{\mathbf{h}}_{t+s}$ as a function of time. Given the linearity of the VAR, one way to interpret these plots is as the magnitude of an expansionary effect that this policy would have produced in the absence of an oil shock. The policy sequence that Bernanke, Gertler, and Watson analyze would have produced 5% faster output growth, 5% higher prices, and so on, four years after the shock. Note that these plots of $\hat{\mathbf{h}}_{t+s}$ equal the difference between the dashed and thin solid curves of Figure 2 by construction.

Leeper and Zha noted that, if the residuals of the VAR are Gaussian, a magnitude such as $\hat{\mathbf{h}}_{t+s}$ would be distributed $N(\mathbf{0}, \mathbf{\Omega}_s)$ where $\mathbf{\Omega}_s = \sigma^2_{FED}(\tilde{\mathbf{c}}_0\tilde{\mathbf{c}}_0' + \tilde{\mathbf{c}}_1\tilde{\mathbf{c}}_1' + ... + \tilde{\mathbf{c}}_s\tilde{\mathbf{c}}_s')$ for σ^2_{FED} the variance of the Fed policy innovation $v_{FED,t}$. Figure 3 also plots ± 2 times the square roots of diagonal elements of $\mathbf{\Omega}_s$, which Leeper and Zha proposed as bounds on what constitutes a plausible policy sequence. This calculation suggests that a modest policy experiment could not cause output to deviate more than 2% from its historical path.[4]

---

[4] One might be concerned that these 95% plausibility intervals could be misleading if the distribution of $v_{FED,t}$ is sufficiently non-Gaussian. We checked for this in the current setting with a simple bootstrap experiment. We drew 10,000 realizations of $\dot{v}_{FED,\tau}$ for $\tau = 1, 2, ..., 60$ from the following distribution: $\dot{v}_{FED,\tau}$ has probability $1/T$ of equalling $\hat{v}_{FED,t}$ for $t = 1, 2, ..., T$, where $\hat{v}_{FED,t}$ is the observed residual from the original OLS estimation of the fifth equation in (1) for some historical t, T is the number of observations in that estimation, and the sampling is with replacement. For each sequence $\{\dot{v}_{FED,\tau}\}_{\tau=1}^{60}$, we calculated $\dot{\mathbf{h}}_s = \sum_{\tau=1}^{s} \tilde{\mathbf{c}}_s \dot{v}_s$ and the upper and lower 2.5% tails of $\dot{\mathbf{h}}_s$ across boostrap simulations for each s and each element of the vector $\mathbf{h}_s$, where $\tilde{\mathbf{c}}_s$ is the vector of restricted impulse-response coefficients that came from the original historical regressions. The resulting confidence intervals typically differ from the Leeper-Zha magnitudes by less than 5%.

Figures 4 and 5 repeat this analysis for the dual oil shocks associated with the Iranian revolution in 1978 and the Iran-Iraq War in 1980, this time reproducing Bernanke, Gertler, and Watson's policy experiment of holding the Fed funds rate constant at 9%. Again under Bernanke, Gertler, and Watson's policy experiment, all six variables follow substantially more expansionary paths than monetary policy has typically been observed to achieve. By contrast, the simulated policy effects for dealing with the Persian Gulf War in 1990 by freezing the Fed funds rate at 7% (Figures 6 and 7) just barely extend outside the 95% confidence regions, though here even with the expansionary policy output still slows to a 1% growth rate.

Bernanke, Gertler and Watson might argue that they were specifically interested in evaluating a policy that was out of the ordinary, in which case perhaps the correct conclusion from these figures is that the Fed really should have been doing something much different from its usual historical pattern. It nevertheless seems fair to conclude that, if one believes that it is possible for the Fed to achieve a policy such as the dashed lines in Figures 2 or 4, the claim for its feasibility is not based on the coefficient values of the historical VAR.

4   The role of lag length

The results of the previous sections all accept Bernanke, Gertler, and Watson's specification of $p = 7$ monthly lags for the VAR. This lag length is substantially shorter than that used in most of the previous literature on the effects of oil shocks. Table 1 summarizes results from those studies of which we are aware in which the actual values for the estimated coefficients

of the VAR were included in the published papers. All use quarterly data, though the table includes a variety of data sets, measures of oil prices, specification of the dynamic relation, and other included explanatory variables. In each case we focus on the equation in the system that explains output growth. The table reports the values of j for which $|\beta_j|$ is biggest in specifications of the form

$$y_t = \sum_{j=1}^{p} \beta_j o_{t-j} + f(\mathbf{x}_{t-1}, \varepsilon_{t-1})$$

for y_t a measure of output, o_{t-j} a measure of oil shocks, and $\mathbf{x}_{t-1}$ a vector of lags of y_t and other variables. Without exception, every quarterly study has reported β_4 to be biggest and β_3 to be second biggest in absolute value. Hence, Bernanke, Gertler, and Watson's choice of $p = 7$ monthly lags rules out the primary effects of oil shocks that have been reported in the literature. The importance of seasonal factors in many time series is another reason one might prefer to use 12 lags in a monthly VAR.

Bernanke, Gertler, and Watson's choice of $p = 7$ is based on the value of p that minimizes the Akaike Information Criterion, which is equivalent to maximizing the difference between the log likelihood and the number of estimated parameters (see Akaike, 1978). The log likelihood of the 7-lag VAR is 2987.84, while that for a 12-lag VAR is 3187.66. The usual likelihood ratio test statistic is

$$2(3187.66 - 2987.84) = 399.64 \tag{3}$$

which under the null hypothesis that the data are adequately represented as a 7th-order VAR would asymptotically have a $\chi^2(245)$ distribution. The probability that a $\chi^2(245)$

11

variate could exceed 399.64 is on the order of 1 in a billion, leading to rather conclusive rejection of the 7-lag specification using standard significance levels. The Akaike criterion is penalizing the extra parameters more heavily than the classical hypothesis test, and would require a likelihood ratio statistic in excess of 490 before favoring the 12-lag specification.

There may be serious problems in using the large-sample distribution for the likelihood ratio test with so many estimated parameters. Sims (1980, p. 17) suggested adjusting the likelihood ratio test for small samples by multiplying (3) by $(T - c)/T$, where T is the number of observations (360) and c the number of estimated parameters per equation (85). This would produce a small-sample-adjusted likelihood ratio $\chi^2(245)$ statistic of 305.28, for which we would reject the 7-lag specification with a more comfortable p-value of 0.005. We also performed a simple bootstrap experiment to evaluate the small-sample p-value. We estimated an unrestricted reduced-form VAR for the observed data $\mathbf{y}_t$ with 7 lags,

$$\mathbf{y}_t = \hat{\mathbf{k}} + \hat{\boldsymbol{\Phi}}_1 \mathbf{y}_{t-1} + \hat{\boldsymbol{\Phi}}_2 \mathbf{y}_{t-2} + ... + \hat{\boldsymbol{\Phi}}_7 \mathbf{y}_{t-7} + \hat{\mathbf{v}}_t$$

for $t = 1, 2, ..., T$. For the ith artificial sample, we generated a sequence of random vectors $\{\dot{\mathbf{v}}_\tau^{(i)}\}_{\tau=1}^{T+100}$ where $\dot{\mathbf{v}}_\tau^{(i)}$ has probability $1/T$ of equalling the first historical residual ($\hat{\mathbf{v}}_1$), $1/T$ of equalling the second historical residual ($\hat{\mathbf{v}}_2$), ..., and $1/T$ of equalling the last historical residual ($\hat{\mathbf{v}}_T$), and where the sequence $\{\dot{\mathbf{v}}_\tau^{(i)}\}_{\tau=1}^{T+100}$ is generated i.i.d. with replacement. We then generated the ith artificial sample $\{\dot{\mathbf{y}}_\tau^{(i)}\}_{\tau=1}^{T+100}$ according to

$$\dot{\mathbf{y}}_\tau^{(i)} = \hat{\mathbf{k}} + \hat{\boldsymbol{\Phi}}_1 \dot{\mathbf{y}}_{\tau-1}^{(i)} + \hat{\boldsymbol{\Phi}}_2 \dot{\mathbf{y}}_{\tau-2}^{(i)} + ... + \hat{\boldsymbol{\Phi}}_7 \dot{\mathbf{y}}_{\tau-7}^{(i)} + \dot{\mathbf{v}}_\tau^{(i)}$$

with starting values equal to the historical initial observations:

$$\dot{\mathbf{y}}_{\tau}^{(i)} = \hat{\mathbf{y}}_{\tau} \quad \text{for } \tau = 0, -1, ..., -6.$$

We threw out the first hundred values $\{\dot{\mathbf{y}}_{\tau}^{(i)}\}_{\tau=1}^{100}$ to reduce the contribution of initial conditions, and fit a VAR for the generated data by OLS:

$$\dot{\mathbf{y}}_{\tau}^{(i)} = \tilde{\mathbf{k}}^{(i)} + \tilde{\boldsymbol{\Phi}}_{1}^{(i)} \dot{\mathbf{y}}_{\tau-1}^{(i)} + \tilde{\boldsymbol{\Phi}}_{2}^{(i)} \dot{\mathbf{y}}_{\tau-2}^{(i)} + ... + \tilde{\boldsymbol{\Phi}}_{7}^{(i)} \dot{\mathbf{y}}_{\tau-7}^{(i)} + \tilde{\mathbf{v}}_{\tau}^{(i)}$$

for $\tau = 101, ..., T + 100$. We did a likelihood ratio test of the null hypothesis of $p = 7$ lags (which is true by construction for these artificial data) against the alternative of $p = 12$ lags. We repeated this procedure to generate $i = 1, 2, ..., 10,000$ different artificial samples each of length T. The likelihood ratio statistic exceeded the value of 399.64 found for the actual historical 7-lag VAR in only 23 of these simulations. This suggests that the Sims small-sample correction used above is if anything too conservative for this application, and the null hypothesis of 7 lags should be rejected with a p-value of less than 0.005.

Where one has no prior information about lag length, a recent study by Ivanov and Kilian (2000) argues that choosing a lag length using the Akaike criterion is a good approach. When, as here, there are strong claims about the value of p based on prior studies, comparing the alternative claims using classical hypothesis tests strikes us as a more natural way to frame the issues. Notwithstanding, one can combine hypothesis testing with lag length selection using a procedure described by Lütkepohl (1993, p. 125). The likelihood ratio test of the null hypothesis of p_0 lags against the alternative of $p_0 + 1$ lags is asymptotically independent of the test of p_1 against $p_1 + 1$ for $p_0 \neq p_1$. Hence one can construct a test

of size 0.05 of the null hypothesis of $p = 7$ lags against the alternative of p somewhere

between 8 and 24, and simultaneously choose the alternative, as follows. First test the null

hypothesis of $p = 23$ lags against the alternative of $p = 24$. If this rejects at the 0.003 level

of significance, choose $p = 24$. If it accepts, proceed to test $p = 22$ against $p = 23$, and if it

rejects at the 0.003 level, choose $p = 23$. If the true value of p is 7, the probability that one

will end up selecting a value of $p = 7$ is given by $(1 - 0.003)^{24-7} = 0.95$; hence this sequential

testing procedure has a cumulative probability of 5% of rejecting the null hypothesis $(p = 7)$

if it is in fact true.

Using the Sims correction to calculate critical values for each individual test, the first

rejection encountered is for the null of $p = 15$ lags against the alternative of $p = 16$ lags.

Hence the procedure would call for rejection of the 7-lag VAR in favor of a 16-lag VAR.

Indeed, three separate tests (11 versus 12, 12 versus 13, and 15 versus 16) each call for rejec-

tion individually, even after the Sims correction, at the 1% level. Given the independence

of the tests, the probability of three or more such outcomes is

$$\sum_{k=0}^{14} \binom{17}{k} (0.99)^k (1 - .0.99)^{17-k} = 0.0006.$$

Hence the evidence against the null hypothesis of $p = 7$ lags in favor of using more lags

would seem to be quite strong.

In light of the above, it seems worthwhile investigating how Bernanke, Gertler, and

Watson's conclusions might change if one uses $p = 12$ or $p = 16$ lags instead of 7.

Bernanke, Gertler, and Watson explored a variety of measures of oil price shocks, includ-

ing the nominal change in crude oil prices, dummy variables for the dates of major shocks suggested by Hoover and Perez (1994), the absolute value of the percentage change in the relative price of oil, as in Mork (1989), and the annual net oil price increase suggested by Hamilton (1996). Figure 8 plots the impulse-response functions relating output growth to an oil price shock, along with ± 1 standard deviation bands. The left-hand panels use $p = 7$ lags, and are identical to Figure 2 in Bernanke, Gertler, and Watson. On the basis of the generally weak effect of any of the measures, the authors suggested that "finding a measure of oil price shocks that 'works' in a VAR context is not straightforward," though they acknowledged that "for some more complex– some might argue, data-mined– indicators of oil prices, an exogenous increase in the price of oil has the expected effects on the economy" (page 105).

The middle panel of Figure 8 reproduces these calculations with the lag order p of the VAR set to 12 months. The effects are typically twice as large as those in the left panels, and the overall pattern (though not the magnitude) is the same for any one of the four measures of an oil price shock. The right-hand panel uses $p = 16$, and is very similar in appearance to the $p = 12$ case.[5] It would appear that lag length, rather than sample period or oil price measure, is the key explanation for the discrepancy between Bernanke, Gertler, and Watson's conclusion and that of other researchers.

Figure 9 reproduces the impulse-response diagrams from Figure 1 using now 12 lags. In

---

[5] We found very similar results when the lag length on the oil variable is set to 12 and all other variables at 7. Hence it appears to be a delayed effect of an oil shock on the economy, rather than changes in the endogenous dynamics of other variables, that accounts for the difference.

this case, we find that even if the Federal Reserve did have the power to prevent the Federal funds rate from rising after an oil shock, such a policy would do little to mitigate the contractionary effects of the shock, though it would nearly double the inflationary consequences for the price level.[6]

Figure 10 analyzes the effects on output growth of three different oil shocks– the OPEC embargo of 1973, the Iranian revolution and Iran-Iraq War in 1978-80, and the Persian Gulf War in 1990. The left three panels use $p = 7$ lags in the VAR, and reproduce Figure 6 in Bernanke, Gertler and Watson and the first panels of our Figures 2, 4, and 6. The right three panels repeat this analysis with 12 lags. Despite significant stimulus from the policy of holding the Fed funds rate constant, an important disruption in output growth would have occurred in all three episodes, according to these simulations. Graphs using $p = 16$ lags are very similar to the $p = 12$ case shown.

5   Conclusion

We conclude that the potential of monetary policy to avert the contractionary consequences of an oil price shock is not as great as suggested by the analysis of Bernanke, Gertler, and Watson. Oil shocks appear to have a bigger effect on the economy than suggested by their VAR, and we are unpersuaded of the feasibility of implementing the monetary policy needed to offset even their small shocks.

A key basis for believing that oil shocks have a bigger effect than implied by the Bernanke,

---

[6] These correspond to the Sims-Zha experiments in Figure 4 of Bernanke, Gertler and Watson. We also found similar effects of expanding the lag length to $p = 12$ on the anticipated policy graphs in their Figure 4.

Gertler, and Watson estimates is that the biggest effects of an oil shock do not appear until three or four quarters after the shock. Investigating the cause of this delay would seem to be an important topic for research. It may be, as Bernanke, Gertler, and Watson proposed, that a response of monetary policy to inflation is an important part of the process, and the role of monetary policy simply is not fully captured by including the Federal funds rate, short rate, and long rate in the vector autoregression or conditioning on a fixed time path for the former. Alternatively, Herrera (2000) suggests that the lag may result from accumulation of inventories immediately following the oil shock.

A separate important question, not considered in this paper, is the role of monetary policy in the inflation associated with historical oil shocks. A number of important recent studies, including DeLong (1997), Barsky and Kilian (1999), Hooker (1999), and Clarida, Galí and Gertler (2000), suggest that different monetary policy (particularly prior to the oil shocks) could have made a substantial difference for inflation, even if we are correct that monetary policy might have had little potential to avert an economic slowdown.

References

Akaike, Hirotugu (1978), "On the Likelihood of a Time Series Model," *The Statistician,* 27, pp. 217-235.

Balke, Nathan S., Stephen P. A. Brown, and Mine Yücel (1999). "Oil Price Shocks and the U.S. Economy: Where Does the Asymmetry Originate?", working paper, Federal Reserve Bank of Dallas.

Barsky, Robert B., and Lutz Kilian (1999), "Money, Stagflation and Oil Prices: A Re-Interpretation," working paper, University of Michigan.

Bernanke, Ben S., Mark Gertler, and Mark Watson (1997). "Systematic Monetary Policy and the Effects of Oil Price Shocks," *Brookings Papers on Economic Activity,* 1:1997, pp. 91-142.

Bohi, Douglas R. (1989), *Energy Price Shocks and Macroeconomic Performance*, Washington D.C.: Resources for the Future.

Brown, Stephen P. A., and Mine K. Yücel (1999), "Oil Prices and U.S. Aggregate Economic Activity: A Question of Neutrality," *Federal Reserve Bank of Dallas Economic and Financial Review*, Second Quarter 1999, pp. 16-23.

Burbidge, John, and Alan Harrison (1984), "Testing for the Effects of Oil-Price Rises Using Vector Autoregressions," *International Economic Review, 25*, pp. 459-484.

Carruth, Alan A., Mark A. Hooker, and Andrew J. Oswald (1998), "Unemployment Equilibria and Input Prices: Theory and Evidence from the United States," *Review of Economics and Statistics,* 80, pp. 621-628.

Clarida, Richard, Jordi Galí, and Mark Gertler (2000). "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *Quarterly Journal of Economics,* 115, pp. 147-180.

Daniel, Betty C. (1997), "International Interdependence of National Growth Rates: A Structural Trends Analysis," *Journal of Monetary Economics, 40*, pp. 73-96.

Davis, Steven J. (1987), "Fluctuations in the Pace of Labor Reallocation," in K. Brunner and A. H. Meltzer, eds., *Empirical Studies of Velocity, Real Exchange Rates, Unemployment and Productivity*, Carnegie-Rochester Conference Series on Public Policy, 24, Amsterdam: North Holland.

_____, and John Haltiwanger (2001), "Sectoral Job Creation and Destruction Responses to Oil Price Changes and Other Shocks," forthcoming, *Journal of Monetary Economics.*

_____, Prakash Loungani, and Ramamohan Mahidhara (1996), "Regional Labor Fluctuations: Oil Shocks, Military Spending, and Other Driving Forces," working paper, University of Chicago.

DeLong, J. Bradford (1997), "America's Peacetime Inflation: The 1970s," in C. Romer and D. Romer, eds., *Reducing Inflation: Motivation and Strategy,* Chicago: Chicago University Press.

Dotsey, Michael, and Max Reid (1992), "Oil Shocks, Monetary Policy, and Economic Activity," *Economic Review of the Federal Reserve Bank of Richmond,* 78/4, pp. 14-27.

Gisser, Micha, and Thomas H. Goodwin (1986), "Crude Oil and the Macroeconomy: Tests of Some Popular Notions," *Journal of Money, Credit, and Banking, 18,* pp. 95-103.

Hamilton, James D. (1983), "Oil and the Macroeconomy Since World War II," *Journal of Political Economy, 91*, pp. 228-248.

____ (1996), "This is What Happened to the Oil Price-Macroeconomy Relationship," *Journal of Monetary Economics*, 38, pp. 215-220.

____ (2000), "What is an Oil Shock?", working paper, UCSD, http://weber.ucsd.edu/~jhamilto.

Herrera, Ana M. (2000), *Inventories, Oil Shocks, and Aggregate Economic Behavior.* Unpublished Ph.D. dissertation, University of California at San Diego.

Hooker, Mark A. (1996), "What Happened to the Oil Price-Macroeconomy Relationship?," *Journal of Monetary Economics, 38, pp. 195-213.*

____ (1997), "Exploring the Robustness of the Oil Price-Macroeconomy Relationship," working paper, Board of Governors of the Federal Reserve System.

____ (1999), "Are Oil Shocks Inflationary? Asymmetric and Nonlinear Specifications versus Changes in Regime," working paper, Board of Governors of the Federal Reserve System.

Hoover, Kevin D., and Stephen J. Perez (1994), "Post Hoc Ergo Propter Once More: An Evaluation of 'Does Monetary Policy Matter?' in the Spirit of James Tobin," *Journal of Monetary Economics, 34,* pp. 89-99.

Ivanov, Ventzislav, and Lutz Kilian, "A Practioner's Guide to Lag-Order Selection for Vector Autoregression," working paper, University of Michigan, http://www.econ.lsa.umich.edu/wpweb/2000-03.pdf.

Keane, Michael P., and Eswar Prasad (1996), "The Employment and Wage Effects of Oil Price Changes: A Sectoral Analysis," *Review of Economics and Statistics, 78*, pp. 389-400.

Lee, Kiseok, and Shawn Ni (1999), "On the Dynamic Effects of Oil Price Shocks" A Study Using Industry Level Data," working paper, University of Missouri, Columbia.

____, ____, and Ronald A. Ratti, (1995) "Oil Shocks and the Macroeconomy: The Role of Price Variability," *Energy Journal, 16,* pp. 39-56.

Leeper, Eric M., and Tao Zha (1999), "Modest Policy Interventions," working paper, University of Indiana, http://php.indiana.edu/~eleeper/Papers/lz112499.pdf.

Loungani, Prakash (1986), "Oil Price Shocks and the Dispersion Hypothesis," *Review of Economics and Statistics, 58,* pp. 536-539.

Lucas, Robert E.Jr. (1976) "Econometric Policy Evaluation: A Critique." In K. Brunner and A. H. Meltzer, eds., *The Phillips Curve and Labor Markets, Carnegie-Rochester Conference Series on Public Policy*, vol. 1, Amsterdam: North-Holland, pp. 19-46.

Mork, Knut A. (1989), "Oil and the Macroeconomy When Prices Go Up and Down: An Extension of Hamilton's Results," *Journal of Political Economy, 91*, pp. 740-744.

Rasche, R. H., and J. A. Tatom (1977), "Energy Resources and Potential GNP," *Federal Reserve Bank of St. Louis Review*, 59 (June), pp. 10-24.

____, and ____ (1981), "Energy Price Shocks, Aggregate Supply, and Monetary Policy: The Theory and International Evidence." In K. Brunner and A. H. Meltzer, eds., *Supply Shocks, Incentives, and National Wealth, Carnegie-Rochester Conference Series on Public Policy*, vol. 14, Amsterdam: North-Holland.

Raymond, Jennie E., and Robert W. Rich (1997), "Oil and the Macroeconomy: A Markov State-Switching Approach," *Journal of Money, Credit and Banking,* 29 (May), pp. 193-213. Erratum 29 (November, Part 1), p. 555.

Rotemberg, Julio J., and Michael Woodford (1996), "Imperfect Competition and the Effects of Energy Price Increases," *Journal of Money, Credit, and Banking, 28* (part 1), pp. 549-577.

Santini, Danilo J. (1985), "The Energy-Squeeze Model: Energy Price Dynamics in U.S. Business Cycles," *International Journal of Energy Systems,* 5, pp. 18-25.

____ (1992), "Energy and the Macroeconomy: Capital Spending After an Energy Cost Shock," in J. Moroney, ed., *Advances in the Economics of Energy and Resources*, vol. 7, Greenwich, CN: J.A.I. Press.

____ (1994), "Verification of Energy's Role as a Determinant of U.S. Economic Activity," in J. Moroney, ed., *Advances in the Economics of Energy and Resources*, vol. 8, Greenwich, CN: J.A.I. Press.

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

____ (1982), "Policy Analysis with Econometric Models," *Brookings Papers on Economic Activity,* 1982:1, pp. 107-152.

____, and Tao Zha (1996), "Does Monetary Policy Generate Recessions?", working paper, Princeton University, http://eco-072399b.princeton.edu/yftp/mpolicy/TZCAS296.PDF.

Table 1

Summary of most important lag coefficients in previous

empirical studies of the effects of oil price shocks

Study	Sample period	Lag with biggest coefficient	Lag with 2nd biggest coefficient
Hamilton (1983)	1949:II-1972:IV	4th quarter	3rd quarter
Hamilton (1983)	1973:I-1980:IV	4th quarter	3rd quarter
Gisser and Goodwin (1986)	1961:I-1982:IV	4th quarter	3rd quarter
Mork (1989)	1949:I-1988:II	4th quarter	3rd quarter
Raymond and Rich (1997)	1952:II-1995:III	4th quarter	3rd quarter
Hamilton (2000)	1949:II-1999:IV	4th quarter	3rd quarter
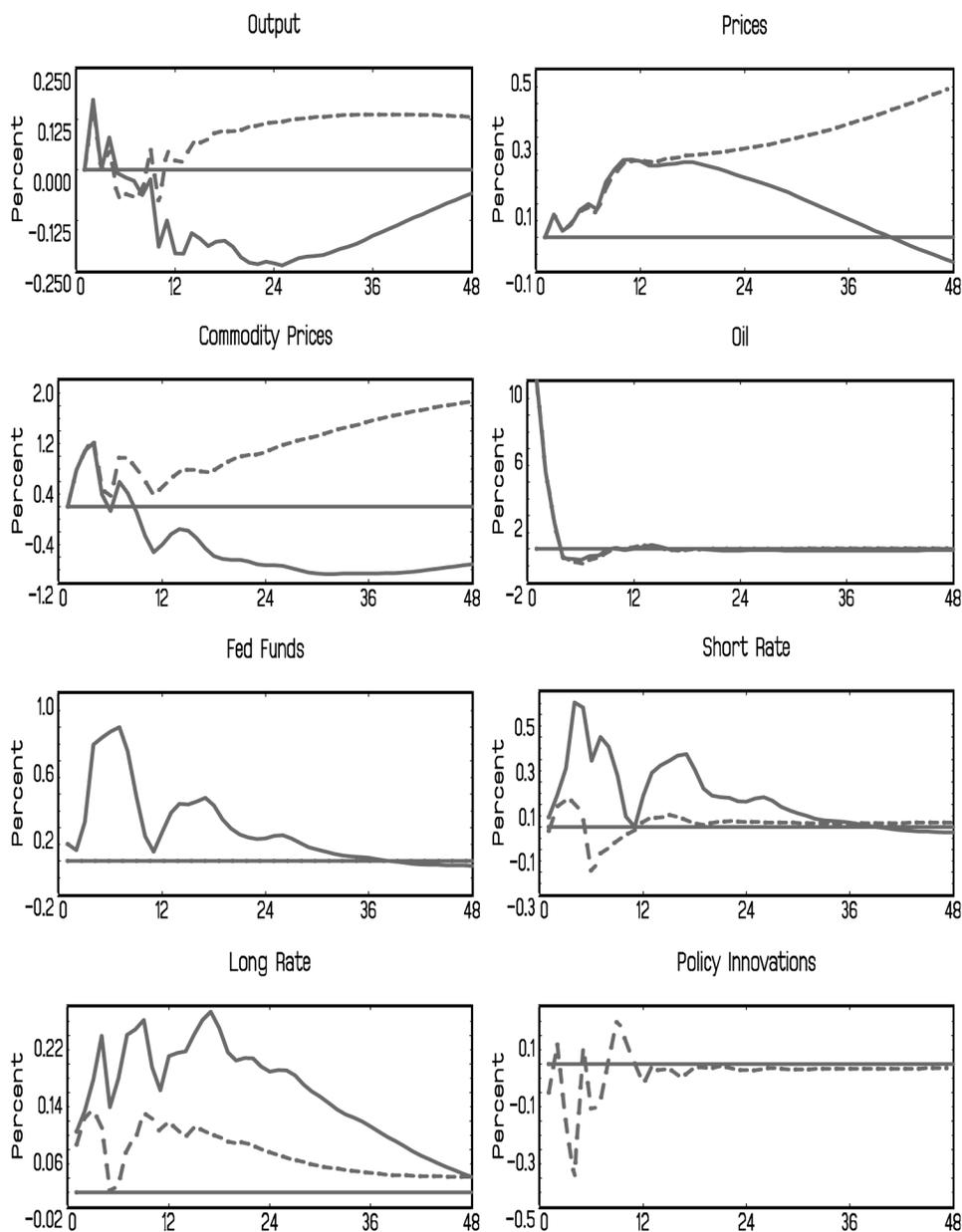
FIG. 1. Impulse-Response functions for effect of 10% oil price increase. Solid line-- $\partial y_{j,t+s}/\partial v_{OIL,t}$ for the system (1) estimated by OLS with $p = 7$ lags. Dashed line-- simulation of a system with the same coefficients but the Fed funds rate equation dropped from the system. Last panel-- sequence of Fed funds rate policy innovations $v_{FED,t+s}$ implied by such a policy.
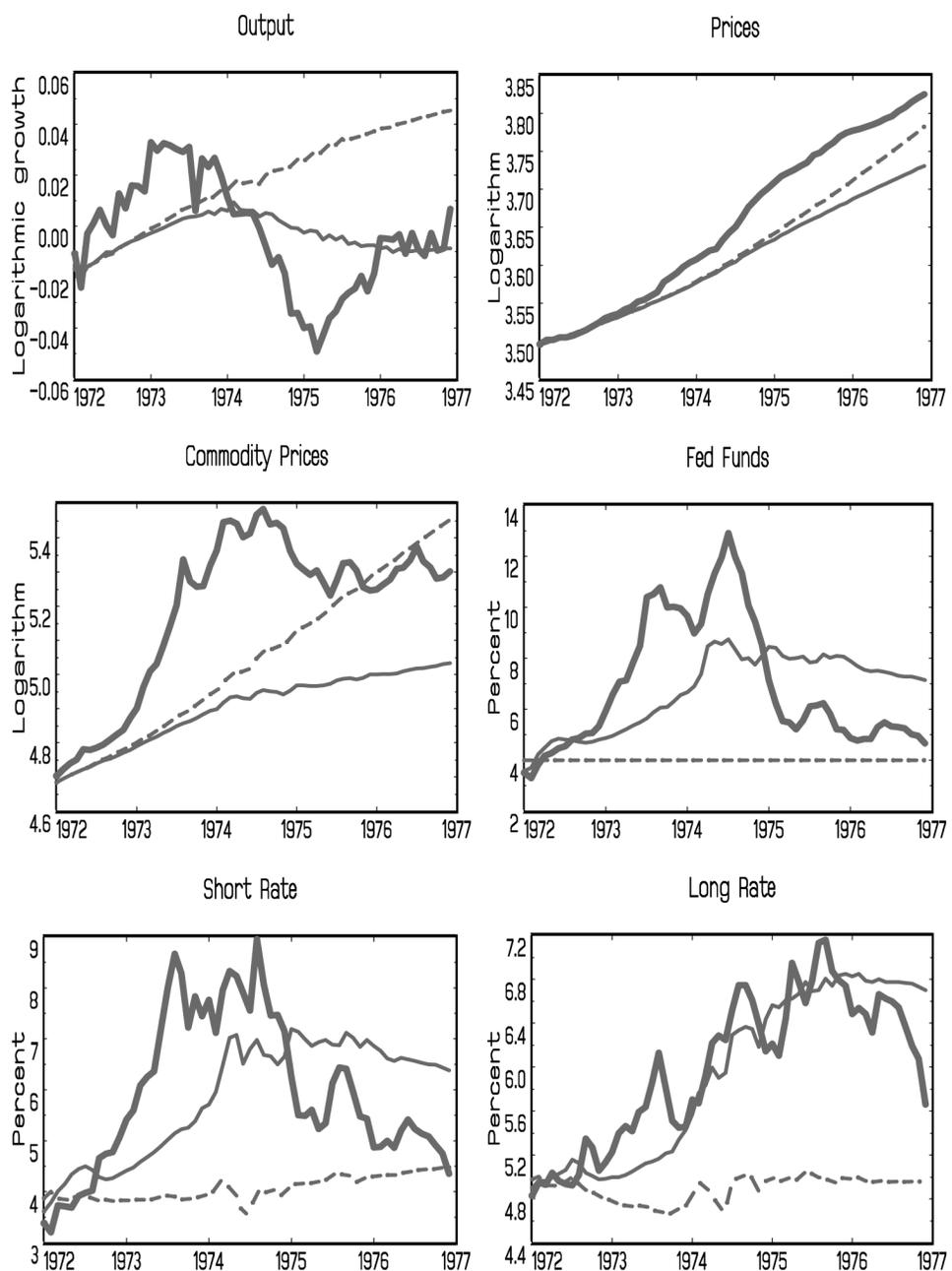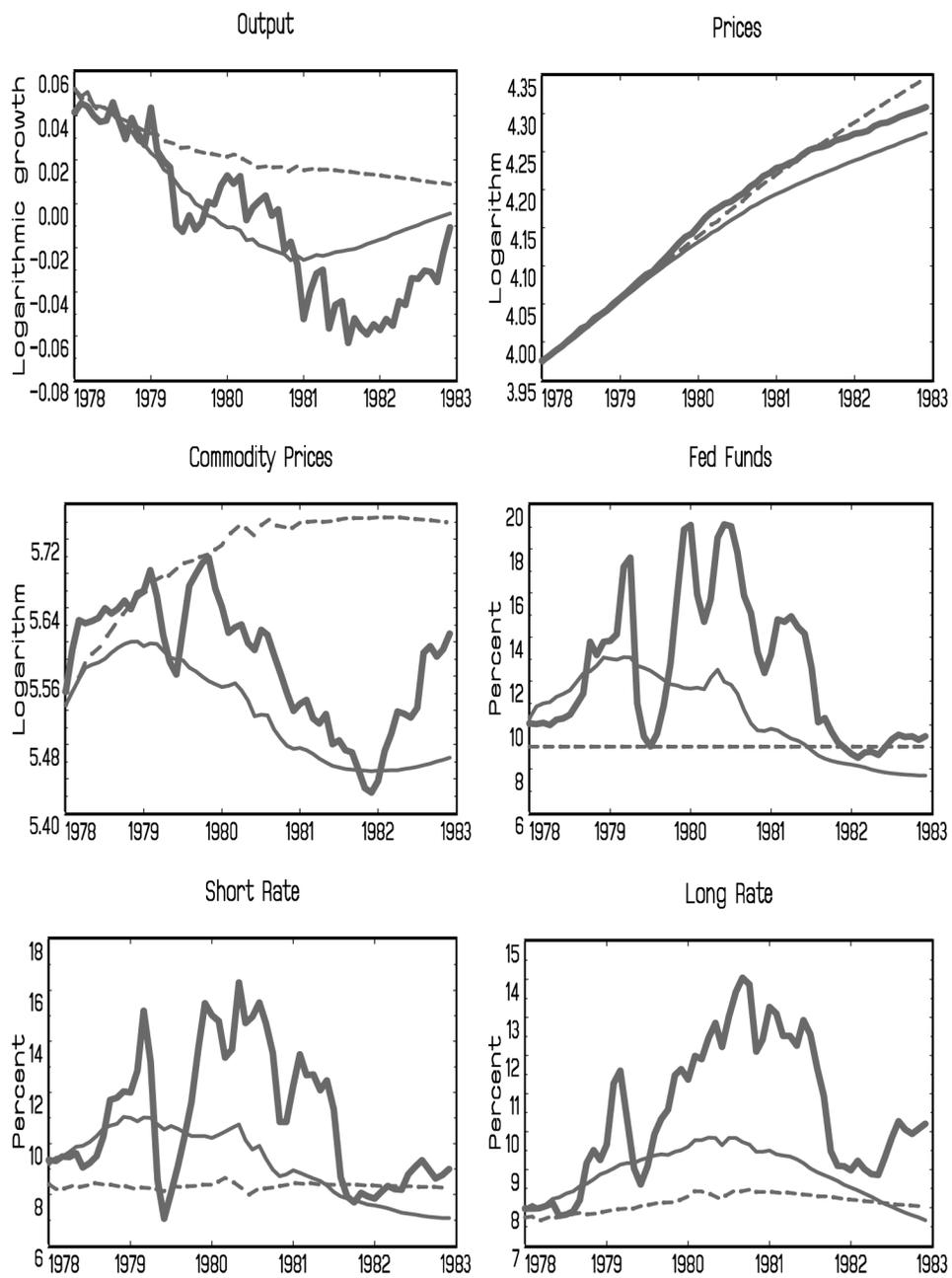
FIG. 2. Simulating policy for the 1973 oil shock. Bold line-- actual behavior of y_{jt} between 1972 and 1977. Solid line-- simulated value for y_{jt} based on OLS estimation of (1) with $p = 7$ lags. Values were obtained by setting initial values to actual values at end of 1971 and setting $v_{OIL,t}$ so that the simulated oil price equals the actual oil price with $v_{jt} = 0$ for $j \neq OIL$. Dashed line-- simulation with $v_{OIL,t}$ so as to set the simulated oil price equal to the actual oil price and $v_{FED,t}$ so as to keep the Fed funds rate at 4.0%
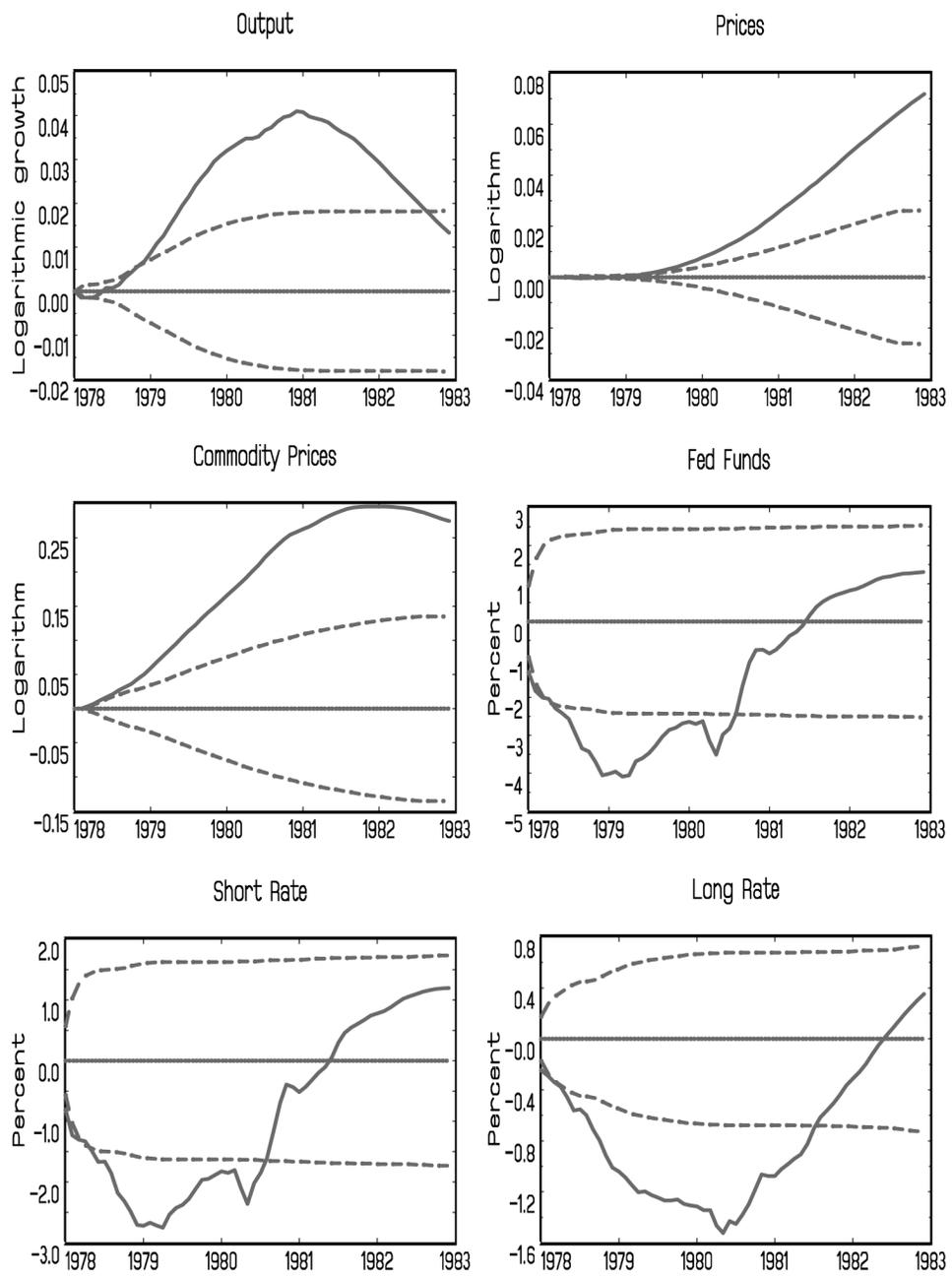
Output

Prices

Commodity Prices

Fed Funds

Short Rate

Long Rate
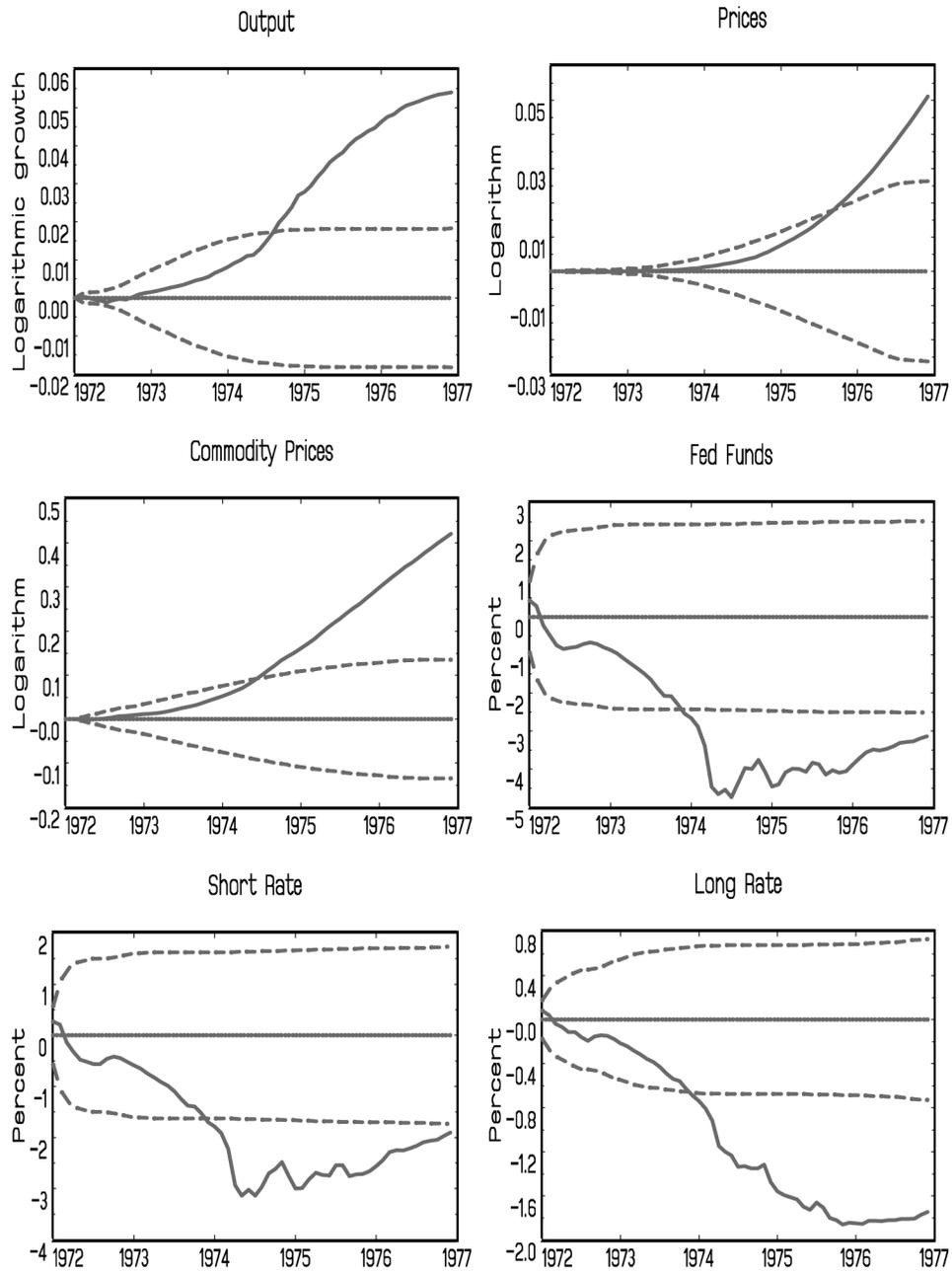
FIG. 3. Implied policy effects and 95% confidence intervals for 1972-77. Solid line-- difference between solid line and dashed line in Figure 2. Dashed lines-- 95% confidence intervals, described in text.
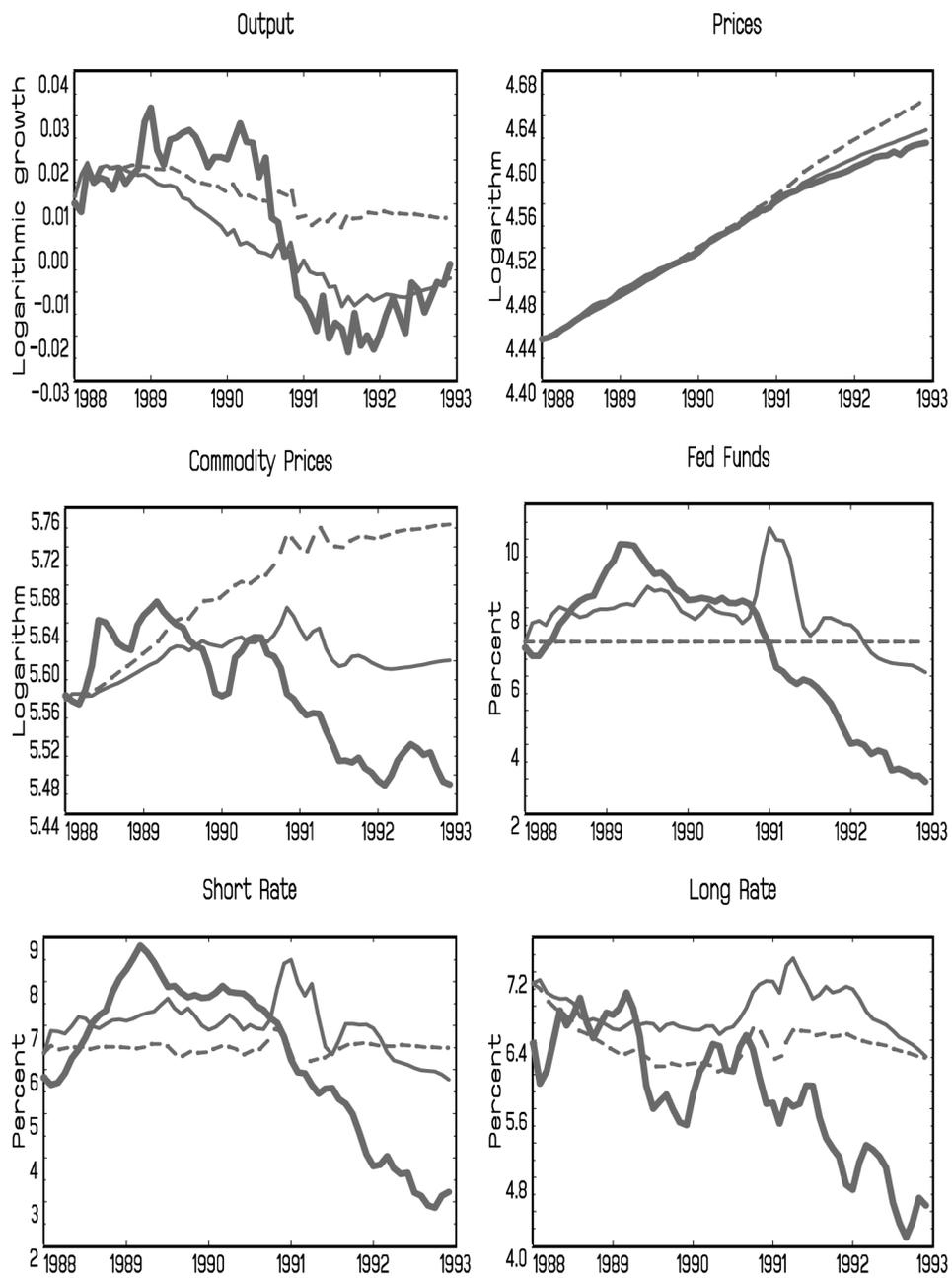
FIG. 4. Simulating policy for the 1979 oil shock. Bold line-- actual behavior of y_{jt} between 1978 and 1983. Solid line-- simulated value for y_{jt} based on OLS estimation of (1) with $p = 7$ lags. Dashed line-- simulation with $v_{OIL,t}$ so as to set the simulated oil price equal to the actual oil price and $v_{FED,t}$ so as to keep the Fed funds rate at 9.0%
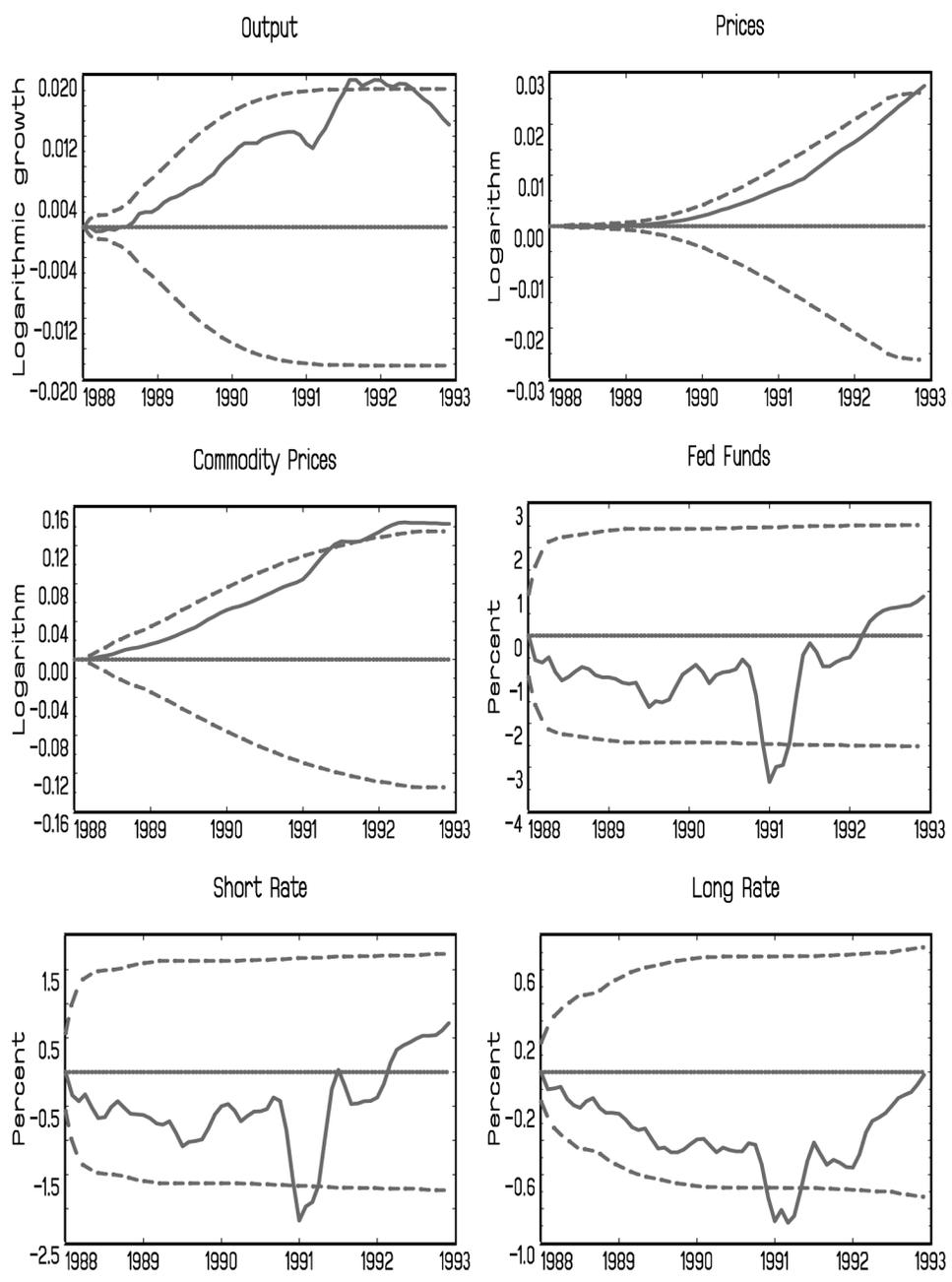
Output

Prices

Commodity Prices

Fed Funds

Short Rate

Long Rate

FIG. 5. Implied policy effects and 95% confidence intervals for 1978-83. Solid line-- difference between solid line and dashed line in Figure 4. Dashed lines-- 95% confidence intervals, described in text.

FIG. 6. Simulating policy for the 1990 oil shock. Bold line-- actual behavior of y_{jt} between 1988 and 1993. Solid line-- simulated value for y_{jt} based on OLS estimation of (1) with $p = 7$ lags. Dashed line-- simulation with $v_{OIL,t}$ so as to set the simulated oil price equal to the actual oil price and $v_{FED,t}$ so as to keep the Fed funds rate at 7.0%

FIG. 7. Implied policy effects and 95% confidence intervals for 1988-93. Solid line-- difference between solid line and dashed line in Figure 6. Dashed lines-- 95% confidence intervals, described in text.
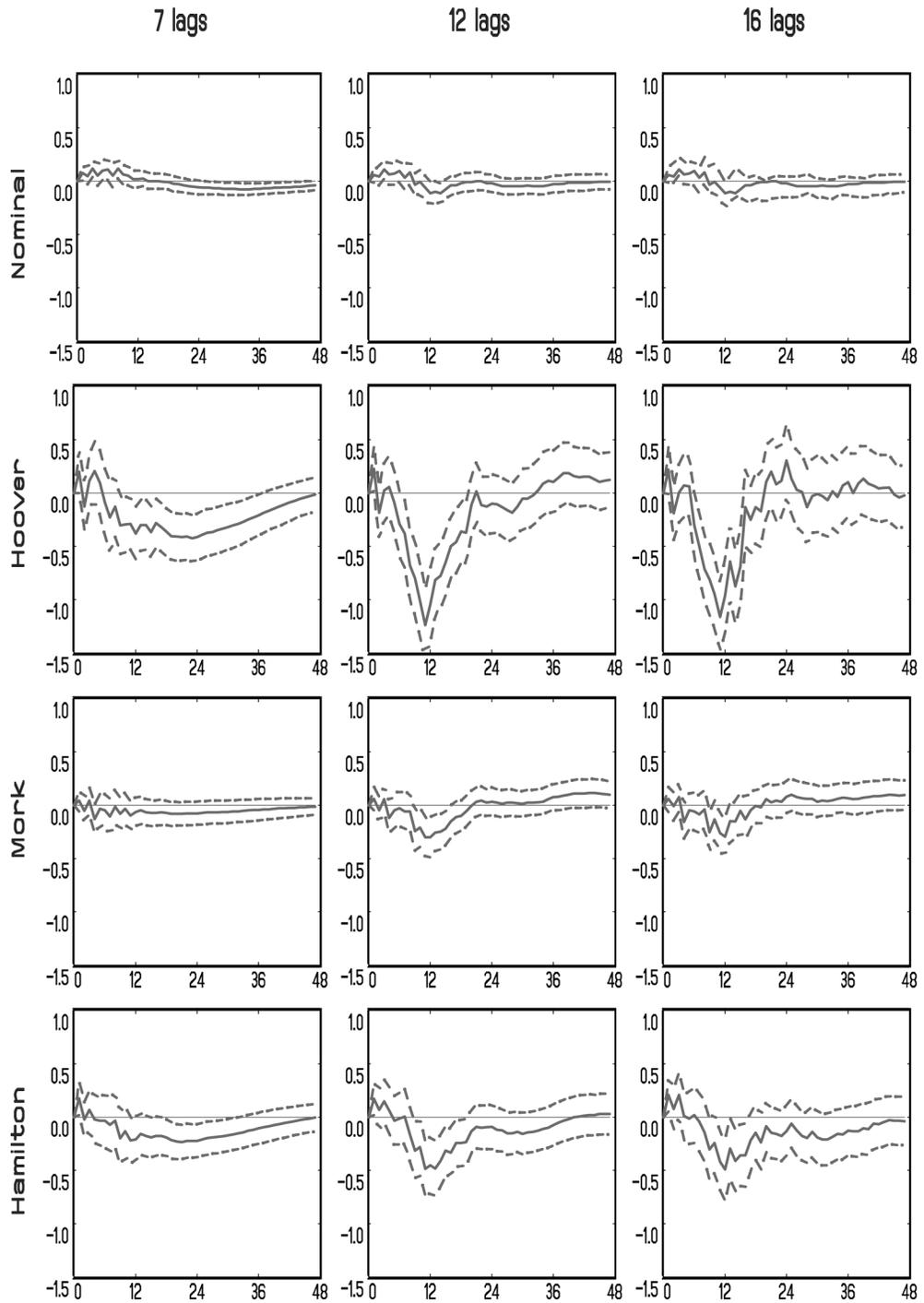
FIG. 8. Output response to an oil price shock. Impulse-response function relating oil shock to subsequent output growth for four different measures of an oil shock. Left panels-- estimates from a 7-lag VAR. Middle panels-- estimates from a 12-lag VAR. Right panels-- estimates from a 16-lag VAR.
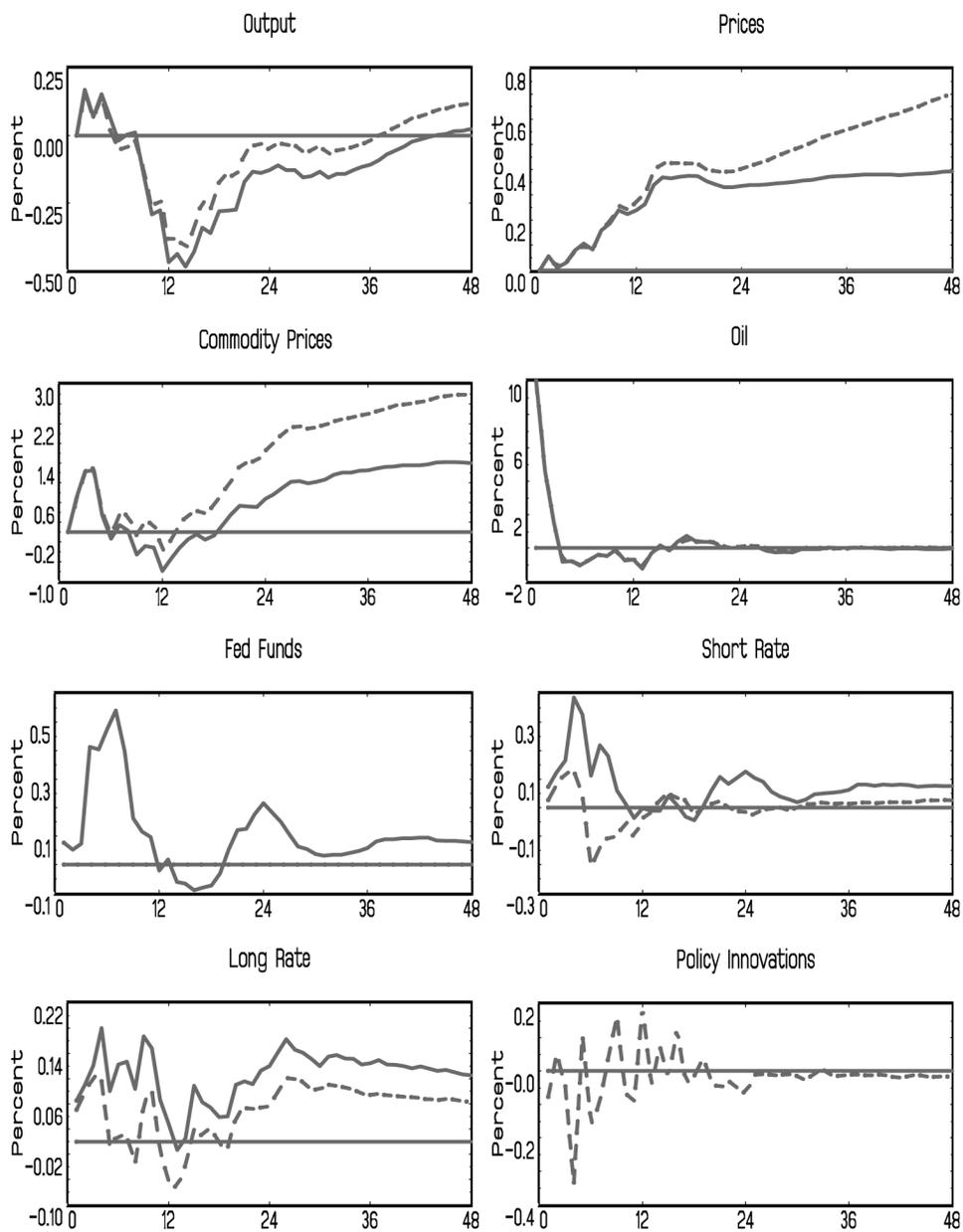
FIG. 9. Impulse-response functions (12 monthly lags). Same as Figure 1 except that the VAR uses $p = 12$.
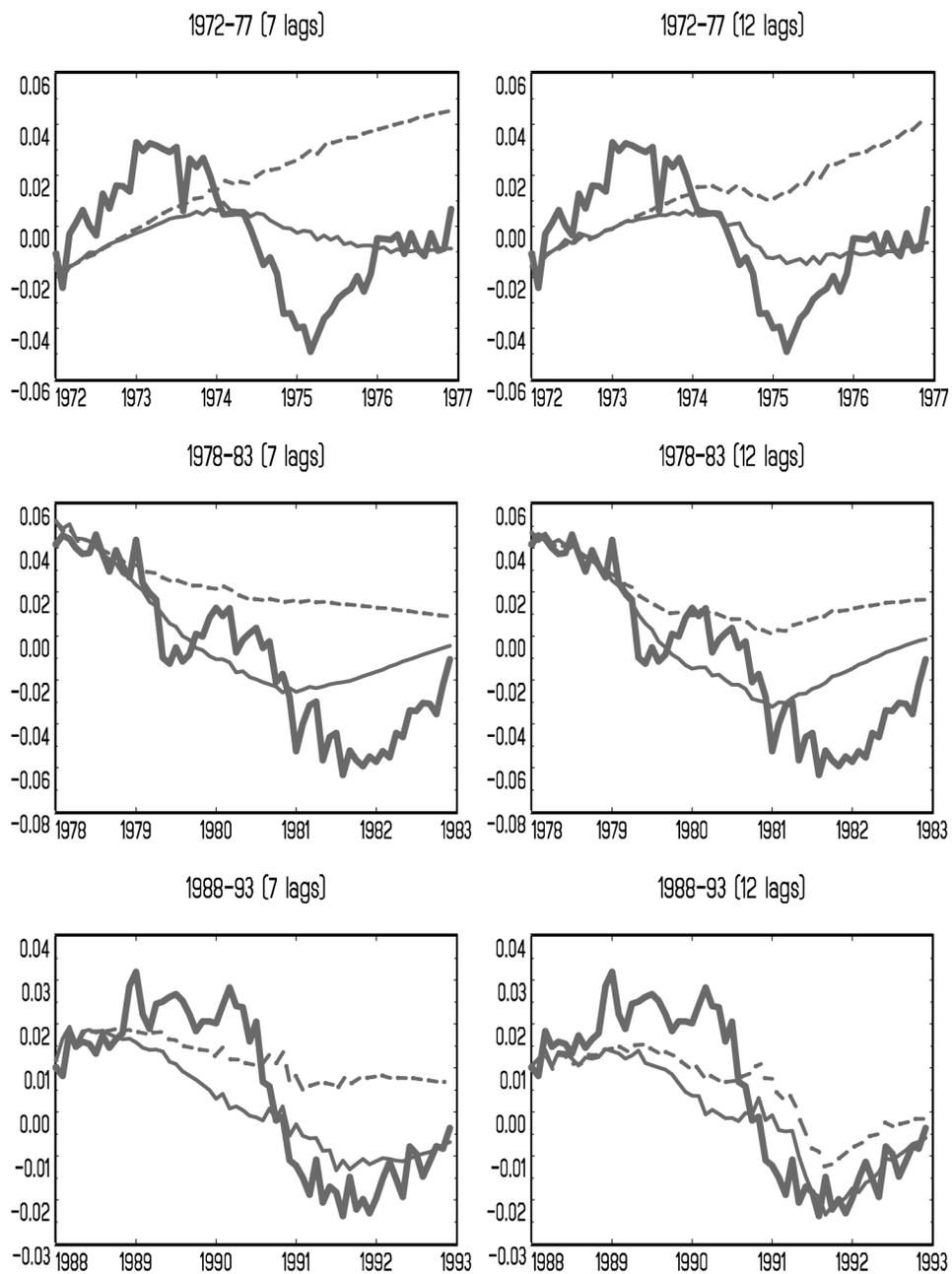
FIG. 10. Simulating policy for three historical oil shocks. Actual
levels of logarithmic growth of output (bold lines), path of output attributed to
oil shocks (solid lines), and path of output if Fed funds rate were constant at
4.0% (1972-77), 9.0% (1988-93), or 7.0% (1988-93), where VAR uses 7 lags
(left panels) or 12 lags (right panels).