Journal of Monetary Economics 60 (2013) 950–966

Contents lists available at ScienceDirect

Journal of Monetary Economics

journal homepage: www.elsevier.com/locate/jme

Monetary policy matters: Evidence from new shocks data

S. Mahdi Barakchian, Christopher Crowe*

Graduate School of Management and Economics, Sharif University of Technology, Tehran, Iran

ARTICLE INFO

Article history:
Received 19 February 2010
Received in revised form
24 September 2013
Accepted 28 September 2013
Available online 9 October 2013

Keywords:
Monetary policy
VAR estimation
Fed Funds futures
FOMC

ABSTRACT

The evidence suggests that monetary policy post 1988 became more forward-looking, invalidating the identifying assumptions in conventional methods of measuring monetary policy's effects, leading to spurious and unlikely results for this period. We propose a new identification scheme that uses factors extracted from Fed Funds futures to measure exogenous changes in policy. Using this shock series in a VAR, we recover the contractionary effect of monetary tightening on output. Moreover, we find that as much as half of the variability in output was driven by monetary policy shocks, and that there is a mild price puzzle.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Identifying the impact of monetary policy on the economy is a central question in empirical macroeconomics. The key identification problem is simultaneity. Hence, the focus has been on the exogenous or 'shock' component of policy changes. For the U.S., a consensus has emerged on the qualitative effects of a monetary policy shock. Christiano et al. (1999) summarize this consensus as follows:

> After a contractionary monetary policy shock, short term interest rates rise, aggregate output, employment, profits and various monetary aggregates fall, the aggregate price level responds very slowly, and various measures of wages fall, albeit by very modest amounts. In addition, there is agreement that monetary policy shocks account for only a very modest percentage of the volatility of aggregate output; they account for even less of the movements in the aggregate price level.

However, this consensus is sensitive to the period used for analysis. In particular, it is dependent on the inclusion of the 1970s and early 1980s, when shocks were large and the policymaking environment was different from the one faced today. When one attempts to identify the effects of monetary policy shocks for the period since the 1980s using the same methodologies one obtains quite different results. Notably, contractionary monetary policy shocks appear to have a small positive effect on output.

This paper presents some evidence on changes to the nature of U.S. monetary policy shocks that would cause conventional identification methods to give misleading results. In particular, we show that U.S. monetary policy has become more forward looking. Hence, VAR identification methods that ignore the role of forecasts in the policymaker's reaction function are misspecified. Identification methods (such as Romer and Romer, 2004) that allow for forward-looking variables in the reaction function but do not allow for the apparent increase in their relative weight will tend to suffer from the same problem.

* Corresponding author. Tel.: +44 2070710924; fax: +44 2070710950.
E-mail addresses: c.w.crowe@gmail.com, ccrowe@capulaglobal.com (C. Crowe).

0304-3932/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jmoneco.2013.09.006

We turn to financial market data in an effort to uncover a measure of monetary policy shocks that is less subject to these criticisms. Following Kuttner (2001), Gürkaynak et al. (2005) and Piazzesi and Swanson (2008) monetary policy shocks are identified as the 'surprise' component of monetary policy actions, estimated using movements in Fed Funds futures contract prices on the day of monetary policy announcements following FOMC meetings.

Factor analysis is employed to efficiently capture the information contained across the maturity spectrum, uncovering the common information from six monthly contracts: the current month and up to 5 months ahead. As in Gürkaynak et al. (2005) two factors are sufficient to summarize the information across the six contracts. Moreover, in keeping with the literature on factor models of the yield curve (e.g. Piazzesi, 2010), the factors have a natural interpretation as level and slope, respectively. The former is employed as the measure of the policy shock.

We enter this new shock measure in a simple monthly VAR, similarly to Romer and Romer (2004), estimated for 1988:12-2008:06.[1] With this new measure, a contractionary monetary policy shock has a statistically significant negative effect on output. While the effect is small in absolute terms, the forecast error variance decomposition suggests that, in an era of low overall output volatility, our new policy shock measure can account for up to half of output volatility at a horizon of 3 years or more—around twice the proportion using existing shock measures. There is some evidence for a 'price puzzle': contractionary monetary policy also leads to a small, and borderline significant, increase in the general price level at a horizon of 1–3 years, although this is subsequently reversed. Efforts to eliminate the price puzzle by including a measure of commodity prices or inflation expectations in the VAR, following suggestions in the literature, are not successful.

1.1. The related literature

Our methodology builds on the insights of an increasingly influential literature on identifying monetary policy shocks using financial market data. Rudebusch (1998) is an early paper advocating the use of Fed Funds futures data, while Kuttner's (2001) focus on one-day changes in futures prices, rather than the difference between the implied futures rate and the actual policy rate, allows for sharper identification. Faust et al. (2004) propose a novel two-stage identification scheme in which the information available from the Fed Funds futures is used to partially identify a structural VAR. Gürkaynak et al. (2005) use a two factor model to combine information from futures contracts at different horizons and separately identify level and slope factors. Hamilton (2008) derives level, slope and curvature factors using three Fed Funds futures contracts, and estimates the impact of the different factors on housing market variables. Thapar (2008) uses 3 month Treasury Bill futures prices as a proxy for market expectations, in a novel identification method that combines these market-based forecasts with Greenbook forecasts of output and price variables. D'Amico and Farka (2011) uses intraday futures data to estimate the contemporaneous relation between monetary policy and stock prices within a VAR framework. Taylor (2010) carries out a slightly different exercise, using intraday Fed Funds futures data to identify the effect of macroeconomic data announcements on market expectations of future monetary policy changes.

While this paper is therefore not the first to turn to Fed Funds futures data, this paper is the first to use shocks extracted from futures contracts to identify the responses of output and inflation to monetary policy shocks. Earlier contributions have focused on the impact of monetary policy on financial rather than macro variables. Because, in our case, the policy shock is identified outside the VAR, one can avoid some of the weaknesses of structural VAR estimation. By contrast, Faust et al. (2004) use the structural VAR model to identify the monetary policy shock and to estimate the impulse responses of the macro variables to the policy shock, and as a result their method is subject to some of these weaknesses. Like Kuttner (2001) and Hamilton (2008), but unlike Rudebusch (1998) and Thapar (2008), this paper focuses on daily innovations in Fed Funds futures prices. Using daily data from policy announcement days helps to remove the impact of other news (such as economic data releases) and more cleanly identifies the impact of exogenous policy shocks. Moreover, as Kuttner (2001) has argued, focusing on innovations to the futures price helps to strip out the impact of fluctuations in term and risk premia.

This paper also contributes to a smaller literature on the instability over time of identified impulse responses from VARs. Boivin and Giannoni (2006) test for instability in a small structural VAR, and find evidence for a structural break. Owyang and Wall (2009) estimate aggregate and regional VARs and find that the estimated impact of monetary policy on output is significantly lower in the Volcker–Greenspan period than earlier. Both papers argue that the apparent change in the impact of monetary policy shocks is a real one, reflecting fundamental changes in the transmission mechanism. Boivin and Giannoni argue that the key change is a stronger Fed response to inflation expectations. Owyang and Wall attribute the change in responsiveness to changes in the propagation mechanism for monetary policy. By contrast, our analysis suggests that although the reduction in the estimated impact reflects a real change in behavior (forecasts playing a greater role in the Fed's decision-making), the key change is to the estimated effect rather than the actual effect, because identification problems become more pronounced when the Fed's policy becomes more forward looking.

The next section briefly reviews the literature on identifying monetary policy shocks and their effects. It focuses in particular on four identification schemes that have received significant attention: Christiano et al.'s (1996) recursive VAR identification; Sims and Zha's (2006) non-recursive VAR; Bernanke and Mihov's (1998) over-identified VAR; and Romer and Romer's (2004) narrative

[1] Because the Fed Funds futures market only started trading in October 1988, we are unable to derive our shock measure for the early portion of the "great moderation". However, the results for the other identification strategies we follow in Section 2 are broadly the same whether the estimation starts in 1982, 1984 or 1988.

identification. We contrast the baseline results in the original papers with results for the recent period (focusing on the post-1988 period to allow a comparison with our new measure), and then analyze how the nature of monetary policy shocks has changed since the early 1980s as policy has become more forward-looking. Section 3 discusses the Fed Funds futures market and outlines the new shock measure. Section 4 uses the new measure to estimate the effects of monetary policy shocks in the post-1988 period, discusses the results and outlines some robustness checks. Section 5 concludes.

2. The failure of conventional identification schemes

Following Christiano et al. (1999), we identify a monetary policy shock as the orthogonal disturbance term s_t in an equation of the form

$$S_t = f(\Omega_t) + s_t \tag{1}$$

where S_t denotes the monetary stance (or more narrowly, the instrument of the monetary authority, e.g. the Fed Funds rate) and f is a linear function relating S_t to the policymaker's information set Ω_t.[2] The remainder of this section illustrates how some existing methods of identifying equation (1) appear to have broken down, and argue that this may reflect a failure to sufficiently control for the introduction of more forward-looking information into Ω_t over the last two to three decades.

The four schemes we consider are Christiano et al.'s (1996) recursive VAR approach, Bernanke and Mihov's (1998) over-identified VAR, Sims and Zah's (2006) non-recursive VAR and Romer and Romer's (2004) narrative approach. The full details of these approaches and our efforts to replicate them are detailed in the appendix. This section provides a brief overview of the results.

Estimated over their original sample periods—from the 1960s to the mid-1990s—all four approaches suggest that monetary policy shocks have an effect in line with the conventional wisdom: a monetary contraction lowers output and other real indicators over the short to medium term, and has a more muted impact—generally negative—on inflation. However, estimating the models over the more recent period yields very different results. Most worryingly, monetary contractions are estimated to have a stimulative effect on output.

2.1. Four identification schemes

Christiano et al. (1996) estimate a quarterly VAR with six variables and four lags over the period 1960Q1–1992Q4. Their results show that a contractionary shock is associated with a persistent decline in output. The price index responds slowly but eventually declines. We replicate their results and report the impulse responses of GDP and the GDP deflator to a contractionary monetary policy shock (Fig. 1 panel a).[3] However, when the same model is estimated for the recent period (1988Q4–2007Q3), neither output nor prices show the expected response (Fig. 1 panel b).[4]

Bernanke and Mihov (1998) develop a monthly model in which contemporaneous identification restrictions are imposed on monetary variables in order to model the Fed's operating procedure. As in all the monthly estimates in this paper, output is measured by industrial production. Fig. 2 (panel a) shows that in the original period the responses of industrial production and prices are as expected, and very similar to Christiano, Eichenbaum and Evans's.[5] However, when this model is estimated for the later period (panel b), again neither output nor prices show the expected response. Both output and prices increase significantly—immediately in the case of output, and over the medium term in the case of prices.[6]

Sims and Zha (2006) include a wider set of variables. We replicate their findings for their original sample (Fig. 3, panel a).[7] Results are similar to those obtained from Christiano et al.'s (1996) recursive identification scheme (the results are not significant due to the wide standard error bands obtained under the bootstrap algorithm). However, when the model is

[2] Hence Eq. (1) can be thought of as the monetary authorities' feedback rule or policy reaction function, although as Christiano et al. (1999) highlight, there are pitfalls in identifying the coefficients in $f()$.

[3] In this paper the size of the monetary policy shock is always equal to one standard deviation and impulse responses are always reported with two standard error bands. Standard errors shown in Figs. 1–4 are obtained via multivariate normal parametric bootstrapping, based on 500 replications.

[4] We end our sample in 2007 Q3 because nonborrowed reserves (NBR) become negative during the fourth quarter of 2007. Our sample is also truncated (at 2007:11) for the Bernanke and Mihov estimation for the same reason. All estimation for the recent period starts at end-1988. This is because the Fed Funds futures data that we rely on for identifying the impact of shocks in Section 4 are only available from this period onwards, and we want to ensure that the results are comparable across methods. However, the finding that the previous methodologies appear to break down for the later period is robust to starting the sample in 1982 or 1984, as already noted.

[5] Bernanke and Mihov estimate different versions of the model, including four that are over-identified and one that is just-identified. We replicate the over-identified model (Federal Funds rate targeting model) since Bernanke and Mihov find that this performs best for the post-1988 period.

[6] Although this paper re-estimates the same VAR, i.e. a monthly VAR with 13 lags and six variables (output, domestic prices, commodity prices, the Federal Funds rate, total reserves and NBR), there are some minor differences between our VAR and Bernanke and Mihov's. They interpolate GDP and the GDP Deflator to convert a quarterly series to a monthly series, while this paper uses monthly Industrial Production and CPI data instead. This paper also employs a different commodity price index. These differences are minor, and comparing the impulse responses from the original period suggests that they have no significant effect on the results.

[7] Due to data constraints, this paper excludes their bankruptcy measure from the VAR. The impulse responses of our model estimated for the original period are almost identical to those in Sims and Zha (2006). In fact Sims and Zha mention that the measure of bankruptcy makes “only a modest contribution” to the results, while Christiano et al. (1999) also re-estimate the Sims and Zha model excluding the bankruptcy measure. Having said this, our confidence intervals are somewhat wider than those reported by Sims and Zha: this is partly cosmetic (they report 68%, or approximately one standard error, CIs, whereas we report two standard error CIs); it may also reflect the exclusion of the bankruptcy measure in our estimates, or possibly differences in the bootstrap algorithms.

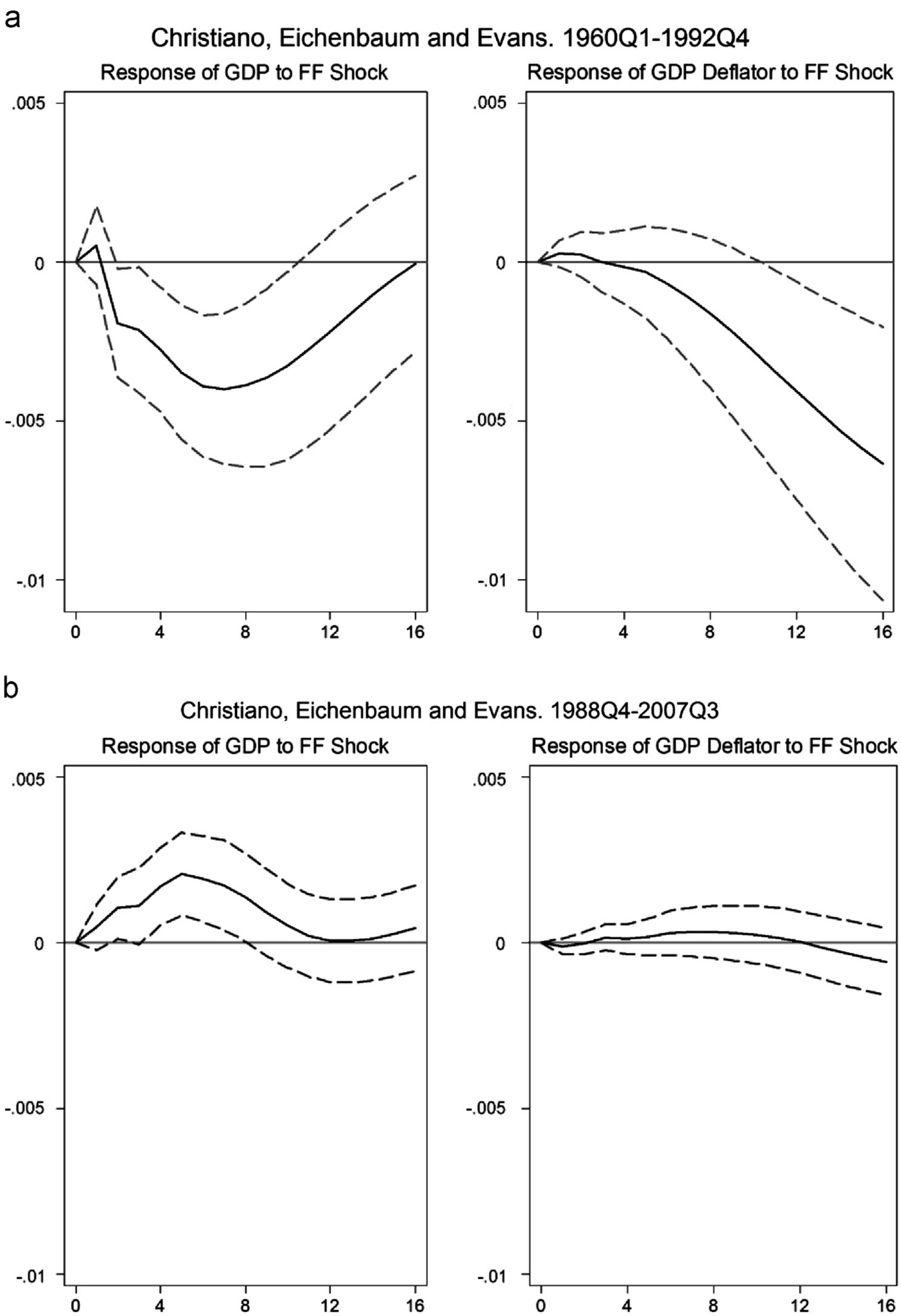

Fig. 1. Structural VAR (quarterly data, 6 endogenous variables plus constant and linear time trend, 4 lags) as described in the text. Variables ordered as GDP, GDP deflator, commodity prices, non-borrowed reserves, Fed Funds rate, total reserves. All variables except for the Fed Funds rate are in logs and seasonally adjusted. Graphs show response of GDP and GDP deflator to a one standard deviation positive shock to the Fed Funds rate. Structural shocks obtained via Cholesky decomposition. Two Standard Error bands produced by parametric bootstrapping (500 replications).

estimated for the 1988:Q4–2008:Q2 period (panel b), the impulse responses are very different. After a contractionary monetary policy shock, output increases significantly over the medium term.

Romer and Romer (2004) argue that VAR identification schemes fail to control for anticipated monetary policy changes and for deviations between desired and actual changes due to endogenous movements in monetary instruments, and develop a narrative approach that seeks to overcome these problems. Romer and Romer estimate a monthly VAR with three variables: the log of industrial production, log PPI for finished goods and a measure of the monetary policy shock derived through their narrative method. Their results replicate those of the VAR identification schemes, although the estimated effect of monetary policy is stronger

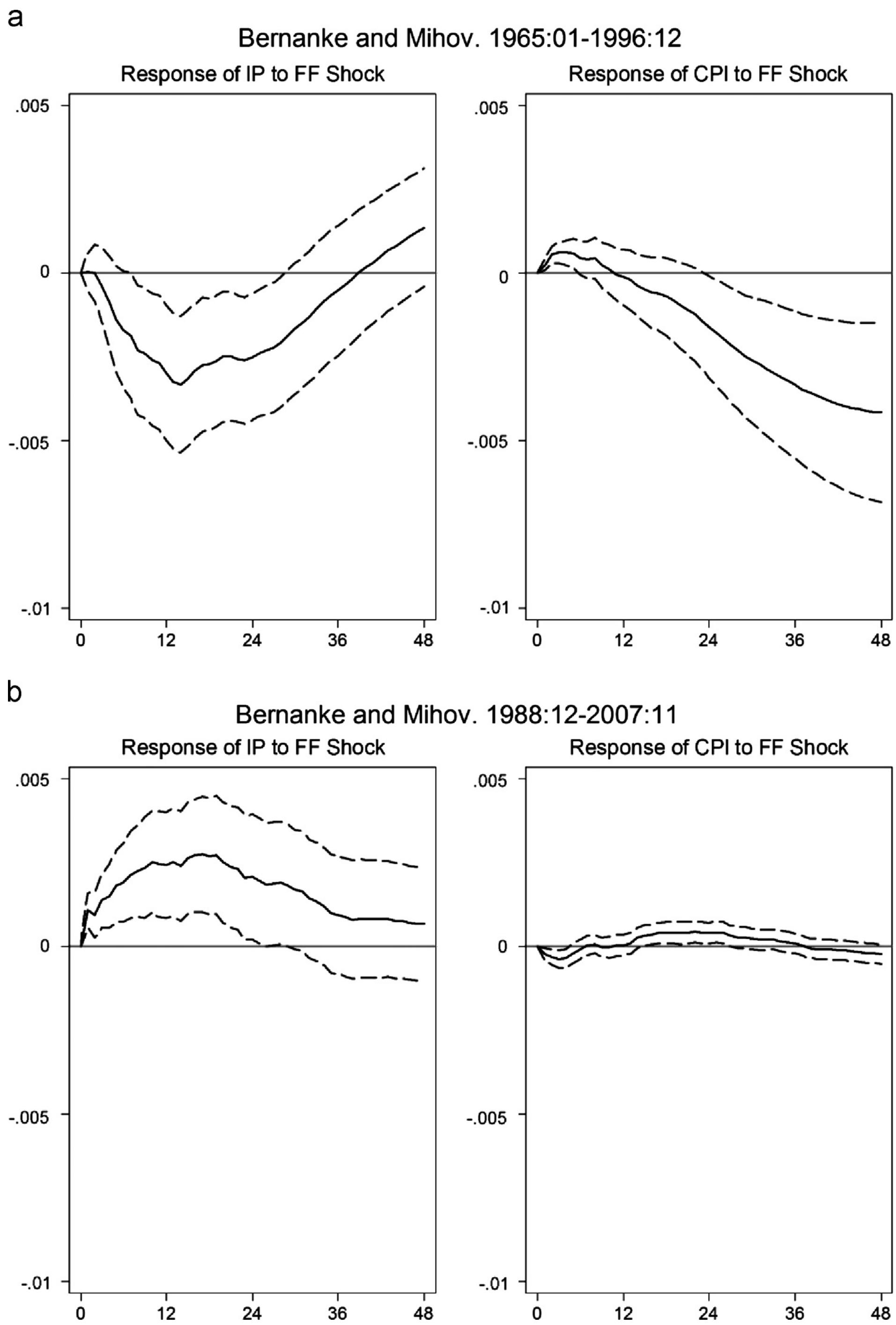

Fig. 2. Structural VAR (monthly data, 6 endogenous variables plus constant and linear time trend, 13 lags) as described in the text. Variables include industrial production, consumer price index, commodity prices, Fed Funds rate, total reserves, non-borrowed reserves. The first 3 variables are in logs and seasonally adjusted. The last two variables are seasonally adjusted and normalized by dividing by the 36-month moving average of total reserves. Graphs show response of output and CPI to a one standard deviation positive shock to the Fed Funds rate. Structural Shocks obtained by imposing the structural decomposition discussed in the text (1 overidentifying restriction) Two Standard Error bands produced by parametric bootstrapping (500 replications).

and quicker than for the VAR identification schemes. Fig. 4 (panel a) illustrates their findings. However, when this model is estimated for the period 1988:12–2008:06, the estimated impulse responses are different, especially for output (panel b).[8]

[8] See the data Appendix for information on how the Romer and Romer index was extended to 2008.

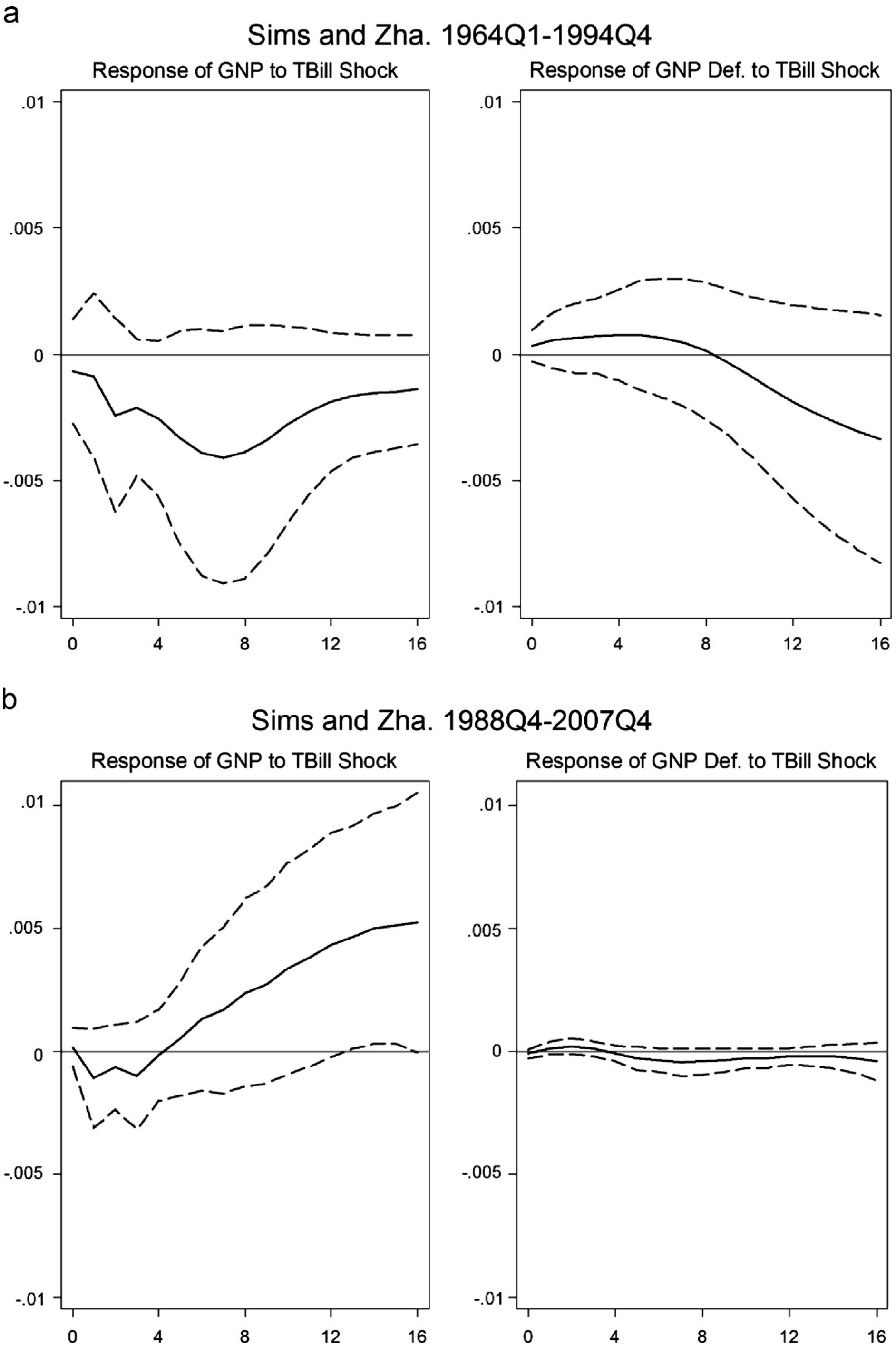

Fig. 3. Structural VAR (Quarterly data, 7 endogenous variables plus constant and linear time trend, 4 lags) as described in text. Variables include Crude Materials Prices, M2, T Bill Rate, Intermediate Materials Prices, GNP Deflator, Wages (private sector workers) and GNP. All variables except the T Bill Rate are in logs and seasonally adjusted. Graphs show response of GNP and GNP Deflator to a one standard deviation positive shock to the T Bill Rate. Structural Shocks obtained by imposing the structural decomposition discussed in the text (2 overidentifying restrictions). Two Standard Error bands produced by parametric bootstrapping (500 replications).

What can one take from these findings? The overall message is that the existing identification schemes lead to estimated impulse responses in the post-1988 sample that are both different to those found in the earlier samples, and counter to what most central bankers would find plausible. However, in our view there are good reasons to doubt the robustness of these empirical results. Several identification problems are likely to have become particularly acute for the recent period.

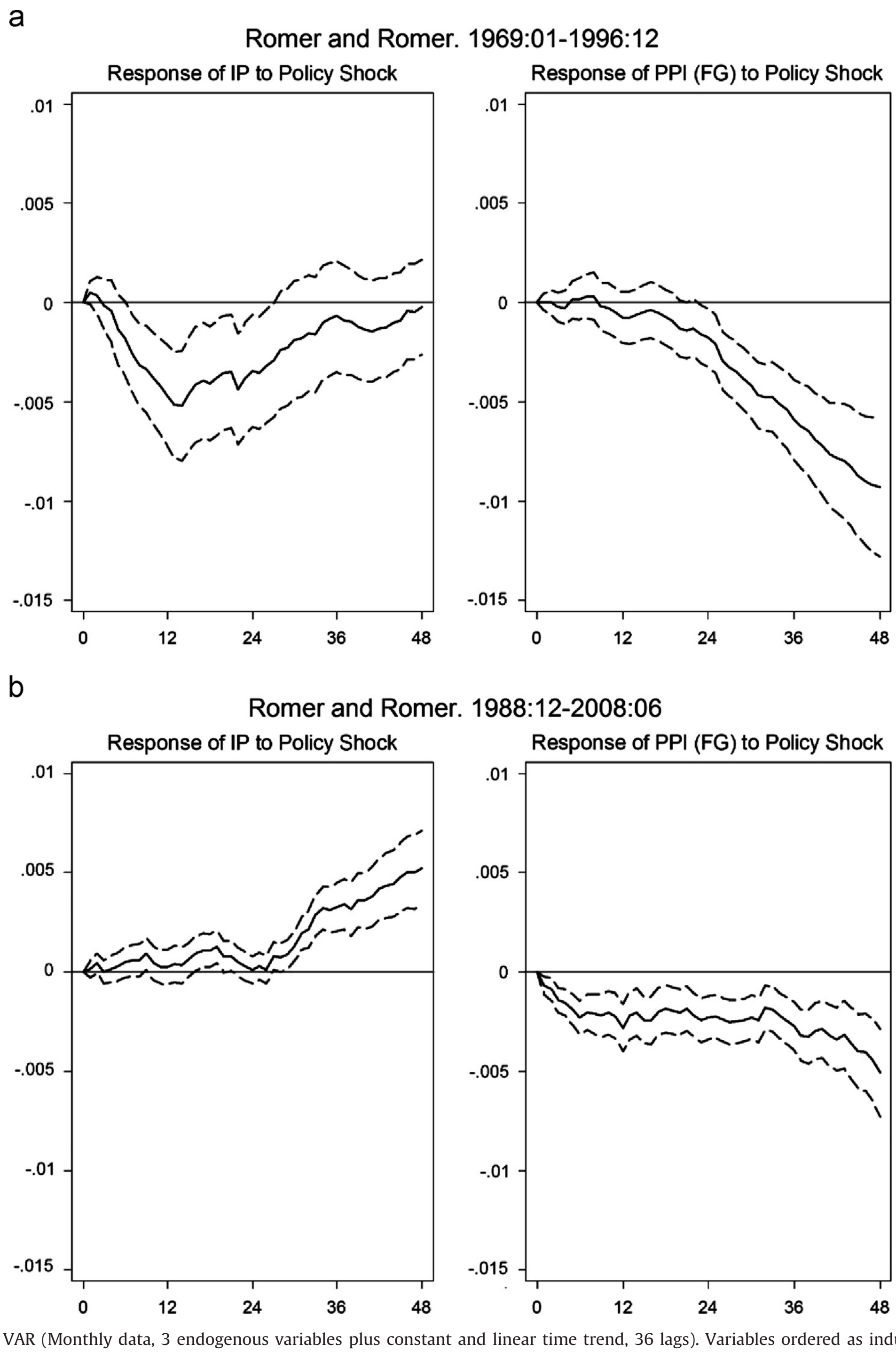

Fig. 4. Structural VAR (Monthly data, 3 endogenous variables plus constant and linear time trend, 36 lags). Variables ordered as industrial production, producer price index (finished goods), both seasonally adjusted and in logs, and Romer and Romer's shock measure, cumulated. Graphs show response of industrial production and PPI (finished goods) to a one standard deviation positive shock to the policy measure. Structural shocks obtained via Cholesky decomposition. Two Standard Error bands produced by parametric bootstrapping (500 replications).

2.2. Identifying policy shocks under changing policy regimes

To shed some light on these issues, this section estimates the Romer and Romer (2004) policy regression over three subsamples, chosen based on the policy regime in place at the time, in order to assess the stability of the parameters on the different elements of the Fed's information set. The principal changes to policy regime took place in late 1979, when the Fed started to target monetary aggregates under chairman Paul Volcker, and late 1982, when the Fed moved back towards targeting the Federal Funds rate. The three subsamples are therefore 1969:1–1979:10 (pre-Volcker); 1979:11–1982:10 (nonborrowed reserves targeting); and 1982:11–2008:06 (Fed Funds rate targeting).[9]

Our first step is to analyze the stability of the regression coefficients via a series of Chow tests comparing each set of adjoining subsamples. There is clear evidence of a structural break at both potential break points. This suggests that Romer and Romer's reaction function, that assumes constant coefficients across the whole sample, could be misspecified.[10] These results are in line with those of Boivin and Giannoni (2006), who undertake a similar exercise for a small structural VAR similar to the systems discussed in Section 2, and find strong evidence for a structural break. Hence, the VAR identification methods discussed above—which like Romer and Romer's method assume time-invariant coefficients in the policy reaction function in order to identify monetary policy shocks—are likely to suffer from very similar problems.

The second step is to test whether specific elements of Ω_t have changed. The focus is on two sets of variables: the eight forward-looking variables (1- and 2-quarter ahead forecasts) and nine backwards-looking variables (current and last quarter estimates) included in Romer and Romer's specification, comparing the post-1988 period with the rest of the sample. Table A2 in the online appendix presents F tests of the joint significance of the variables for the two subsamples. Policymaking appears to be unambiguously forward-looking in the post-1988 period, but one cannot reject the null hypothesis of no forward-looking variables in Ω_t during the pre-1988 period. This finding corroborates other analyses of Fed policymaking over the period (Orphanides, 2003; Boivin and Giannoni, 2006).

These results shed some light on the findings presented in Section 2. Failure to allow for structural breaks—under all four methods of identification—will tend to give biased estimates of the shocks themselves, and hence biased estimates of the impact of the shock on other macroeconomic variables. For instance, by increasing the measurement error associated with the Romer and Romer shock series, it will lead to attenuation (bias toward zero) in the shocks' estimated macroeconomic impact.

The fact that policymaking appears to have become more forward looking in recent years has particularly serious implications for the VAR identification methods, since these do not include any forward-looking elements in Ω_t. If Fed policymakers react to an expected increase in output growth above the economy's potential by tightening monetary policy to partially offset it, then a monetary contraction will appear to cause higher growth if these anticipatory movements are not explicitly allowed for. Since anticipatory movements appear to have become more important for the recent period than earlier, this might explain why VAR identification methods identify the expected contractionary impact of monetary tightening for the earlier period, but for the later period generate the counterintuitive expansionary effects shown in Section 2. Although Romer and Romer's methodology attempts to control for anticipatory movements, by imposing equal coefficients throughout the sample it may not adequately capture the stronger effects in the recent period.

3. A new Fed Funds futures-based shock measure

Conventional methods of identifying monetary policy shocks—which require the estimation of (1) with suitable proxies for Ω_t—will perform badly if either Ω_t or $f()$ are misspecified. An alternative approach is to use financial market data to obtain the private sector's contemporaneous beliefs about $f(\Omega_t)$ at the time of each meeting, and use these as a proxy for the true reaction function and its elements. This circumvents the need to estimate $f(\Omega_t)$ directly, and therefore does not require that we impose restrictions on the variables in Ω_t or the functional form $f()$.

3.1. Overview

To illustrate this approach in general terms, assume that there are two measures of the private sector's expectation for the policy stance S_t for a particular policy meeting: one in the immediate run-up to the meeting, ${}_{t-1}\widehat{S}_t$, and one immediately after the announcement of the policy stance decided at the meeting, ${}_t\widehat{S}_t$. Each is a noisy measure of the private sector's true expectation:

$$ {}_{t-1}\widehat{S}_t = E^P_{t-1}[S_t] + \xi_{t-1} = E^P_{t-1}[f(\Omega_t)] + \xi_{t-1} \tag{2} $$

$$ {}_t\widehat{S}_t = E^P_t[S_t] + \xi_t = S_t + \xi_t \tag{3} $$

where the private sector's actual expectations at time τ of the stance at time t are denoted by $E^P_\tau[S_t]$. The noise ξ can arise from several sources, including time-varying risk premia as well as measurement or rounding errors. We make the following two identifying assumptions:

$$ E^P_{t-1}[f(\Omega_t)] - f(\Omega_t) = 0 \tag{4} $$

[9] Bagliano and Favero (1998) identify five regimes. We extend the last period from 1996:3 and start the first period in 1969:1 rather than 1966:1, reflecting the coverage of the original Romer and Romer series. We also combine their first two and last two periods, as we do not find the distinction meaningful in either case.

[10] See Table A1 in the online appendix, which also presents a test of a structural break at the end of 1988, matching the subsample with available Fed Funds futures data, that suggests that our sample is broadly representative of the Fed Funds rate targeting period as a whole.

$$\xi_t - \xi_{t-1} = 0 \tag{5}$$

The first assumption (4) states that the private sector's beliefs prior to the announcement about the Fed's information set are correct.[11] The second assumption (5) states that the noise term is unchanged around the time of the policy announcement. Then, subtracting (2) from (3) yields

$$_t\widehat{S}_t - {}_{t-1}\widehat{S}_t = s_t \tag{6}$$

This implies that a suitable proxy for the shock, s_t, is given by the change in the measure of the private sector's beliefs about the policy stance around the time of a policy announcement, $_t\widehat{S}_t - {}_{t-1}\widehat{S}_t$.

3.2. Fed funds futures data

Our measures of the private sector's beliefs about the policy stance $\widehat{S}_t$ are derived from Fed Funds futures contracts. The Federal Funds futures market was established at the Chicago Board of Trade (CBOT) in October 1988 (see Söderström, 2001, Kuttner, 2001 and Faust et al., 2004 for further information). The price of a contract for month $m+h$ (i.e. at a horizon h from the current month m) is a bet on the monthly average effective Fed Funds rate in month $m+h$ (denoted $\overline{r}^e_{m+h}$). Note that the average target Fed Funds rate ($\overline{r}_{m+h}$) might differ from the effective rate due to targeting errors on the part of the Fed:

$$\overline{r}^e_{m+h} = \overline{r}_{m+h} + \varepsilon_{m+h} \tag{7}$$

These errors are typically small and mean zero. For a given contract price p^h_d on day d in month m, the futures rate f^h_d is simply given by $1-p^h_d$. Then standard no-arbitrage conditions imply that the futures rate is equal to the average effective Fed Funds rate in month $m+h$, $E_d\overline{r}^e_{m+h}$, plus a risk (or hedging or term) premium ρ^h_d:

$$f^h_d = E_d\overline{r}^e_{m+h} + \rho^h_d \tag{8}$$

Assuming that the risk premium ρ^h_d remains constant and that there is also no change in the expected average targeting error $E_d[\varepsilon_{m+h}]$, then the change in the expected target rate during subsequent calendar months ($h \geq 1$) following a policy announcement on day d of month m is given by

$$\Delta E_d\overline{r}_{m+h} = f^h_d - f^h_{d-1} \tag{9}$$

while the change for the remainder of the current month (whose length is M days) is given by

$$\Delta E_d\overline{r}_m = \frac{M}{M-d}\left(f^0_d - f^0_{d-1}\right) \tag{10}$$

The innovation to the expected target rate in a given month then serves as a good proxy for the underlying monetary policy shock s_t under four assumptions. First, the average target rate $\overline{r}_{m+h}$ should be correlated with the policy stance S_t. If this holds then $f^h_d - f^h_{d-1}$ provides an estimate of $_t\widehat{S}_t - {}_{t-1}\widehat{S}_t$, while the noise term ξ_t is given by the sum of the risk premium ρ^h_d, the expected Fed targeting error $E_d[\varepsilon_{m+h}]$ as well as data errors. Second, there should be no predictable changes in the noise terms that make up ξ_t, e.g. due to predictable effects of policy announcements on risk premia: this is a necessary condition for (5) to hold. Third, there should be no other 'news' that might affect the expected futures rate (such as macroeconomic data announcements that might have implications for rate changes in the future) during the 24-hour period associated with the policy decision. Last, the policy announcement itself should not reveal information about the Fed's private information set Ω_t or its reaction function $f()$. These last two assumptions are necessary for (4) to hold.[12] Assuming that these assumptions are valid, then the policy 'surprise' is a good measure of the shock. The evidence, discussed in Section 3.4, provides strong support for the first three assumptions, while evidence on the fourth is more mixed.

Following Kuttner (2001), the impact of policy announcements (or non-announcements) following FOMC meetings is estimated by comparing the end of day price on the day following the (last) day of the meeting with that on the meeting day for meetings occurring before February 1994, and comparing the price on the day of the meeting with that on the day before the meeting for subsequent meetings. Our analysis focuses only on FOMC meeting dates, rather than on all dates that the Fed announced changes to the target Fed Funds rate, including inter-meeting changes.[13]

[11] These assumptions are stated in their strongest form to clarify the exposition. A weaker assumption would be that, conditional on the realization of Ω_t and s_t, (4) and (5) hold in expectations. A more serious problem—simultaneity bias—will arise if (4) and (5) do not hold even in their weaker, conditional expectations, form, e.g. because the private sector makes systematic errors in forecasting the Fed's policy reaction function. This issue is addressed in more detail later in the paper.

[12] For instance, a negative macroeconomic news release (one that tends to revise down output and inflation expectations) that occurred concurrently with a policy announcement would imply lower rates in the future, implying that (4) is contradicted. Similarly, if a policy announcement provides new information about the Fed's information set, e.g. so that a rate cut signals that the Fed expects a recession, then the private sector's beliefs prior to the announcement were incorrect and again (4) does not hold.

[13] Like Faust et al. (2004), who also focus on regular policy announcements, we believe that intermeeting changes are more likely to be associated with the simultaneous release of macroeconomic information rather than reflecting exogenous shocks to policy.

3.3. Constructing the shock series

The simplest signal of the policy stance S_t is the futures rate for the current month, f_d^0. However, we argue that there are several reasons to focus on a range of maturities. First, combining the information from several sources—essentially taking a sample mean of the shock measures obtained from contracts at different horizons—should help to minimize the effect of noise in a specific contract. This averaging may be particularly important since the risk premium is likely to be more volatile at shorter horizons (as shown in the data appendix, the market for the current month contract is not the most liquid, and intra-month trading volumes are in fact particularly volatile for this contract, which could lead to a more volatile liquidity premium and hence introduce more noise into the shock measure). Moreover, since the Fed's policy decisions are relatively persistent over time, a policy change in the current period will be reflected in higher expected rates several months ahead, so that futures contracts settling several months in the future will also contain information about the current shock. Indeed, shocks which are expected to be permanent might be expected to have a greater impact on the economy. But some shocks to current rates might have little impact on longer term expectations (for instance, if the shock was to the immediate timing of the rate change rather than to the long-term direction of rates, as Gürkaynak, 2005 argues). Hence, a measure of shocks that combines the innovations to rates in the current (spot) month with those anticipated in the future is likely a better measure of the overall policy stance. While contracts are now available for more than a year into the future, longer-dated contracts have not been available for the whole period and even now are typically relatively illiquid. Hence, we focus on contracts for the current month and up to 5 months ahead.

In order to combine the information available in the estimated forecast innovations at all six horizons, one can estimate a simple factor model via maximum likelihood. Denoting the vector of innovations at the six horizons (normalised to have mean zero and variance of one) as **s**, the vector of factors as ϕ, the factor loading matrix as Λ and the vector of unique factors as **e**, the factor model is given by

$$\mathbf{s} = \boldsymbol{\phi}\Lambda' + \mathbf{e} \tag{11}$$

This method of extracting the common shocks in the contract prices at different horizons has several advantages. While one is principally interested in extracting the common levels shock (which captures unexpected policy tightening or loosening), because more than one factor is extracted one we can also potentially analyze shocks to the term structure. The method is well suited to the data, which includes futures contracts with differing levels of liquidity and hence volatility. The prices of some contracts will tend to comove more than others, reflecting a lower unique variance for these series. The factor model allows one to capture this explicitly, putting more weight on those series that exhibit a greater degree of comovement in extracting the factors. At the same time, this method is relatively simple and does not require us to formulate and estimate a fully specified model of the term structure. Two factors adequately capture the information in the futures shocks.[14] The two factors summarize the new information on the medium term evolution of policy rates that is revealed by the policy rate announcement. Indeed, the factors turn out to have an intuitive interpretation. The first factor, which is highly positively correlated with all the individual innovations, can be thought of as a levels effect: that portion of the new information related to the policy announcement that causes vertical shifts in the expected medium-term trajectory for policy rates. Since the transmission of monetary policy is generally thought to occur via the impact of short rate changes on longer term (real) rates, it is this portion of the new information on rates that corresponds most closely to the relevant policy shock. We therefore use this factor as our measure of the underlying policy shock.

The second factor, whose correlation with the individual innovation series at different maturities decreases monotonically from positive to negative as the maturity increases, can be thought of as a slope or yield curve effect: that portion of the new information relating to the policy announcement that leads to differential effects on expected policy rates in the near term and further out. While this factor captures an important portion of the news relating to policy announcements, it does not capture the notion of a policy shock that is the focus of the current paper.

3.4. Assessing the shock series

Our new shock series is presented in Fig. 5. Our factor-based shock measure has a mean of 0 and a standard deviation of 1 by construction. To aid interpretation, in Fig. 5 it is scaled to be a weighted average of the deviations from the mean of the six underlying monthly “shock” series. Two standard deviation bars are shown, and the 27 June 2001 meeting is indicated by a vertical bar to aid the discussion in Section 3.5.

The validity of our shock measure depends on the validity of the underlying assumptions. The first assumption, that the Fed Funds target rate at the relevant horizons (0–5 months) is correlated with the ‘true’ monetary stance, seems uncontroversial. Bernanke and Mihov (1998) have demonstrated that a Fed Funds targeting model best describes monetary policy in the post-1988 period, while it is intuitive that, in an economy with forward-looking agents making irreversible

[14] Estimating a principal factor model with up to six factors, the first factor accounts for 92% of the total variance, the second factor for a further 9%, and the third factor for 0.4%. The eigenvalues of the first three factors are 5.2, 0.52 and 0.02, respectively (the last three factors have negative eigenvalues and make a cumulative contribution to the variance of -1%). Hence, a model with two factors appears to adequately and parsimoniously capture the main patterns of correlation in the data, and it is this parsimonious specification that is then estimated via Maximum Likelihood. Tables A3 and A4 in the online appendix present further details of the factor model and the estimated shock series.

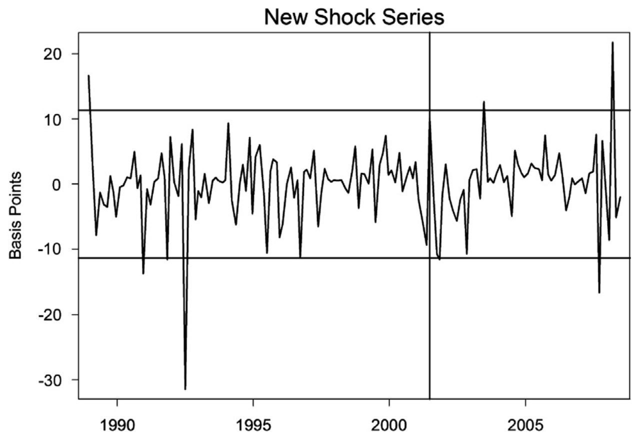

Fig. 5. New shock series, in basis points. To make it comparable in size to the 6 underlying shocks, the first factor (SD = 1 by construction) is divided by the sum of the 6 coefficients from the factor model. Two standard error bands shown by horizontal lines; vertical line identifies the June 2001 FOMC meeting discussed in Section 3.5.

economic decisions, the overall stance of policy depends not only on the current target rate but also on the rates expected in the immediate future. With respect to the second assumption—that there should be no predictable innovations to the noise component of the private sector's expectations about the policy stance in the short run—Piazzesi and Swanson (2008) show that anticipated changes to risk premia in the Fed Funds futures market occur mainly at business cycle frequency. With respect to the third assumption—that other information that could be conflated with the policy announcement and bias our results is not released on the same day—Gürkaynak et al. (2005) show that some FOMC meeting and intermeeting dates associated with policy announcements coincide with macroeconomic data releases. However, they show that only *Employment Report* releases have any independent effect on Fed Funds futures. Bernanke and Kuttner (2005) identify ten observations, all before 1994, for which *Employment Report* releases coincide with policy announcements or FOMC meetings. But our decision to focus only on FOMC meetings helps to alleviate this problem, since only three of these dates coincide with FOMC meetings (the others coincide with intermeeting changes).[15] We provide some empirical evidence that the inclusion of these dates is not driving our results in the robustness checks in Section 4.2.

To test the fourth assumption, one can regress our (scaled) shock measure on the difference between the Fed's Greenbook forecasts and high-quality private sector (*Blue Chip*) forecasts for the 17 variables used in Romer and Romer's (2004) estimated reaction function, where this difference is used as a proxy for the Fed's internal information. Since the Greenbook forecasts are only made public with a 5-year lag, the shock measure should only be correlated with the Fed's internal information to the extent that the latter is revealed indirectly by the policy rate, the announcement and any related communication. As we show in Table 1, the joint hypothesis of zero coefficients on all 17 variables cannot be rejected at the 10% level. This suggests that our shock measure should be relatively uncorrelated with the Fed's exclusive information, and simultaneity bias should therefore not be a significant problem.

However, an inspection of the coefficient estimates in Table 1 points to evidence that our shock measure may be contaminated by the impact of the Fed tightening policy in response to near term output and price pressures, since our shock measure responds positively to current quarter output and inflation forecasts. We investigate further the implications of this for our results in Section 4.3.

To illustrate how our shock measure compares to others in the literature, Table 2 presents correlation coefficients for our shock measure (New), the change in the target Federal Funds rate (ΔFF) and Romer and Romer's shock measure (*R&R*; all on a per-meeting basis, for 157 meetings); the final row presents correlation coefficients between the per-quarter average of these three measures and the monetary policy shock obtained from a Cholesky decomposition of Christiano, Eichenbaum and Evans's quarterly VAR specification (*CEE*), for 76 quarterly observations (1988Q4–2007Q3). Our new shock measure is positively and significantly correlated with all three measures (at least at the 10% level).

3.5. *Our new shock series: an illustrative observation*

Our shock measure, although correlated with existing measures, can differ significantly from these for some observations. These differences can help illustrate some of the relative strengths of our approach. For instance, the FOMC decided at its 26-27 June 2001 meeting on a 25 basis points reduction in the Fed Funds rate. The cut followed five successive 50 basis point cuts (three at the three preceding meetings and two cuts between meetings), as part of a rate-cutting cycle that saw the Fed Funds rate fall from 6.5% to 1.75% over the course of the year. While the VAR and narrative identification

[15] The three dates in question are 7 July 1989 and 2 July 1992 (the day after the meeting), and 4 February 1994 (the day of the meeting).

Table 1
Regression results and *F*-test statistics for policy shock measure and Greenbook variables.

Variable	Coefficient
Unemployment_0	−4.26
$\text{Output Growth}_{-1}$	−1.31
Output Growth_0	2.37***
Output Growth_1	−0.783
Output Growth_2	1.19
GDP Deflator_{-1}	−0.92
GDP Deflator_0	2.34**
GDP Deflator_1	−1.49
GDP Deflator_2	−0.323
$\Delta\text{OutputGrowth}_{-1}$	0.541
$\Delta\text{OutputGrowth}_0$	−1.14
$\Delta\text{OutputGrowth}_1$	0.803
$\Delta\text{OutputGrowth}_2$	−1.44
$\Delta\text{GDP Deflator}_{-1}$	0.300
$\Delta\text{GDP Deflator}_0$	−1.31
$\Delta\text{GDP Deflator}_1$	−0.117
$\Delta\text{GDP Deflator}_2$	1.22
Constant	−0.610
R^2	0.185
$F(17)$	1.50
p-value	0.132

The dependent variable is the scaled shock measure in basis points; the independent variables are the difference between the Greenbook and Blue Chip forecasts for the 17 variables identified by Romer and Romer (variables are estimates for the previous or current quarter or forecasts one or two quarters ahead, except for variables denoted "Δ" which are the change in the forecast from the previous meeting; all variables are then differenced between the Greenbook and Blue Chip consensus forecasts). The regression is run over 113 FOMC meetings between 1988 and 2002. The *F*-test statistic shown is for the joint null hypothesis that the coefficient on all 17 variables is zero. Standard errors are robust to heteroskedasticity (but are omitted from the table for brevity). *10% level of significance.

** 5% level of significance.

*** 1% level of significance.

Table 2
Correlation between shock measures.

	New	Δ*FF*	R&R	CEE
New	1			
Δ*FF*	0.39***	1		
R&R	0.23***	0.73***	1	
CEE	0.22*	0.26**	0.09	1

Correlation coefficients for our new shock measure (*New*) and existing measures: the change in Fed Funds Rate (Δ*FF*), Romer and Romer's narrative measure (*R&R*), and Christiano, Eichenbaum and Evans's measure (*CEE*; based on Cholesky decomposition of VAR residuals). Coefficients in rows 1–3 based on 157 per-meeting values; coefficients in last row based on 76 quarterly values.

* 10% level of significance.

** 5% level of significance.

*** 1% level of significance.

methods identify a negative policy shock, it is clear from reading the Fed's statements as well as from market reaction that the Fed's interest rate cuts were largely an endogenous response to the economic slowdown in the wake of the bursting tech bubble and concerns that the economy was set to slow further.

By comparison, our shock measure is large and positive (almost 2 standard deviation bands above zero, or 10 basis points when suitably scaled). The intuition for this is that market participants were anticipating another 50 basis point cut in rates. Reuters (June 28) reports that "the market had priced in the prospect for 50 basis points." The smaller cut therefore represented a positive shock to Fed Funds rate expectations. Market reaction to the move supports our interpretation of the June 27 rate cut as a policy tightening. Reuters (June 27) reports that "the dollar climbed to a 10-week high on the yen on Thursday, helped by a raft of factors, including the… rate cut." The dollar also gained ground against the euro. Meanwhile, bond yields rose significantly (particularly for two-year government paper). These reactions are more consistent with a contractionary than an expansionary monetary policy shock.

4. Identifying the effect of monetary policy shocks

Following Romer and Romer, we identify the effect of monetary policy shocks using a small 3 variable monthly VAR (as they do, we let the shocks series take a value of zero for months without FOMC meetings).[16] The variables are ordered so that monetary policy is allowed to respond to, but not affect, output and inflation contemporaneously. We use the log of industrial production as our measure of output and the log consumer price index as our measure of prices.[17] As with Romer and Romer's shock measure, our measure captures unanticipated *changes* in policy rates. Hence, like Romer and Romer, we enter our shock measure cumulated in the VAR, since here it is the level, not the change, in policy that is the appropriate variable.[18] The baseline VAR includes 36 monthly lags. However, the results are fully robust to shorter lag specifications that match the kind of lag structure in the other VAR results cited in Section 2 and make fewer demands on the data given the relatively short sample available. Results for 6, 12 and 24 months, which are almost identical to the baseline impulse responses, are discussed in Section 4.2.

4.1. Baseline results

Impulse response functions are shown in Fig. 6. We show a 95% confidence interval estimated using a system bootstrap of the VAR and factor model (to deal with the generated regressor problem). After almost one year, a contractionary monetary policy shock shows a sustained negative effect on output that has its maximum impact at a horizon of around two years. The output response is very similar to the baseline results for the earlier period reviewed in Section 2 (although with greater persistence), but very different from the results obtained for the 1988–2008 period using the same methodologies.

The response of prices to a monetary contraction is more problematic. The effect becomes significantly negative only after four years; the positive response over the medium term, although small, contrasts with the negative effect that has generally been found in the literature. Castelnuovo and Surico (2006), like Hanson (2004), find evidence that the price puzzle is limited to the pre-1979 period, arguing that this is due to equilibrium indeterminacy when monetary policy responds weakly to inflation expectations, and that the inclusion of a variable capturing the persistence of expected inflation under indeterminacy can eliminate the price puzzle. However, our baseline results suggest evidence for a price puzzle even in the post-Volcker period, when the reaction of interest rates to expected inflation should be sufficiently strong to guarantee equilibrium determinacy. Other studies (e.g. Christiano et al., 1996) have included a measure of commodity prices as a means of eliminating the price puzzle (although their argument for including this variable, that commodity prices help to forecast inflation, has been criticized by Hanson, 2004).[19] In the following section we add a proxy for inflation expectations and a commodity price index to our baseline VAR as two of a series of robustness checks; neither helps to resolve the price puzzle. However, this apparently robust finding of a significant price puzzle is consistent with other recent work that uses Fed Funds futures to identify policy shocks (Thapar, 2008).

Received wisdom about the “great moderation” period is that less pronounced monetary policy shocks helped to contribute to the general moderating in macroeconomic volatility. In order to shed some light on this issue, we analyze the percentage of the forecast error variances of output and prices which can be attributed to our shock measure and two other measures over the recent period, a Federal Funds rate shock and the Romer and Romer shock (Fig. 7).[20] Results for our shock measure are shown with a solid line; results for Fed Funds rate shock (dashed line) and Romer and Romer shock (dotted line) are shown for comparison; two standard error bands for our shock measure are also shown.

The estimated impact of monetary policy shocks on the variance of the price level is broadly similar across the three measures, although the Romer and Romer method identifies the largest effect, particularly at longer horizons, which is intuitive given the impulse response shown in Fig. 4. However, at horizons of more than two years the estimated impact on output volatility is considerably higher for our shock measure—around 2 times as high as with either of the alternative measures. In fact, the results using our new measure suggest that almost half of forecast error variance at horizons of around 3 years can be accounted for by monetary policy shocks. Hence, while monetary policy shocks may have moderated in absolute terms, their relative contribution to output volatility in recent years may need to be reassessed.

---

[16] Romer and Romer's baseline specification employs a single equation approach. We apply this methodology as one of a series of robustness checks in Section 4.2

[17] This follows much of the literature, but differs from Romer and Romer (2004) who use the log of the producer price index for finished goods as their price measure. Our VARs also include an exogenous time trend.

[18] An additional rationale for using the cumulated shock series, which is I(1) by construction, is that the output and price series are generally considered I(1); hence, if the I(0) shock series were included the VAR would be statistically unbalanced, leading to nonstationary, highly persistent, residuals. Including the I(1) cumulated series allows for implicit cointegration between the variables in the VAR.

[19] Giordani (2004) argues that the price puzzle arises because the VAR system, by including output rather than the output gap (which enters in theoretical models), is misspecified. However, since our VAR model includes a linear time trend, we are in effect dealing with an output gap measure, assuming that (log) potential output follows a linear trend. This explanation is therefore unlikely to account for the estimated price puzzle in our model.

[20] In order to make the results comparable, we estimate in each case a small recursive VAR including industrial production, CPI and one of three variables: the Federal Funds rate, the Romer and Romer (cumulated) shock measure and our (cumulated) shock measure. The sample period is 1988:12–2008:06. This approach is similar to that of Romer and Romer (2004), who estimate the first two VARs to compare impulse responses using their shock measure with those using a standard recursive VAR shock measure (with the Fed Funds rate as the monetary instrument). However, we use CPI as our price measure, whereas Romer and Romer use the PPI for finished goods.

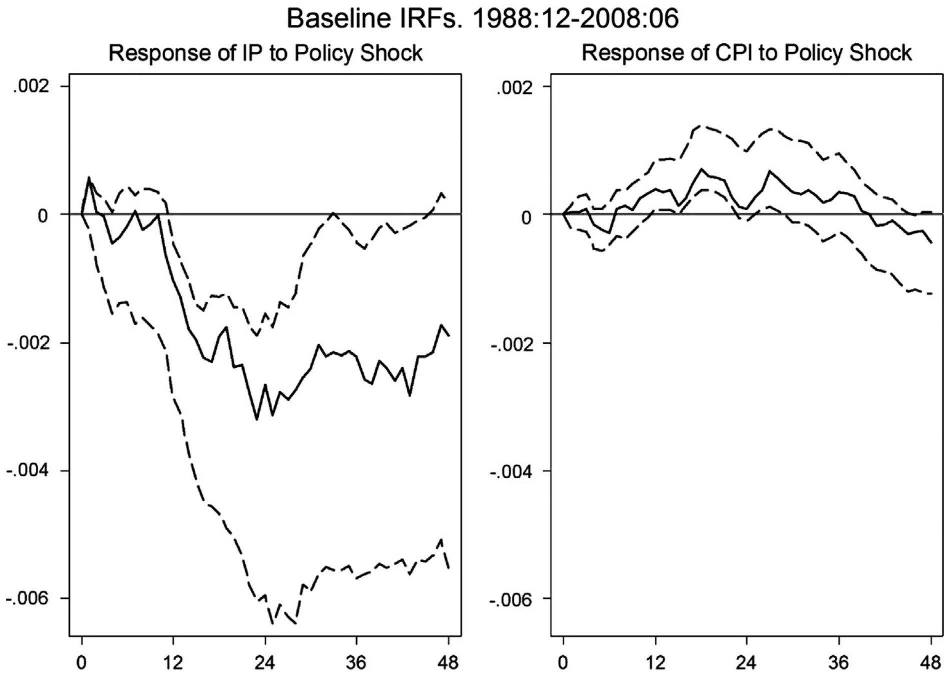

Fig. 6. Structural VAR (Monthly data, 3 endogenous variables plus constant and linear time trend, 36 lags). Variables ordered as industrial production, consumer prices, both seasonally adjusted and in logs, and our shock measure, cumulated. Graphs show response of industrial production and CPI to a one standard deviation positive shock to the policy measure. Structural shocks obtained via Cholesky decomposition. 95% confidence intervals produced by bootstrapping the combined VAR and factor model system (500 replications).

4.2. Robustness

We report here results for eight robustness checks and one further comparison. The first changes the ordering in our baseline VAR, allowing our monetary policy shock to have an instantaneous impact on output and prices. Impulse responses remain qualitatively identical, although the price puzzle is more pronounced. The second uses Romer and Romer's price measure (PPI for finished goods). Again, the only (modest) difference is in the strength and persistence of the price puzzle. The third modifies the lag structure to include 6, 12 or 24 lags rather than 36. The estimated impulse responses are essentially unchanged. The fourth assesses subsample stability by estimating the baseline VAR (with lag length reduced to 12 in light of the shorter sample) for two truncated time periods, dropping pre-1990 or post-2001 data. Results are qualitatively identical to those for the sample as a whole.

As a fifth robustness check, we include a commodity price index, ordered first in the recursive VAR. As already discussed, this has helped to eliminate the price puzzle in some studies. However, the price puzzle remains, while the output response to the policy shock is unchanged. The sixth exercise includes a measure of inflationary expectations to test the robustness of the price puzzle. Following Castelnuovo and Surico (2006), we use one quarter ahead expected inflation from the Fed's Greenbook (replaced by the corresponding *Blue Chip* forecast for 2003 onwards), and order this variable first in the recursive VAR. This exercise does not help to eliminate the price puzzle either, and the output response is also unaffected. The seventh robustness check assesses whether the inclusion of FOMC meeting dates that coincide with *Employment Report* releases is critical to the results, by including dummies for these meeting dates. The output response to the policy shock remains the same as under the baseline. Because our shock measure is identified outside the VAR it seems likely that our results are robust to other modifications to the VAR framework.

Finally, we estimate single-equation systems for output and prices similar to those estimated by Romer and Romer (2004). In keeping with the VAR results, we find a negative and persistent effect on output (with a point estimate of between 1% and 2%) and a small positive effect on the price level (although, due to wide estimated standard error bands, both effects are only at the border of statistical significance).[21]

This section closes with a final comparison exercise. To shed some light on how our factor-derived shock measure compares with the simple Kuttner (2001) spot-month shock, one can estimate the baseline model with the (cumulated) spot-month innovation in place of our shock measure. In this case, the impulse response for output is closer to that for the other identification schemes, with a small, albeit insignificant, positive output response to a 'contractionary' policy shock. These results support the view that shocks to the spot month futures contract often reflect new information about

[21] Results of all robustness checks are available from the authors on request.

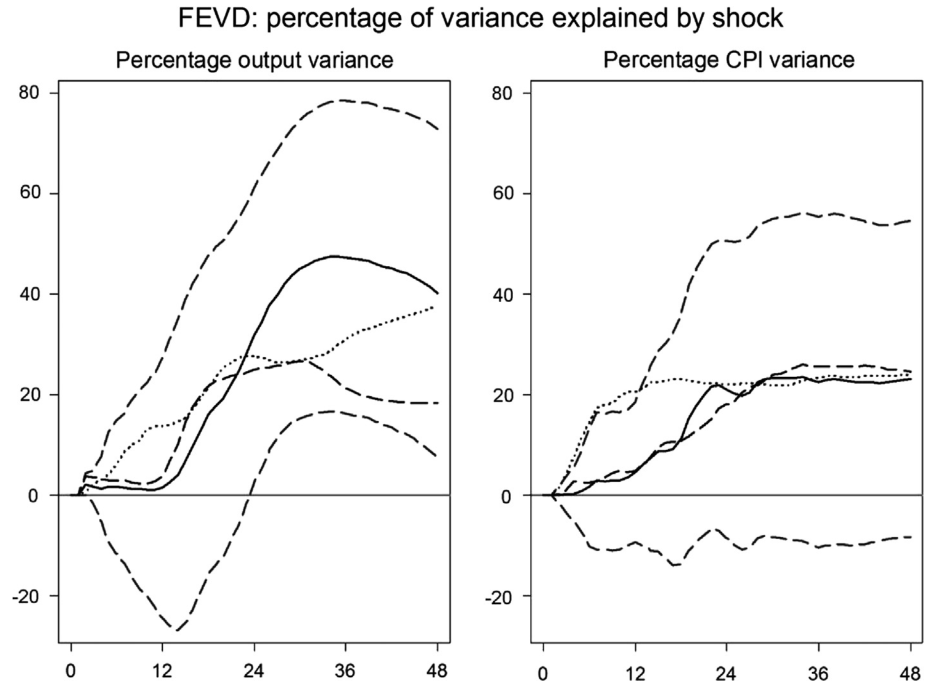

Fig. 7. Structural VAR (Monthly data, 3 endogenous variables plus constant and linear time trend, 36 lags). Variables ordered as industrial production, consumer prices, both seasonally adjusted and in logs, and one of three policy measures: our shock measure; Romer and Romer's measure (both cumulated); and the Federal Funds rate. Graphs show Cholesky FEVDs: the percentage of the forecast error for output and CPI accounted for by each policy measure. The FEVD for our shock measure is shown in bold, with two standard error bands produced by bootstrapping the combined VAR and factor model system. FEVDs for the Fed Funds rate (dashed line) and Romer and Romer shock (dotted line) are shown for comparison (SE bands not shown).

the timing, rather than the general direction, of policy. Hence, it is not surprising that the IRFs associated with this noisy shock measure are imprecisely measured. As with the other identification schemes discussed in Section 2, the apparently perverse sign of the estimated effect of policy on output is suggestive of simultaneity bias, perhaps because timing shocks are particularly associated with the Fed's communication of internal information.

4.3. Decomposing our shock measure

Section 3.4 presented evidence that our shock measure may be contaminated by the Fed's reaction to its own information on near term inflationary pressure. Romer and Romer's (2000) analysis of Fed and private sector forecasts suggests that the Fed's forecasts are likely to include some accurate exclusive information. To shed some additional light on this issue, we regress the Fed's exclusive information (the difference between the Fed's forecast and the private sector forecast) on the private sector's overall forecast error (the difference between the actual outcome and the private sector forecast), for both real GDP and the GDP deflator and at forecast horizons of 0–2 quarters. The R^2s from these regressions have the interpretation of the share of the Fed's internal information that turns out to be correct ex post. This share varies from 1% to 6% for real GDP and from 3% to 20% for the GDP deflator (results reported in the appendix, Table A6). This points to potential positive bias in our estimate of the effect of policy on output and inflation. Note though that the Fed's accurate information accounts for a relatively small share of the difference between its forecast and the private sector's, suggesting that the bias is small.

To provide some additional evidence on the likely impact of this bias on our results, we decompose our shock measure using the results of the regression of the shock on the Fed's internal information presented in Table 1. The residuals from this equation give an estimate of the ‘pure’ shock component, while the fitted values give an estimate of any remaining portion of the systematic component $f(\Omega_t)$. However, simultaneity bias is not the only likely source of bias in the results. Attenuation bias (bias towards zero) due to measurement error is also likely to be present. While the residual should be cleansed of simultaneity problems, if $f(\Omega_t)$ is correctly specified then the fitted value will be cleansed of measurement error (it will all be captured by the residual term). When the two decomposed shock measures are entered in the baseline VAR system (Fig. 8), both the ‘predicted’ portion of the shock (bottom panels) and the residual portion (top panels) have a significant negative effect—of strikingly similar magnitude—on output. This suggests that, for output, the likely bias resulting from the news about the Fed's own information set being included in our shock measure is of around the same order of magnitude as the bias due to measurement error, where this latter bias is likely to be small. Moreover, since both sources of bias should tend to drive the estimated effect towards zero the true effect is likely somewhat larger. Note that the fitted portion of the “shock” measure accounts for 17% of output variation at a 3 year horizon, while the residual portion

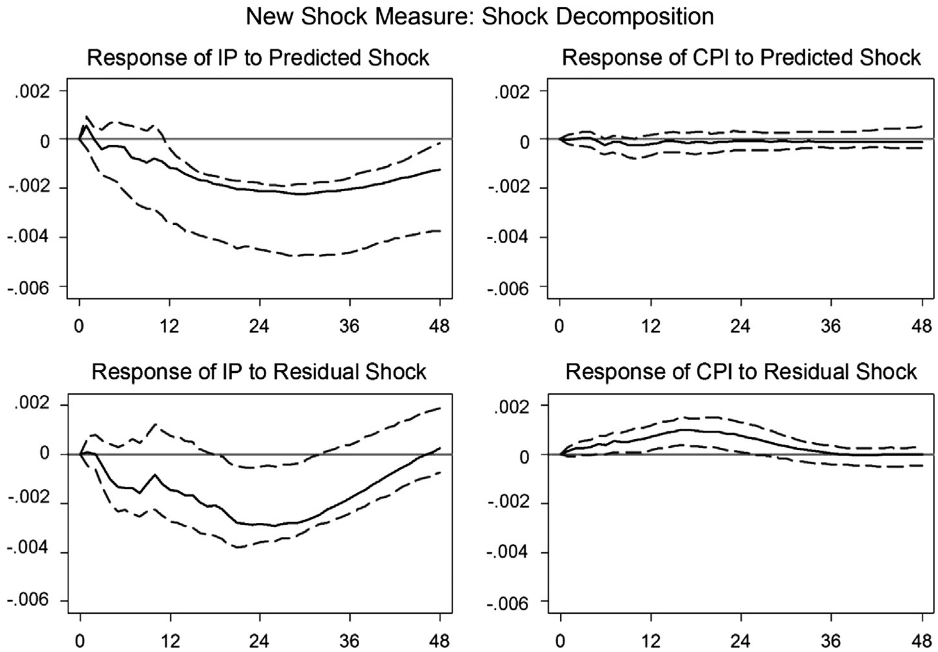

Fig. 8. Structural VAR (Monthly data, 4 endogenous variables plus constant and linear time trend, 12 lags). Data sample 1988:12–2002:12. Variables ordered as industrial production, consumer prices, both seasonally adjusted and in logs, and the predicted and residual components of the regression of our shock measure on the Fed's private information described in the text, both cumulated. Graphs show response of industrial production and CPI to a one standard deviation positive shock to each policy measure. Structural shocks obtained via Cholesky decomposition. Two Standard Error bands produced by bootstrapping the combined VAR and factor system (500 replications).

accounts for 30%, reflecting the fact that most of the variation in our shock measure cannot be accounted for by the Fed's internal information.

5. Conclusion

Conventional VAR and non-VAR identification schemes for estimating the effect of U.S. monetary policy shocks on the wider economy are sensitive to the sample period under consideration. In particular, these schemes generate unrealistic impulse response functions for output, and to a lesser extent prices, for the quarter century starting in the mid-1980s known as the “great moderation”. These apparently perverse results may be generated by a failure to properly identify the Fed's reaction function to allow for changes in its parameters over time, particularly a greater weight placed on forward-looking variables.

This paper outlines a new measure of monetary policy shocks derived from Fed Funds futures contracts that is less prone to these problems. As a result, our new measure generates a more realistic impulse response function for output, with a small but statistically significant negative effect whose maximum impact is felt at a horizon of two years following a monetary contraction. There is also evidence of a “price puzzle” over the medium term. Almost half of output variability (at a 3 year horizon) can be explained by monetary policy shocks using our new identification strategy, twice the share under other identification schemes for the same period.

While our shock measure may be contaminated by the Fed's systematic policy reaction to its internal forecasts, this is likely to bias our estimated impulse responses towards zero, so that the estimated output response may represent an underestimate. Moreover, while this simultaneity bias appears to be small under our identification scheme, it is likely to be more important for VAR-based identification methods.

One can rationalize the high share of output volatility accounted for by our shock measure by a combination of substantive and econometric factors. Substantively, the Fed exercised more effective control over the economy during the ‘great moderation’ period covered in our analysis, partly via an improved focus on forward-looking indicators, helping to minimize the impact of exogenous demand shocks so that a greater share of the remaining shocks is accounted for by policy itself. Although the absolute effect of the shocks is small, their relative impact is large in a period of relatively low overall volatility. In addition, our shock measure captures only policy changes that were truly unanticipated by the private sector, and it is these unexpected monetary policy changes that are generally believed to have the largest impact on output. Additional econometric factors include the fact that our shock variable is not a pure measure of shocks but also includes the Fed's systematic response to its internal forecasts. While the inclusion of the Fed's response to exclusive information will tend to reduce the magnitude of the estimated coefficients, it may increase the overall effect by increasing the size of the estimated shocks, although we find that this systematic component accounts for only one-third of our shock measure's total estimated contribution to output variation.

Acknowledgments

Substantive work was completed while the corresponding author worked at the International Monetary Fund (IMF). This work reflects the views of the authors alone and does not reflect the views of the IMF, its Executive Board or Management. Similarly, it does not necessarily reflect the views of Capula Investment Management LLP, and should be attributed accordingly. The authors would like to thank, subject to the usual caveats, Olivier Blanchard and Larry Christiano for useful discussions and the editor, associate editor and an anonymous referee, seminar participants at the IMF, Federal Reserve Bank of Chicago and Bank of England, Julian Di Giovanni, Ken Kuttner, Hashem Pesaran, Paolo Surico, Eric Swanson and—in particular—David Romer for useful comments on an earlier draft.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.jmoneco.2013.09.006.

References

Bagliano, F.C., Favero, C.A., 1998. Measuring monetary policy with VAR models an evaluation. European Economic Review 42, 1113–1140.
Bernanke, B.S., Kuttner, K.N., 2005. What explains the stock market's reaction to federal reserve policy? Journal of Finance 60, 1221–1257.
Bernanke, B.S., Mihov, I., 1998. Measuring monetary policy. Quarterly Journal of Economics 113, 869–902.
Boivin, J., Giannoni, M.P., 2006. Has monetary policy become more effective? Review of Economics and Statistics 88, 445–462.
Castelnuovo, E., Surico, P., 2006. The Price Puzzle: Fact or Artefact? Bank of England Working Paper No. 288, Bank of England, London.
Christiano, L.J., Eichenbaum, M., Evans, C.L., 1996. The effects of monetary policy shocks: evidence from the flow of funds. Review of Economics and Statistics 78, 16–34.
Christiano, L.J., Eichenbaum, M., Evans, C.L., 1999. Monetary policy shocks: what have we learned and to what end?. In: Taylor, J.B., Woodford, M. (Eds.), Handbook of Macroeconomics, vol. 1. Elsevier B.V., Amsterdam.
D'Amico, S., Farka, M., 2011. The fed and the stock market: an identification based on intraday futures data. Journal of Business and Economic Statistics 29, 126–137.
Faust, J., Swanson, E., Wright, J., 2004. Identifying VARs based on high-frequency futures data. Journal of Monetary Economics 6, 1107–1131.
Giordani, P., 2004. An alternative explanation of the price puzzle. Journal of Monetary Economics 51, 1271–1296.
Gürkaynak, R.S., 2005. Using Federal Funds Futures Contracts for Monetary Policy Analysis. Federal Reserve Board Finance and Economics Discussion Series #2005-29, Federal Reserve Board, Washington, DC.
Gürkaynak, R.S., Sack, B., Swanson, E., 2005. Do actions speak louder than words? The response of asset prices to monetary policy actions and statements. International Journal of Central Banking 1, 55–94.
Hamilton, J.D., 2008. Daily monetary policy shocks and new home sales. Journal of Monetary Economics 55, 1171–1190.
Hanson, M.S., 2004. The “price puzzle” reconsidered. Journal of Monetary Economics 51, 1385–1413.
Kuttner, K.N., 2001. Monetary policy surprises and interest rates: evidence from the federal funds futures market. Journal of Monetary Economics 47, 523–544.
Orphanides, A., 2003. Historical monetary policy analysis and the Taylor rule. Journal of Monetary Economics 50, 983–1022.
Owyang, M.T., Wall, H.J., 2009. Regional VARs and the channels of monetary policy. Applied Economics Letters 16, 1191–1194.
Piazzesi, M., 2010. Affine term structure models. In: Hansen, L.P., Ait-Sahalia, Y. (Eds.), Handbook of Financial Econometrics. North-Holland, Amsterdam.
Piazzesi, M., Swanson, E.T., 2008. Futures prices as risk-adjusted forecasts of monetary policy. Journal of Monetary Economics 55, 677–691.
Romer, C.D., Romer, D.H., 2000. Federal reserve information and the behavior of interest rates. American Economic Review 90, 429–457.
Romer, C.D., Romer, D.H., 2004. A new measure of monetary shocks: derivation and implications. American Economic Review 94, 1055–1084.
Rudebusch, G.D., 1998. Do measures of monetary policy in a VAR make sense? International Economic Review 39, 907–931.
Sims, C.A., Zha, T., 2006. Does monetary policy generate recessions? Macroeconomic Dynamics 10, 231–272.
Söderström, U., 2001. Predicting monetary policy with Federal funds futures prices. Journal of Futures Markets 214, 377–391.
Taylor, N., 2010. The determinants of future U.S. monetary policy: high-frequency evidence. Journal of Money, Credit and Banking 42, 399–420.
Thapar, A., 2008. Using private forecasts to estimate the effects of monetary policy. Journal of Monetary Economics 55, 806–824.