NBER WORKING PAPER SERIES

ARE THE EFFECTS OF MONETARY POLICY SHOCKS BIG OR SMALL?

Olivier Coibion

Are the Effects of Monetary Policy Shocks Big or Small?
Olivier Coibion
NBER Working Paper No. 17034
May 2011
JEL No. E3,E4,E5

<u>ABSTRACT</u>

This paper studies the small estimated effects of monetary policy shocks from standard VAR's versus the large effects from the Romer and Romer (2004) approach. The differences are driven by three factors: the different contractionary impetus, the period of reserves targeting and lag length selection. Accounting for these factors, the real effects of policy shocks are consistent across approaches and most likely medium. Alternative monetary policy shock measures from estimated Taylor rules also yield medium-sized real effects and indicate that the historical contribution of monetary policy shocks to real fluctuations has been significant, particularly during the 1970s and early 1980s.

Olivier Coibion
Department of Economics
College of William and Mary
115 Morton Hall
Williamsburg, VA 23188
and NBER
ocoibion@wm.edu

A central question of monetary economics is the sensitivity of the economy to policy instruments. Quantifying this sensitivity, however, requires disentangling endogenous and exogenous changes in the policy instrument. Much of the evidence on the quantitative effects of interest rate changes has come from the VAR literature, most commonly relying on the identifying assumption that policy innovations have no contemporaneous effects on macroeconomic variables. The predominant finding has been that the effects of unexpected policy innovations are relatively small.[1] Romer and Romer (2004, R&R henceforth) reach a very different conclusion using a novel approach to identify monetary policy innovations by first constructing a historical series of interest rate changes decided upon at meetings of the FOMC and then isolating the innovations to these policy changes that are orthogonal to the Federal Reserve's information set. R&R show that this new measure of policy shocks points to much larger effects of monetary policy shocks than previously thought.

In fact, their results imply not just greater effects of monetary policy shocks, but also a dramatically different historical interpretation of U.S. business cycle fluctuations than what one would reach based on the VAR approach. To see this, Figure 1 plots the implied contribution of monetary policy shocks to macroeconomic fluctuations between 1970 and 1996 from a standard VAR and the baseline R&R approach.[2] The results from the standard VAR approach are well known: monetary policy shocks appear to account for very little of the fluctuations in the real economy, measured either via industrial production or unemployment, and play only a small role in accounting for inflation patterns over this time period. Strikingly, even the 1980-1982 recessions, commonly attributed to the policy changes enacted by then newly appointed Chairman Volcker, cannot be explained by monetary policy shocks according to the standard monetary VAR. The R&R methodology, on the other hand, implies that the brunt of fluctuations in industrial production, unemployment, and inflation have been driven by monetary policy innovations. Even the 1990 recession, for example, is attributed to contractionary monetary policy according to the baseline R&R methodology, a result at odds with the conventional wisdom.[3] A hybrid approach, which incorporates the R&R shocks into a standard VAR yields an interpretation in which monetary policy innovations can account for much of the trend component of inflation and a smaller, albeit non-negligible, component of fluctuations in the real side of the economy.

Given the stark contrast implied by the different approaches not only for measuring the direct effects of monetary policy shocks via impulse responses but also the historical interpretation of U.S. business cycle fluctuations, a more detailed analysis of these methods is needed to identify why the results differ so dramatically. I find that three key elements play a significant role in accounting for the

---

[1] See for example Leeper, Sims and Zha (1996), Bernanke, Gertler and Watson (1997), Bernanke and Mihov (1998), Christiano et al. (1999), and Bernanke, Boivin and Eliasz (2005).
[2] The methods used for the results of Figure 1 are described in detail in section 1.
[3] See Blanchard (1993), Hall (1993) and Walsh (1993).

1

difference in the estimated effects of monetary policy shocks across the different methods. First, by applying a common approach, namely the single equation approach of R&R, to estimating impulse responses to the different shocks, I show that the R&R innovations identify episodes in which the change in monetary policy is larger: the Federal Funds Rate (FFR) rises much more and for a longer period after the baseline R&R shocks than after either standard VAR shocks or the implied shocks from the hybrid VAR. Controlling for this difference in how contractionary monetary policy is in response to a shock, both sets of VAR shocks yield very similar effects of monetary policy innovations as the baseline R&R results.

Second, I show that the period in which the Federal Reserve abandoned targeting the Federal Funds Rate between 1979 and 1982 plays an important role in accounting for the different results across methods. In particular, the large effects of monetary policy shocks identified by R&R are disproportionately driven by two historical episodes: the rapid decline in the FFR in April-June 1980 and the subsequent rapid increase in September-November of 1980. Dropping either of these episodes lowers the estimated effects of monetary policy shocks using the baseline R&R method significantly, by about a half in the latter case and by about a third in the former case. Both episodes were in the midst of the non-borrowed reserves targeting period by the Fed, during which the Federal Funds Rate experienced extreme swings. Thus, both periods are likely to be particularly noisy for measuring policy innovations. Consistent with this, I document that the R&R shocks are predictable using lagged macroeconomic variables, but that this predictability is eliminated once one omits the period of non-borrowed reserves targeting. Additional evidence comes from alternative measures of the stance of monetary policy during these episodes, such as the Bernanke and Mihov (1998) measure or the Boschen-Mills (1995) index, which point to much smaller policy reversals than implied by the R&R measure. If one omits the period of non-borrowed reserves targeting from the estimation, all three approaches (standard VAR, hybrid VAR, and R&R method) yield very similar estimates of the real effects of monetary policy shocks: approximately mid-way between the standard VAR and the benchmark R&R results.

Third, I find that R&R's approach to estimating impulse responses is particularly sensitive to the lag structure. While they assumed fairly long lag specifications (2 years of autoregressive lags for each macroeconomic variable and 3-4 years of lags of the shock) in estimating impulse responses, I show that the estimated peak effects of monetary policy shocks using their approach are nearly monotonically increasing in the assumed lag length. In addition, AIC lag selections consistently point to much shorter lag specifications than assumed by R&R and yield estimated peak effects over the whole sample that are between the results of R&R and those of the baseline VAR. Similar results obtain from a model-averaging procedure designed to take into account uncertainty about the true lag length. Accounting for the period of non-borrowed reserves targeting eliminates much of this sensitivity to lag length and yields

2

estimates of the peak real effects of monetary policy shocks which are more consistent across the three approaches: a one-hundred basis point innovation to the Federal Funds Rate lowers industrial production by 2-3% and raises the unemployment rate by approximately half a percentage point.

To provide additional evidence on the size of the effects of monetary policy shocks, I consider three alternative ways of identifying monetary policy shocks. First, I re-estimate the same reaction function as R&R but follow the suggestion of Hamilton (2008) and Sims and Zha (2006b) of explicitly allowing for heteroskedasticity in the innovations (via GARCH). Using this very close substitute to the original R&R shocks, I find that most of the sensitivity issues that were associated with the original R&R shocks are mitigated. Second, following Boivin (2006) and Coibion and Gorodnichenko (2011), I extract a measure of monetary policy shocks from an estimated Taylor rule with time-varying parameters. The resulting shock series, which should identify a narrower set of monetary policy shocks, namely innovations to the policy rule rather than changes in the coefficients of the rule, yields real effects of monetary policy shocks which are again between the VAR and original R&R estimates. Finally, I use the monetary policy shocks estimated using a DSGE model by Smets and Wouters (2007) and show that these shocks also point to medium-sized effects on real variables. In short, the analysis points to a clear answer to the question posed in the title of the paper: the effects of monetary policy shocks are medium.

Despite the fact that the estimated effects of monetary policy shocks appear to be smaller than originally found by R&R using their own measure of shocks as well as with the GARCH and time-varying coefficient based shocks, the implied historical contribution of monetary policy innovations to production remains nontrivial and differs qualitatively from that implied by standard VAR's. This is particularly the case during the 1970s in which "stop-go" monetary policy likely played an important role in affecting fluctuations in production, as argued in Romer and Romer (2002), and during the Volcker disinflation. This supports the broader claim of R&R that monetary policy innovations are historically significant and contrasts with the prediction of standard VAR's that monetary policy shocks have mattered little for historical U.S. macroeconomic fluctuations.

Section 1 provides details on the standard VAR approach, the baseline R&R framework, and their hybrid VAR, as well as baseline estimates of the effects of monetary policy shocks from each. Section 2 focuses on the effects of different approaches to estimating impulse responses as well as sensitivity to specific time periods and lag selection issues for these measures of monetary policy shocks. Section 3 considers alternative measures of monetary policy shocks. Section 4 concludes.

I   Estimating the Effects of Monetary Policy Shocks

The most traditional approach to identify and estimate the effects of monetary policy shocks remains the timing assumption that a current innovation to the instrument used by monetary policymakers, most

commonly assumed to be the Federal Funds Rate in the U.S., has no contemporaneous effect on macroeconomic variables such as production, employment and prices. This assumption has motivated a long literature that estimates the effects of monetary policy shocks using a vector autoregression (VAR) ordered with the policy variable last. Following Christiano, Eichenbaum and Evans (1999), I will use the following ordering of variables as representative of this literature: the log of industrial production (seasonally adjusted), the unemployment rate, the log of the Consumer Price Index (CPI, seasonally adjusted), a log commodity price index, and the effective Federal Funds Rate (FFR).[4] While including the unemployment rate is not typical, I do so to have an additional measure of the real effects of monetary shocks. For this purpose, the unemployment rate is particularly useful because so much of its fluctuations are at the business cycle frequency.[5] Estimates of the effects of a one hundred basis point innovation to the FFR, based on monthly data from 1970:1 to 1996:12, are presented in Figure 2, allowing for 12 lags in the VAR.[6] Confidence intervals are generated by repeatedly (1,000 times) drawing from the asymptotic distribution of the parameters which yields a distribution of impulse responses. As commonly noted, the estimated real effects of a monetary policy shock are relatively small: the peak effect on industrial production is less than one percent (-.7%) and the unemployment rate rises by 0.16 percentage points.

These estimates are in line with much of the empirical findings of the VAR literature. For example, Christiano et al. (1999) find a peak drop of approximately 0.70% in production in response to a 100 basis point increase in the Federal Funds Rate.[7] Similarly, Leeper, Sims and Zha (1996) estimate a peak drop in GDP of 0.35% and an increase in the unemployment rate of approximately 0.10 percentage points in their thirteen variable model. Bernanke and Blinder (1992) find somewhat higher estimates, with unemployment rising by approximately 0.60 percentage points. Bernanke, Gertler and Watson (1997) find a peak effect on interpolated monthly GDP of approximately -0.4% while Bernanke and Mihov (1998) estimate a peak effect of approximately -0.5% for monthly interpolated GDP. Bernanke, Boivin and Eliasz (2005) estimate a Factor-Augmented VAR and find that a 100 basis point increase in the FFR lowers industrial production by a maximum of approximately 0.6% and raises the unemployment rate by 0.20% points. Gorodnichenko (2006) proposes an alternative factor-based VAR analysis which predicts a peak drop in real GDP of approximately 0.8%. Uhlig (2005) uses sign restrictions to estimate

---

[4] The commodity price index is from the Commodity Research Bureau. All variables in the VAR are in levels. The CPI housing costs until 1983 were measured in large part using monthly payments on new mortgages, so contemporaneous interest rates had a direct effect on the overall CPI. All results in the paper are robust to using the CPI excluding shelter instead of the overall CPI.
[5] None of the results are sensitive to the inclusion of the unemployment rate in the VAR.
[6] The time sample is set to be comparable to the R&R sample of monetary policy shocks.
[7] These and subsequent figures concerning the results of previous work come from impulse responses displayed in each paper and are thus only meant to be approximations. In each case, I compute implied responses to a 100 basis point increase in the FFR based on the reported point estimates.

the effects of monetary policy shocks and finds a peak drop in monthly interpolated GDP of approximately 0.30% when he imposes the assumption of zero response of real GDP on impact. Faust, Swanson and Wright (2005) estimate the effects of monetary policy shocks using Futures markets for the FFR and find a peak drop in GDP of 0.60%.[8] Thus, despite different identification strategies, variables, time samples, lag structures and other factors, the consensus finding from this literature is that the effects of monetary policy shocks are relatively small.[9]

A notable exception to this finding of small effects of monetary policy shocks is Romer and Romer (2004). They propose a novel procedure to identify monetary policy shocks. First, they use the narrative approach to extract measures of the change in the Fed's target interest rate at each meeting of the FOMC between 1969 and 1996. They then regress this measure of policy changes on the Fed's real-time forecasts of past, current and future inflation, output growth and unemployment. The residuals from this regression constitute their measure of monetary policy shocks. This procedure has several advantages over the identification scheme of the VAR literature. First, their measure of policy eliminates much of the endogenous movement between the effective FFR and macroeconomic variables.[10] Second, by conditioning on the Fed's real-time forecasts of macroeconomic variables which are formed using a multitude of economic variables, their policy innovations identify changes in policy that are independent of current expectations of future economic conditions. Because the standard VAR only controls for a small set of macroeconomic variables, the R&R measure of monetary policy shocks is therefore likely to be relatively freer of anticipatory movements than measures from a standard VAR.

Having derived a measure of monetary policy shocks, R&R then assess the effects of these shocks via regressions of each macroeconomic variable on lags of itself and lags of the policy shock:

$$\Delta x_t = c + \sum_{i=1}^{I} \beta_i \Delta x_{t-i} + \sum_{j=1}^{J} \gamma_j \varepsilon_{t-j}^{mp} + v_t \tag{1}$$

where x is the macroeconomic variable and ε^{mp} is their measure of monetary policy shocks. Note that this specification applies the same timing restriction as the standard VAR, namely that policy shocks have no contemporaneous effect on macroeconomic variables. With estimates of the parameters of (1), impulse responses can readily be constructed. Given monthly data, they set I=24 and J=36 for industrial production and I=24 and J=48 for prices.[11] R&R use the PPI for finished goods, but I will focus on the CPI for consistency with most of the literature. Figure 2 presents the impulse responses of industrial

---

[8] The method of Faust et al. (2005) does not yield point estimates but rather confidence intervals. The estimate here is based on the midpoint over their confidence intervals.

[9] These papers also consistently find that monetary policy shocks account for a relatively small component of forecast error decompositions of real macroeconomic variables, with the exception of Christiano et al. (1999) who find that monetary policy shocks can account for up to 44% of the output forecast error variance.

[10] One potential limitation of relying on changes in the target rate rather than the effective rate is that policy changes may be enacted prior to the meetings of the FOMC.

[11] R&R estimate (1) using unseasonally-adjusted data with monthly dummies. Because this has no important effects on the results, I focus on the specification with seasonally-adjusted data.

production, unemployment and prices, to a permanent 100 basis point shock using this approach. The time period for the analysis is 1970:1-1996:12, with shocks prior to 1969:3 set equal to zero. For the unemployment rate, which R&R did not consider, I estimate (1) in levels with the same lag structure as industrial production. Confidence intervals are generated by repeatedly (1,000 times) drawing from the asymptotic distribution of the parameters which yields a distribution of impulse responses. The contrast between this approach and the standard VAR is striking: the R&R methodology yields a peak effect on production of -4.3% and a rise in the unemployment rate of nearly a full percentage point (0.93). The effect on production and unemployment is thus more than four times as large as that of the standard VAR approach. R&R show that these results are robust to longer lag lengths in (1) and alternative specifications of the regression used to derive the shock series.

In addition to their single equation specification, R&R also consider a hybrid approach in which they integrate their new measure of monetary policy shocks into a standard VAR a la Christiano et al. (1999). Specifically, in place of the effective FFR, they substitute the cumulative sum of their monetary policy shocks. The advantage of this procedure is that one should be able to more precisely identify the effects of monetary policy shocks than in the standard VAR, given that the R&R measure controls for much of the endogenous fluctuations in the interest rate as well as the Fed's information set, while also simultaneously estimating the joint dynamics of macroeconomic variables, including controlling for commodity prices.[12] Figure 2 presents estimates from this hybrid procedure, using the baseline VAR with the cumulative R&R shocks in lieu of the FFR, again using a year's worth of lags. The estimated real effects of monetary policy shocks are in between the standard VAR and the baseline R&R results, with industrial production falling a maximum of 1.6% and unemployment rising by 0.40 percentage points. The response of the price level is also between that of the standard VAR (and does not exhibit a pronounced price puzzle) and the baseline result of R&R.

These approaches thus point to rather startling differences in the estimated effects of monetary policy shocks and, as illustrated in Figure 1, even more striking differences in historical interpretations. The standard VAR approach finds small real effects of monetary policy shocks and concludes that these innovations account for very little of historical business cycle fluctuations. The R&R approach, on the other hand, identifies large effects of monetary policy shocks and indicates that these shocks can account for much of the historical fluctuations at business cycle frequencies in production, employment and inflation. The hybrid approach yields a more moderate view, in which policy shocks have larger effects

---

[12] R&R also reproduce estimates of (1) controlling for commodity prices. This yields slightly smaller effects of monetary policy shocks on production. However, because commodity prices are endogenous variables which should respond to monetary policy (see Barsky and Kilian 2002 and Kilian 2009), these impulse responses omit the endogenous response of commodity prices to the shock and the subsequent effects of these commodity price changes on domestic production. The hybrid VAR, on the other hand, controls for this interaction.

than in the standard VAR framework, and can account for the gradual rise of inflation in the 1970s, the disinflation of the 1980s, but much less of the business cycle fluctuations in production and employment than implied by the baseline R&R approach.

II What Accounts for the Different Estimated Effects of Monetary Policy Shocks?

To disentangle the differences among the approaches, I first compare the implied monetary policy shocks of each approach. Second, I apply the same single equation procedure to generate impulse responses, that of R&R, to each set of shocks. Third, I study the influence of the period of non-borrowed reserves targeting. Fourth, I present sensitivity results to lag selection issues as well as a model-averaging exercise to address this sensitivity.

2.1 Identified Monetary Policy Shocks

The first step to understanding the difference across specifications is to consider the implied monetary policy shocks of each. The top panels of Figure 3 plot the cumulative sum of monetary policy shocks from each approach. As can immediately be seen, there is substantial comovement between these series. The correlation between actual shocks (cumulative shocks) from the standard VAR and the R&R approach is 0.31 (0.53) while that between the actual shocks (cumulative shocks) from the hybrid VAR and the R&R approach is 0.83 (0.28). In the mid to late 1970s, all three measures are falling, which points to a systematic period of loose monetary policy, i.e. interest rates being set lower than one would expect given macroeconomic conditions. This is reversed in the early 1980's, with a sequence of positive monetary policy innovations associated with the Volcker disinflation. Despite these similarities, there are also some visible differences. First, identified monetary policy shocks from the standard VAR are more volatile than the other measures: the standard deviation of shocks from the standard VAR is almost 50% greater than the R&R shocks and 75% greater than the hybrid shocks. Second, the cumulative R&R series exhibit a much more pronounced decline in the 1970s than the other series. Third, while the R&R shocks point to increasingly contractionary monetary policy (conditional on the Fed's forecasts) in the mid-1980s to early 1990s, both the standard VAR and the hybrid VAR exhibit mostly negative sequences of policy shocks over the same time period. The bottom panels of Figure 3 plot the difference between the identified shocks from the standard VAR and the R&R baseline shocks, as well as the difference between the identified shocks from the hybrid VAR and the R&R baseline shocks. Not surprisingly, the identified shocks are most different between 1979 and 1982, the period in which the Federal Reserve abandoned targeting the Federal Funds Rate. I return to this feature of the data in section 2.3.

7

2.2   Impulse Responses using Single Equation Estimation

One way to address how much of the difference in impulse responses comes just from the shock series is to apply the single equation approach to constructing impulse responses of R&R to each shock series. Figure 4 plots the impulse responses extracted from estimates of equation (1) for industrial production, unemployment (in levels) and prices applied to the different measures of monetary policy shocks. In addition, I estimate the impulse response of the effective Federal Funds Rate to each shock measure, again using equation (1) in levels with the same lags as for industrial production, to assess whether the different shocks lead to quantitatively similar monetary contractions.[13] Note first that the results for the baseline R&R specification are identical to those in Figure 2 since we are applying exactly the same procedure as before. Using the shocks from the standard VAR, the results are larger than those in Figure 2: a 100 basis point innovation leads to a peak drop in industrial production of approximately 2% (instead of 0.7%) and a peak increase in the unemployment rate of almost 0.4 percentage points (instead of 0.16), although the standard errors are also now much larger. The response of the price level is again more pronounced, with prices falling 2% after five years. For the hybrid VAR, the results are quantitatively very similar to the original impulse responses. Thus, while using the same procedure as R&R to estimate impulse responses does raise the implied quantitative effects of monetary policy shocks using the standard VAR shocks, there remains a sizable difference between the results of R&R and the two other specifications.

One factor underlying this difference is apparent from the average response of the Federal Funds Rate to each type of innovation. After a standard VAR shock, the FFR rises by approximately 100 basis points for a few months and then begins to decline. The response of the FFR to an innovation from the R&R shock series, on the other hand, is much larger. First, the FFR rises by over 200 basis points, and remains high for an extended period of time relative to the effects of a VAR shock. The hybrid VAR shocks are in between: the FFR rises by 200 basis points in the first few months after the shock, but then rapidly return to zero. Thus, a key difference between the R&R shocks and the other shock measures is that *the R&R shocks are associated with much larger and longer changes in monetary policy: i.e. the contractionary impetus is not the same across shocks*. As a result, it should not be surprising to observe larger effects of monetary policy innovations from the R&R shocks on macroeconomic variables.

To quantify this notion, I construct a counterfactual set of impulse responses for the two VAR shocks in which the initial innovations are designed to match the quantitative response of the FFR to the baseline R&R innovations. Specifically, for each set of VAR shocks, I select a sequence of innovations used to generate the impulse response of the FFR (innovations are allowed to be non-zero over the first 24 months) which minimizes the distance between the impulse response of the FFR to these innovations and

---

[13] In addition, I allow the shocks to have a contemporaneous effect on the interest rate.

8

the impulse response of the FFR to R&R shocks over the first 36 months. One can then construct alternative impulse responses for industrial production, unemployment and prices by feeding this sequence of innovations into the single equation estimates of (1) for these variables, in the spirit of Sims and Zha (2006a). The results, illustrated in Figure 4, indicate that once one controls for the different response of the FFR to each shock measure, the impulse responses are broadly similar across specifications.[14] Both the baseline VAR shock approach and the hybrid VAR point to peak effects of monetary policy shocks on industrial production of approximately -4%, unemployment increases of almost 1% point, and drops in the price level of around 4-6% after five years. In other words, one reason why the baseline R&R specification is yielding bigger implied effects of monetary policy innovations is because the contractionary impetus associated with each type of shock is not directly comparable across shocks.

2.3 The Period of Non-Borrowed Reserves Targeting

Another possibility for why alternative approaches yield such different estimated effects of monetary policy shocks is that the results could be driven by outliers from each specification. Figure 3 showed, for example, that while there is a high correlation between the different shock series, there are some substantial differences as well, particularly between 1979 and 1982 when the biggest monetary policy shocks in the sample occurred. This period is problematic for the identification of monetary policy shocks. On the one hand, the Volcker disinflation era is commonly cited as the clearest example of a monetary-policy driven recession and so should be particularly instructive for identifying the effects of monetary policy shocks. On the other hand, because the Federal Reserve officially stopped targeting the FFR during this time period, it is also the period in which the identification procedure from each methodology is most likely to be misspecified.

To assess the sensitivity of the empirical results to outliers and specific episodes, I replicate the estimation procedure of section 2.2 (i.e. applying single equation estimation for alternative shock measures) but set rolling 3-month intervals of shocks equal to zero. For each specification, I then extract the implied peak effect of monetary policy shocks on industrial production, unemployment, and prices, as well as the cumulative effect on the FFR over the first 12 months of the impulse response. The latter provides a simple metric for the size of the contractionary impetus associated with each shock measure which incorporates both the magnitude of the FFR's response to policy shocks and its persistence.

---

[14] The counterfactual responses display larger responses but otherwise maintain very similar stochastic properties as the original responses. For example, the standard deviation of interest rate changes during the first 36 months of the impulse responses goes from 0.21 (0.32) for the standard VAR (hybrid VAR) to 0.27 (0.30) in the counterfactual responses, bringing each more in line with the standard deviation of 0.28 from using the R&R shocks.

The results are presented in Figure 5. The first result to note is that the peak effects of monetary policy shocks using the hybrid shocks are quite robust. For example, the peak effects on unemployment are consistently between 0.4 and 0.6 percentage points, with little sensitivity to individual shocks. Similar results hold for the response of prices and industrial production. On the other hand, both the baseline R&R and the specifications based on standard VAR shocks are quite sensitive to certain observations. The results using the standard VAR shocks vary when certain shocks over the entire sample are dropped, but are most noticeably sensitive to observations between 1979 and 1982.[15] The R&R results are also sensitive, but almost exclusively to observations between 1979 and 1982. Dropping some of these observations has significant implications on the peak estimated effects of monetary policy shocks, sometimes reducing them to the values one gets using the hybrid shocks. The results in Figure 5 also suggest that the larger and more persistent response of the FFR to R&R shocks relative to the other shock measures identified in the previous section is also a reflection of a limited number of observations between 1979 and 1982. Omitting some of these episodes reduces the size of the contractionary impetus of R&R shocks as well as the associated real effects of monetary policy shocks.

In particular, there appear to be two episodes that play a leading role in driving the empirical results of R&R. These include the period from April 1980 to June of 1980 and that from September 1980 to November 1980. If we drop the latter 3-month episode, the peak effect of monetary policy shocks on industrial production (unemployment) falls to -1.9% (0.49% points), while dropping the former time period reduces the peak effect on industrial production (unemployment) to -2.9% (0.70% points). These episodes correspond to the period in which the FFR fell precipitously in the first half of the year then rose dramatically in the second half of the year. Interpreting changes in the FFR during this time period is complicated by the fact that the Fed was targeting alternative measures of monetary policy and that transcripts from this time period are often uninformative about the Fed's expectations for the FFR.[16]

---

[15] This sensitivity of VAR shocks only arises when impulse responses are estimated using the methodology of R&R. If instead one extracts the impulse responses directly from the VAR while dropping the relevant time periods, there is little sensitivity of the estimated peak effects to individual periods.

[16] Further complications come from other policy measures in this time period. For example, between March and July of 1980, the Federal Reserve imposed credit constraints (Credit Restraint Program) per the request of the Carter administration. While the details of the program appear to have been designed to minimize its effects, some evidence indicates that it had much more substantial effects than expected. According to the Board, "It is difficult, if not impossible, to say how much of the weakness in bank loans [under the program] … was due to the recession, how much to the reaction to fiscal announcements and general credit conditions (including expectational effects), how much to the cumulative effects of earlier overall restraints, and how much to the credit restraint programs. But the timing and abruptness of the change in loan growth trends suggest that the announcement of the programs played a significant role." The CRP was removed in July of 1980 as it became clear that the economy was in a recession (see Schreft 1990 for more details). Note that because this program was in place exactly during the time period when the FFR fell precipitously, this is one indication that monetary policy may not have been as loose as might be gauged from changes in the FFR over this time period.

Given the volatile nature of the FFR over this time period, it is non-trivial to disentangle market-driven fluctuations in the FFR from changes in monetary policy.

One way to discern whether these identified shocks are reasonable is to compare them to alternative measures of the stance of monetary policy over this time period. For example, while the baseline R&R shocks point to a cumulative policy-driven decrease in the FFR of nearly 4% between April and June of 1980 followed by an increase of nearly 4% between September and November, the corresponding identified policy changes from the standard VAR approach are of a slightly less than 3% cumulative decrease in the first episode and an increase of only 128 basis points in the second. This is particularly striking because the VAR uses the effective FFR as the measure of policy while R&R use intended policy changes, with the latter being substantially less volatile than the former, particularly during this time period. Similarly, the hybrid VAR points to a cumulative policy-driven decrease in interest rates of slightly under 3% during the first episode followed by a cumulative increase of 176 basis points between September and November. Because the hybrid VAR shocks are based on the cumulative R&R shock measure, but purified of the predictive component of lagged macroeconomic variables, one interpretation could be that the R&R shocks are overstating the degree of policy changes by not sufficiently controlling for the endogenous market-driven fluctuations in the FFR over this time period.

To assess this possibility, I regress the R&R shocks on current and lagged macroeconomic variables

$$\varepsilon_t^{rr} = c + \sum_{i=0}^{I} \beta_i x_{t-i} + v_t. \tag{2}$$

Under the null that the Romer and Romer shocks (ε_t^{rr}) are exogenous, they should be unpredictable using lagged macroeconomic variables. Equation (2) includes contemporaneous values of macroeconomic variables to be consistent with the identification restrictions imposed in both the VAR and single equation approaches. Table 1 reports the F-statistics for the null that all of the β's are equal to zero using the monthly change in industrial production, the unemployment rate, and the monthly inflation rate as regressors. Over the whole sample, we can reject the null of exogeneity of R&R shocks using industrial production and the unemployment rate. However, once we remove the October 1979 to September 1982 period, all of the F-statistics decline substantially and we can no longer reject the null hypothesis that all of the β's are zero for each of the macroeconomic variables. These results support the notion that the R&R shocks contain some element of endogenous responses to macroeconomic conditions during the period in which the Federal Reserve abandoned targeting the Federal Funds Rate.

In addition, one can compare the implied change in the stance of monetary policy from the R&R shocks to alternative measures from the literature. One such measure, similar in spirit to the R&R approach, is the Boschen-Mills (1995) index. This indicator, derived largely from FOMC directives, classifies policy according to the importance that policymakers assigned to reducing inflation relative to

promoting real GDP growth but does not attempt to identify whether changes in the stance of monetary policy reflect macroeconomic conditions or exogenous policy decisions. Another measure of the stance of monetary policy comes from Bernanke and Mihov (1998). They extract their measure from a VAR which allows for time-varying estimates of the central bank's operating procedures. Their empirical results indicate that FFR targeting is an appropriate description of the Federal Reserve's *with the exclusion of the 1979-1982 period*. As a result, their measure of the stance of monetary policy has the key advantage of being based on the most relevant instrument of monetary policy over time.[17] Figure 6 plots these two alternative measures of the stance of monetary policy along with the cumulative R&R shocks, normalized by the 36-month moving average of lagged values.[18] Because R&R shocks represent deviations of the intended FFR from the average response of the Fed to its real-time forecasts of macroeconomic variables, the normalized cumulative shock measure can also be interpreted as a measure of the stance of monetary policy (i.e. periods with a sequence of positive innovations to the FFR can be interpreted as periods in which the Fed raised interest rates more than it normally would have given its information set).

Overall, the three measures reveal broadly similar patterns of switches between loose and tight monetary policy, although some differences stand out. For example, both the Boschen-Mills and Bernanke-Mihov measures point to monetary policy tightening more rapidly in the late 1970s than does the R&R measure. More relevant to the issue at hand is the behavior of the series during 1980. The R&R measure exhibits a dramatic decrease in early 1980 followed by an almost identical reversal in the latter part of the year, both of which dwarf in magnitude all other policy changes over the time sample. The Boschen-Mills indicator, on the other hand, points only to a slight loosening of the stance of monetary policy in August of 1980, with policy remaining primarily concerned with inflation over GDP growth until November of 1981. The Bernanke-Mihov measure, like the R&R measure, identifies a loosening of monetary policy in early 1980 followed by a tightening in the latter half of the year, but the quantitative magnitude of these policy changes is much smaller (relative to policy changes in other periods) than the R&R measure. For example, the Bernanke-Mihov index implies that the monetary loosening in early 1980 was comparable in magnitude to the loosening in late 1974 to early 1975 and that of the end of 1982. In addition, the monetary contraction of the latter half of 1980 was, according to Bernanke and Mihov, quite small in historical perspective, in contrast to the R&R measure.

---

[17] On the other hand, their measures of monetary policy are likely somewhat contaminated by variations in the policy instrument arising from forward-looking behavior on the part of the Federal Reserve, as emphasized in R&R.
[18] The normalization by the 36-month moving average of lagged values is done to be comparable to the Bernanke-Mihov measure, which is based on such a normalization. However, the scales of the three measures are not directly comparable. Note also that both the Boschen-Mills and Bernanke-Mihov measures have the signs reversed from their original values, so that for all three measures, an increase is associated with a tightening of monetary policy.

These results illustrate the difficulty of properly interpreting the magnitude of policy changes during the era of non-borrowed reserves targeting. Combining this with the fact that the empirical results are particularly sensitive to these values implies that one should generally be wary of placing too much weight on results driven by this time period. As a result, I consider estimates from each approach that omit the early Volcker years by restricting the time period for the analysis to 1970:1-1979:9 and 1984:1-1996:12.[19] The results are presented in Figure 7. Dropping the period of non-borrowed reserves targeting eliminates most of the differences in estimated peak real effects across specifications. For the standard VAR approach, the peak effects of monetary policy shocks on real variables increase in absolute value, yielding a maximum drop in industrial production of almost 2% and a rise in the unemployment rate of almost 0.40 percentage points. The price puzzle is more pronounced, but prices fall by significantly more in the long-run than when estimated over the whole sample. For the hybrid VAR estimates, there is only a modest increase in the effects of monetary policy shocks relative to the results over the whole sample. Finally, the results using the baseline R&R approach are significantly revised downwards, with the peak effects on industrial production and unemployment falling to less than 3% and 0.66 percentage points respectively relative to initial results of more than a 4% drop in industrial production and over 0.90 percentage point increase in the unemployment rate. Thus, eliminating the period in which the FFR is generally perceived to be a poor proxy for the stance of monetary policy eliminates much of the quantitative difference in the real effects of monetary policy shocks across the three methods. Overall, these effects appear to be moderate and close to the original estimates from the hybrid VAR procedure.

2.4   Lag Specification

A final factor that can potentially account for the differences in results across specifications is the lag specification. The traditional VAR literature has commonly employed a lag length of one year, a practice which I have so far followed when estimating both VAR approaches. The single equation R&R specification, on the other hand, uses much longer lag lengths in estimating their second step to generate impulse responses: 24 months of autoregressive lags and 36 (48) months of lags of their shock measure for industrial production (prices). To assess the sensitivity of each approach to lag length, Figure 8 presents the estimated peak effects of monetary policy shocks on industrial production, unemployment

---

[19] The resumption date of January 1984 is significantly later than the actual end of the NBR targeting period but ensures that estimation procedures with long lag lengths do not overlap too much with the NBR period. When the latter occurs, I set monetary policy shocks during the period to be zero, which is consistent with R&R's procedure of setting shocks prior to 1969:3 equal to zero.

13

and prices from each specification for different lag lengths using both the whole sample and the restricted sample that omits the period of non-borrowed reserves targeting.[20]

While the estimated peak effects from the standard VAR are robust to lag length, the R&R specification is remarkably sensitive to the number of lags of shocks included in equation (1). Particularly noteworthy is the fact that the estimated peak effects of monetary policy shocks are increasing almost monotonically with the lag length. The hybrid VAR displays a similar sensitivity for lag lengths of less than one to two years. So while R&R employ long lag lengths to construct impulse responses, Figure 8 suggests that one could potentially reconcile the results of the VAR and R&R approaches by using shorter lag specifications, thereby yielding a consensus finding of small real effects of monetary policy shocks. However, this would require relying on lag lengths of well under a year, which would be at odds with both standard lag length specifications and the conventional wisdom that monetary policy shocks have very delayed macroeconomic effects.

One way to address this issue is via lag length selection criteria such as the AIC or BIC. However, some well-known issues arise in this context: while the BIC is consistent, it tends to select too few lags in short samples while the AIC asymptotically selects lag lengths that are too long (Ng and Perron 2005). Monte Carlo simulations of the single equation approach applied to industrial production confirm that the BIC fares very poorly in recovering the correct lag specification when the sample length is equal to that available in the data.[21] The AIC does much better on average, with only minor underestimates of the correct lag specification, but there is substantial uncertainty associated with AIC selections. For example, the Monte Carlo results suggest that if one were to observe an AIC lag selection of 12 months in our data, this would be within the 95% confidence interval of AIC selections for true models with lag lengths ranging from 3 months to 32 months.

As a result of the uncertainty associated with lag selection in the current context, I also consider a model-averaging procedure that takes into account these features of the data. Specifically, I generate a weighted average of impulse responses from specifications using different lag lengths. The weights come from a bootstrap of each specification under different lag lengths which quantifies the probability that this

---

[20] For the single equation approach, I hold the autoregressive lags constant at 24 but vary the number of lags of the shocks. Varying both the autoregressive lags and lags of the shock simultaneously yields very similar results as the ones in Figure 8.

[21] The specific Monte Carlo is as follows. I estimate equation (1) for industrial production assuming $I = 12$ but letting J vary from 3 to 36 months in 3 month increments. For each J, I take the estimated parameters and residuals of equation (1) and generate 1,000 times series of the same length as in the data with a burn-in period of 100 months. For each simulated time series, I then extract the optimal lag length according to the AIC and BIC. Results of the mean lag selections from each information criterion for the different true lag specifications are reported in Appendix Figure 1.

particular specification would yield the same AIC lag selection as identified in the data.[22] Appendix Table 1 presents results from a Monte Carlo exercise documenting that the model averaging approach a) leads to a smaller bias than the AIC, b) reduces the standard deviation across simulations of the estimated peak effect of monetary policy shocks relative to the AIC and c) reduces the mean squared error of estimated impulse responses relative to the AIC criterion.   The model-averaging procedure also dramatically improves upon the performance of specifications with too few lags and specifications based on the BIC while modestly improving upon the performance of specifications with too many lags.[23]

Table 2 presents the estimated peak effects of monetary policy shocks on macroeconomic variables from each empirical approach using the baseline lag specification of Romer and Romer (2004), the AIC lag selection, and the model-averaging approach. Using the AIC over the whole sample yields significantly shorter lag selections than assumed in the baseline specifications. As a result, the estimated peak effects of monetary policy shocks from the Romer and Romer VAR and single equation methods are reduced relative to the baseline lag length assumptions.   However, with the single equation approach, the peak effect of monetary policy shocks on industrial production falls only from -4.3% to -3.4% while the effect on the unemployment rate goes from 0.93% points to 0.61% points. Very similar results obtain with the model-averaging procedure. Thus, one cannot readily reconcile the standard VAR and R&R methods by appealing to shorter lag lengths, since both the AIC and model-averaging approaches to lag selection still point to sizeable real effects of monetary policy shocks.

On the other hand, Figure 8, which also plots the sensitivity of estimated peak effects controlling for the period of non-borrowed reserves targeting, suggests that the lag sensitivity is largely driven by the early Volcker era. This result mirrors the finding that the differential contractionary impetus across shock measures identified in section 2.2 stems from the NBR targeting period. Controlling for this time period eliminates much of the sensitivity to lag length, particularly for industrial production and unemployment. Table 2 presents estimated peak effects using the AIC and model-averaging approaches to lag selection over the restricted sample and documents that the estimated peak effects of monetary policy shocks are now broadly similar across methods, despite different lag specifications. For example, the estimated decline in industrial production ranges from -1.8% to -2.8% across methods and lag specifications. *Thus,*

---

[22] Note that this is similar in spirit to Bayesian model-averaging with equal priors on each lag specification, as in Kass and Raftery (1995). The model-averaging procedure is described in more detail in Appendix 1.

[23] Asymptotically, empirical specifications with too many lags should be unbiased but inefficient. In small samples, however, these specifications have an upward (in absolute value) bias in their estimates of the peak real effects of monetary policy shocks, as documented in Appendix Table 1, although this bias is small relative to that from underfitting the model. The fact that the model-averaging approach has a downward bias (in absolute value) also reflects the asymmetry from underfitting versus overfitting the model. By placing a disproportionate weight on longer lag lengths via the conditional probabilities, the model-averaging approach mitigates the downward bias from underfitting relative to the AIC but does not completely eliminate it. I am grateful to David Romer for pointing this out.

once one accounts for the period of non-borrowed reserves targeting, the real effects of monetary policy shocks appear to be in the medium range, with industrial production dropping by approximately 2-3% and unemployment rising by around half a percentage point.

III   Alternative Taylor-Rule Based Measures of Monetary Policy Shocks

To provide additional evidence on the size of the effects of monetary policy shocks, I consider three alternative measures of shocks, all linked by their adherence to the Taylor rule framework. The first approach follows R&R almost perfectly, but applies the suggestion of Hamilton (2008) and Sims and Zha (2006b) that one should explicitly allow for heteroskedastic errors in estimating policy rules. The second approach is based on Boivin (2006) and Coibion and Gorodnichenko (2011). Like R&R, they estimate Taylor rules using Greenbook forecasts for each meeting of the FOMC, but allow for time-varying parameters. This methodology therefore allows us to differentiate changes in the policy rule from innovations to the rule. The third approach follows Smets and Wouters (2007), who identify innovations to the Taylor rule from an estimated DSGE model. In addition, I revisit the implied contribution of monetary policy shocks to historical U.S. macroeconomic fluctuations.

3.1   Romer and Romer Shocks from GARCH Estimation of the Taylor Rule

In Figure 3, we can observe a common feature of most identified monetary policy shock measures: a sharp increase in the volatility of these shocks between 1979 and 1982, the period in which the Federal Reserve abandoned targeting the Federal Funds Rate. Hamilton (2008) and Sims and Zha (2006b) each forcefully argue that this type of heteroskedasticity is so severe that it matters not just for the calculation of standard errors but also for the estimation of the parameters of the Fed's response function. Hamilton (2008), for example, argues that once one estimates the Taylor rule by GARCH instead of OLS, there is much weaker evidence of a change in the Fed's response to inflation and output when Volcker became Chairman, a result that reflects not just different standard errors but also different point estimates. Similarly, Sims and Zha (2006b) estimate regime-switching VAR's and find that, once one allows for heteroskedasticity, what changes across regimes is only the variances of structural disturbances, with the Volcker reserves-targeting episode standing out as a period of high variance in the disturbances to the policy rule.

To assess the importance of time-varying volatility of monetary policy innovations in the Taylor rule estimated by R&R, I replicate their OLS estimation and find that one can overwhelmingly reject the null of no heteroskedasticity using an ARCH test with one lag at the 1% significance level. As a result, I replicate their baseline Taylor rule regression using GARCH(1,1) and extract the residuals, which I refer to as R&R GARCH shocks. The correlation between this measure of policy shocks and the original R&R

shocks is very high, 0.93, with Appendix Figure 2 presenting a scatter plot of this new measure of policy innovations relative to the original R&R shocks.[24] The shocks line up very closely with the 45 degree line, with the exception of the period of non-borrowed reserves targeting, in which the GARCH shocks are consistently higher than the original R&R shocks. Appendix Figure 3 plots the cumulative sum of the two shock measures. While they track each other closely overall, the GARCH shocks point to much more contractionary policy between 1979 and 1982, and looser monetary policy in the early 1970s.

Figure 9 reproduces the sensitivity results of section 2 using this alternative measure of policy shocks. Using the whole sample and the same lag specification as R&R, the estimated peak effects of monetary policy shocks are medium: a one-hundred basis point innovation lowers industrial production by 2 percent and raises the unemployment rate by 0.60 percentage points. The effect on prices is also substantially dampened. The results display little sensitivity to specific historical episodes, so dropping shocks from any 3-month period has little effect on the estimated peak effects of monetary policy shocks. The results are also insensitive to dropping the early Volcker years. Finally, the estimated peak effects of monetary policy shocks, while still increasing in lag length, are less sensitive to lag selection issues than the original R&R series. In short, explicitly controlling for the GARCH effects associated with the period of non-borrowed reserves targeting in the estimation of the Taylor rule, eliminates most of the sensitivity issues identified in section 2, and yields estimated real effects of monetary policy shocks which are again in the medium range.

3.2   Shocks from an Estimated Taylor Rule with Time-Varying Parameters

Both the original R&R shocks as well as the GARCH measure considered in section 3.1 are defined as the residuals from an estimated Taylor rule over the whole sample. R&R argue that the goal of this regression is not to estimate the Fed's reaction function as well as possible, but rather to purge the intended funds rate series of movements taken in response to useful information about future economic developments. The resulting measure of monetary policy shocks includes factors such as the evolution of the Fed's operating procedures, policy-makers' beliefs about the workings of the economy, the tastes and goals of the Federal Reserve, politics, the pursuit of other objectives (such as affecting the exchange rate), and random variation in policy. Conceptually, the latter is equivalent to the innovations to the central

---

[24] Appendix Table 2 presents the estimated coefficients from the original R&R specification versus those from the GARCH specification. Because there are some missing observations before November 1972, the GARCH equation is estimated over the restricted sample of 1972:11-1996:12. Appendix Table 1 also presents coefficient estimates from estimating the same Taylor rule by OLS over this restricted sample. The GARCH shocks are extrapolated back to 1969:3 using the coefficients from the estimates in Appendix Table 1. Note that the correlation between the original R&R shocks and the shocks extracted from estimating the Taylor rule by OLS over the restricted sample is 0.996. Finally, Appendix Table 2 also reports coefficient estimates from the conditional variance equation, in which the coefficients on both the lagged squared residual and the lagged forecast variance are significant at the 1% level.

bank's rule, while the others can be interpreted as changes in the rule itself. The potential problem with grouping these different factors together is that variations in the rule may lead to different dynamics than innovations to the rule. For example, the effects of a change in the policy rule will generally depend on how quickly agents realize that a regime change has occurred, how long the regime change is expected to last, etc…

To address this possibility, I consider an alternative approach to extract those monetary policy shocks which might more precisely distinguish innovations to the rule from innovations in the rule. Specifically, I follow Boivin (2006) and Coibion and Gorodnichenko (2011) and estimate a restricted version of the Taylor rule in R&R with time-varying parameters:

$$\Delta f_t = c_t + \rho_{1,t} f b_m + \rho_{2,t} f b_{m-1} + \varphi_{\pi,t} F_t \pi_{t+1,t+2} + \varphi_{gy,t} F_t g y_t + \varphi_{ue,t} F_t u e_t + \varepsilon_t. \quad (3)$$

As in R&R, the left-hand side variable is the intended change in the Federal Funds Rate decided upon at the FOMC meeting at time t. The control variables include the level of the intended FFR immediately preceding the meeting (fb_m), the FFR preceding the previous meeting (fb_{m-1}), the forecast of average inflation over the subsequent two quarters ($F_t \pi_{t+1,t+2}$) , the forecast of output growth in the current quarter ($F_t g y_t$), and the forecast of the unemployment rate in the current quarter ($F_t u e_t$). Like the R&R Taylor rule, this is estimated at the frequency of FOMC meetings using Greenbook forecasts of macroeconomic variables and the R&R measure of interest rate changes at each meeting. Unlike R&R, this Taylor rule allows for time-varying response coefficients (which are assumed to follow a random walk, as in Boivin 2006).[25] Thus, those policy changes which would commonly be interpreted as regime change, i.e. changes in the coefficients of the Fed's response function, will generally not be classified as monetary policy shocks in this approach. In addition, changes in the central bank's target rate of inflation or target for GDP growth will be incorporated in the time-varying intercept.[26] Only the random innovations to the rule will be classified as monetary policy shocks. However, this specification of the Taylor rule imposes a more restricted information set than that used in the baseline R&R approach. For example, the Taylor rule estimated by R&R controls for forecasts of lagged, contemporaneous and future values of inflation and output growth while equation (3) imposes specific horizons to which the Fed responds.[27] The Taylor rule estimated by R&R also controls for revisions in the forecasts of inflation and output growth, while

---

[25] The estimation allows for two breaks in the volatility of shocks to the parameters in 1979 and 1982. The time-varying coefficients are presented in Appendix Figure 4.

[26] This helps address the concern, raised by Sims and Zha (2006b), that time-invariant Taylor rules are implausible because they imply that monetary authorities react to expected inflation of 3% in exactly the same way, whether the recent past level of inflation has been 1.5% or 6%. The time-varying Taylor rule only implies that the Fed responds to deviations from trend inflation in the same way, but allows for trend inflation to vary.

[27] The specific horizons chosen are based on Coibion and Gorodnichenko (2011), who employ an AIC selection criterion over a wide variety of forecasting horizons.

the time-varying specification focuses exclusively on contemporaneous forecasts. I will refer to the monetary policy shocks from the time-varying coefficients approach as R&R TVC shocks.

The correlation between the original R&R and the R&R TVC shocks is quite high at 0.82 (0.86 with R&R GARCH shocks). Consistent with the notion that the R&R TVC shocks are excluding some component of interest rate innovations as regime changes, the standard deviation of R&R TVC shocks is ten percent lower than that of the original R&R shocks. Appendix Figure 5 plots a scatter of the two shocks and Appendix Figure 6 plots the cumulative sum of the two shock measures. The cumulative measures exhibit some small differences. For example, the sequence of negative policy shocks between 1969 and 1972 is only half as big with the R&R TVC shocks, indicating that a significant component of the loosening in monetary policy over this time period can be attributed to a changing response to macroeconomic variables. Similarly, the original R&R shocks point to unusually tight monetary policy in the mid 1990s, while the TVC approach attributes much of these interest rate changes to innovations in the Fed's endogenous response to macroeconomic variables.

Figure 10 presents the results from applying the sensitivity analysis from section 2 to this alternative measure of policy shocks. The results are broadly similar to those using the original R&R shocks. First, the estimated peak effects of monetary policy shocks using the baseline lag specification of R&R are close to the baseline result of R&R: a one-hundred basis point innovation to the FFR yields a peak decline in output of approximately 4% and a rise in the unemployment rate of nearly 0.75 percentage points, while the price level falls by about 4%. The results are less sensitive to any 3-month sequence of shocks than the original R&R shocks, and the peak effects of monetary policy shocks decrease (in absolute value) when dropping the period in which the Fed abandoned targeting the FFR. Finally, the peak effects of monetary policy shocks on industrial production and unemployment are much less sensitive to the lag specification: once the lag length exceeds one year, the results are essentially unchanged for longer lag lengths. Thus, this setup presents an alternative interpretation for the sensitivity of both VAR and R&R shocks to the 1979-1982 period: the difficulty of separating between those policy changes that should be characterized as a change in the rule versus those that should be characterized as a pure innovation to the rule.

3.3 Shocks from a Taylor Rule in an Estimated DSGE Model

Along with VAR's and single equation estimates of the central bank's reaction function, another very common approach to studying monetary policy is within the context of estimated DSGE models. A well-known example is the New Keynesian model of Smets and Wouters (2007) which includes price and wage frictions along with a plethora of shocks. Monetary policy is modeled via a Taylor rule in which the monetary authority is assumed to respond to contemporaneous inflation, output growth and the output

gap. Like the standard monetary VAR, however, this approach does not explicitly incorporate the actual forecasts of the central bank.[28]

For comparison with the previous measures of monetary policy shocks, I extract the filtered monetary policy shock series from Smets and Wouters (2007). These shocks are available quarterly over the same sample as the R&R shocks. The correlation between the two shock series is surprisingly high (0.40) given the very different methodologies, and the Smets-Wouters measure is 69% more volatile, as measured by the standard deviation of the shocks. Appendix Figures 7 and 8 display a scatter of the two shock series and their cumulative sums respectively. While the latter tell qualitatively similar stories, the Smets-Wouters shocks point to much more contractionary innovations since the early 1980s than the original R&R shocks, with the sum of the shocks rising from approximately 5% in 1984 to over 20% by the end of 1996. Figure 11 reproduces the sensitivity results of section 2 but applied to the Smets-Wouters shocks using quarterly data. The results closely resemble those found using the R&R GARCH shocks. First, the estimated peak effects using the lag structure from R&R (converted to quarterly frequency) are in the small to medium range: a one-hundred basis point innovation lowers industrial production by a maximum of 1.5% and raises the unemployment rate by slightly less than 0.40 percentage points, when estimated over the whole sample, while the price level is predicted to fall by approximately 1.5%. These smaller estimates of monetary policy shocks likely reflect the same features as in the standard VAR: 1) the use of the actual FFR rather than intended policy changes and 2) not controlling for the real-time information available to the Fed. The results are insensitive to any quarterly shock being dropped from the estimation and are also insensitive to dropping the period of reserves targeting. Finally, the results are much less sensitive to the assumed lag length than the original specification of R&R.

3.4   Revisiting the Historical Contribution of Monetary Policy Shocks

The starkest contrast between the traditional VAR literature and the Romer and Romer (2004) analysis of monetary policy shocks, as emphasized in the introduction and illustrated in Figure 1, lays in the historical contribution of monetary policy shocks to macroeconomic fluctuations implied by each approach. Whereas standard monetary VAR's point to little contribution of monetary policy shocks to historical US fluctuations, the baseline approach of R&R implies that the brunt of the annualized changes in production, employment and prices can be explained by monetary policy innovations. In part, the latter reflects the large estimated macroeconomic effects of monetary policy shocks from the baseline R&R

---

[28] A third approach to estimating the Taylor rule is the framework proposed in Clarida, Gali and Gertler (2000). However, because this approach replaces ex-ante forecasts with ex-post values, the estimated residual combines monetary policy innovations and the rational expectations error.

approach.[29]   In this section, I revisit the implied historical contribution of monetary policy shocks to macroeconomic fluctuations taking into account the different issues identified in previous sections. Using the original R&R shocks, the GARCH R&R shocks, the TVC R&R shocks, and the Smets-Wouters shocks, I use estimates of equation (1) dropping the early Volcker period as in section 2.3 to construct counterfactual time series of industrial production, unemployment and inflation driven exclusively by exogenous monetary policy shocks. While the estimation excludes the 1979:10-1983:12 period, the counterfactuals are constructed using the shocks over the entire sample.[30]

As illustrated in Figure 12, all four measures point to a nontrivial contribution of monetary policy shocks to historical U.S. macroeconomic fluctuations. The estimated contributions are considerably smaller than implied by the original R&R results but also considerably larger than those implied by the baseline VAR. In particular, much of the volatility in production during the 1970's is attributable to monetary policy shocks, consistent with the "stop-go" description of monetary policy during this period emphasized in Romer and Romer (2002). Every shock measure also points to an important contribution of monetary policy shocks to the recessions of the early 1980s. Across methods, exogenous monetary policy actions appear to have contributed approximately 2-3% points out of the 5% point increase in unemployment between 1979 and 1982.

In addition, one should note that the contribution of monetary policy shocks to real fluctuations has drastically declined since the early 1980s. This is consistent with the argument made in Clarida et al. (2000), Lubik and Schorfheide (2004) and Boivin and Giannoni (2006) that improved monetary policy (in this case via smaller innovations to the policy rule) contributed to the Great Moderation. For example, all four shock measures imply that the 1990 recession was not directly caused by exogenous contractionary monetary policy. Finally, monetary policy shocks can account for a significant component of the historical volatility in inflation. In particular, the disinflation of the early 1980s can readily be attributed to exogenous monetary policy, a feature not shared by standard monetary VAR's.

---

[29] The large role assigned to monetary policy fluctuations using the baseline R&R approach in Figure 1 reflects both the large estimated effects of policy shocks and the specific sequence of policy shocks identified in R&R. If one feeds the VAR shocks into R&R's estimate of equation (1), the implied contribution of monetary policy shocks to historical fluctuations remains small. Thus, the difference in identified shocks plays a role above and beyond just generating different impulse responses of policy shocks.

[30] For industrial production and the unemployment rate, I estimate equation (1) using 24 autoregressive lags and 24 lags of monetary policy shocks. As illustrated in Figure 8, the effects of monetary policy shocks are relatively insensitive to the lag structure once the period of non-borrowed reserves targeting is accounted for. Because this is not the case for prices, counterfactual inflation paths are based on applying the model-averaging procedure from section 2.4 to each set of monetary policy shocks excluding the 1979:10-1983:12 period. The displayed counterfactual paths are weighted averages of counterfactuals using different lag specifications where the weights come from the model-averaging estimates.

IV   Conclusion

This paper revisits a basic quantitative issue for monetary policy analysis: how large are the effects of monetary policy shocks? The current state of the literature is mixed, with monetary VAR's commonly pointing to small effects of monetary policy shocks (and a minimal contribution of these innovations to historical US macroeconomic fluctuations) while the Taylor rule approach of Romer and Romer (2004) implies much larger effects of monetary shocks, both for impulse responses and historical decompositions. I find that these differences are driven by several factors. First, unanticipated monetary policy contractions in a VAR are associated with smaller and less persistent increases in interest rates than equivalent innovations from the R&R methodology: adjusting for the size of the contractionary impetus in interest rates eliminates much of the quantitative differences in the impulse responses across methods. Second, the assumed lag structure also affects the magnitude of estimated peak effects of policy shocks. While the VAR estimates are largely robust to this issue, the two-step procedure of R&R is not: their estimated peak effects are nearly monotonically increasing in the lag length of the shocks. Using either the AIC or the model-averaging procedure considered in this paper points to smaller effects of policy innovations than originally found by R&R, albeit still significantly larger than traditional estimates from the VAR literature.

Third, the period in which the Federal Reserve abandoned targeting the Fed Funds Rate plays a disproportionate role in driving the estimated peak effects of policy shocks. This is problematic since each approach relies on the FFR as the primary measure of the stance of monetary policy-makers. Dropping this time period from the analysis helps reconcile the VAR and R&R estimates and yields moderate effects of policy shocks. Similarly, explicitly controlling for the time-variation in the shock structure when estimating the Taylor rule also leads to more moderate effects of policy shocks than originally uncovered by Romer and Romer (2004). Nonetheless, my results confirm the broader point of Romer and Romer (2004) that monetary policy shocks have likely played a more important role in accounting for historical macroeconomic fluctuations than is commonly implied by monetary VAR's, particularly during the 1970s and early 1980s.

Finally, the apparent sensitivity of the results to the period of non-borrowed reserves targeting during the early Volcker years presents a cautionary tale for future monetary policy analysis that will have to wrestle with the current period of zero interest rates and the associated unorthodox monetary policy actions pursued by the Fed and other central banks. Properly measuring the stance of policy and identifying policy innovations is rife with difficulties even in calm periods. Doing so during extraordinary circumstances such as the Volcker experiment or the current period presents a particularly challenging exercise for monetary policy analysis.

References

Barsky, Robert B. and Lutz Kilian. 2002. "Do We Really Know that Oil Caused the Great Stagflation? A Monetary Alternative?" In *NBER Macroeconomics Annual 2001,* ed. Ben S. Bernanke and Kenneth S. Rogoff, 137-183. Cambridge, MA: The MIT Press.

Bernanke, Ben S and Alan S. Blinder. 1992. "The Federal Funds Rate and the Channels of Monetary Transmission." *American Economic Review*, 82(4): 901-21.

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: 91–142.

Bernanke, Ben S., Jean Boivin, and Piotr Eliasz. 2005, "Measuring the Effects of Monetary Policy: A Factor-Augmented Vector Autoregressive (FAVAR) Approach." *Quarterly Journal of Economics,* 120(1): 387-422.

Bernanke, Ben S. and Ilian Mihov. 1998. "Measuring Monetary Policy." *Quarterly Journal of Economics,* 113(3): 869-902.

Blanchard, Olivier, 1993. "Consumption and the Recession of 1990-1991." *American Economic Review,* 83(2): 270-274.

Boivin, Jean, and Marc Giannoni. 2006. "Has Monetary Policy Become More Effective?" *Review of Economics and Statistics,* 88(3): 445-462.

Boivin, Jean. 2006. "Has U.S. Monetary Policy Changed? Evidence from Drifting Coefficients and Real-Time Data." *Journal of Money, Credit and Banking,* 38(5): 1149-1173.

Boschen, John F. and Len O. Mills. 1995. "The Relation Between Money Market and Narrative Indicators of Monetary Policy." *Economic Inquiry,* 33(1): 24-45.

Christiano, Lawrence, Martin Eichenbaum, and Charles Evans. 1999. "Monetary Policy Shocks: What Have We Learned, and To What End?" In *Handbook of Monetary Economics 1999,* ed. John B. Taylor and Michael Woodford, 65-148. Amsterdam: Elsevier Science.

Clarida, Richard, Jordi Galí, and Mark Gertler. 2000. "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory." *Quarterly Journal of Economics,* 115(1): 147-180.

Coibion, Olivier and Yuriy Gorodnichenko. 2011. "Monetary Policy, Trend Inflation and the Great Moderation: An Alternative Interpretation." *American Economic Review,* 101(1): 341-370.

Faust, Jon, Eric T. Swanson, and Jonathan H. Wright. 2004. "Identifying VARs based on High Frequency Futures Data." *Journal of Monetary Economics,* 51(6): 1107-1131.

Gorodnichenko, Yuriy. 2006. "Reduced rank identification of structural shocks in VARs." http://www.econ.berkeley.edu/~ygorodni/.

Hall, Robert E. 1993. "Macro Theory and the Recession of 1990-1991." *American Economic Review,* 83(2): 275-279.

Hamilton, James D. 2008. "Macroeconomics and ARCH." In *Volatility and Time Series Econometrics: Essays in Honor of Robert Engle*, ed. Tim Bollerslev, Jeffry R. Russell and Mark Watson. Oxford: Oxford University Press.

Kass, Robert E. and Adrian E. Raftery. 1999. "Bayes Factors." *Journal of the American Statistical Association,* 90(430): 773-795.

Kilian, Lutz. 2009, "Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market." *American Economic Review*, 99(3): 1053-1069.

Leeper, Eric M., Christopher A. Sims, and Tao Zha. 1996, "What Does Monetary Policy Do?" *Brookings Papers on Economic Activity,* 1996(2): 1-78.

Lubik, Thomas A., and Frank Schorfheide. 2004. "Testing for Indeterminacy: An Application to U.S. Monetary Policy." *American Economic Review,* 94(1): 190-217.

Ng, Serena and Pierre Perron. 2005, "A Note on the Selection of Time Series Models." *Oxford Bulletin of Economics and Statistics,* 67(1): 115-134.

Romer, Christina D. and David H. Romer. 2004. "A New Measure of Monetary Shocks: Derivation and Implications." *American Economic Review,* 94: 1055-1084.

Romer, Christina D. and David H. Romer. 2002. "The Evolution of Economic Understanding and Postwar Stabilization Policy." Federal Reserve Bank of Kansas City *Proceedings*, 2002: 11-78.

Schreft, Stacey L.. 1990. "Credit Controls: 1980." Federal Reserve Bank of Richmond *Economic Review*, 1990(Nov): 25-55.

Sims, Christopher A. and Tao Zha. 2006a. "Does Monetary Policy Generate Recessions?" *Macroeconomic Dynamics,* 10(2): 231-272.

Sims, Christopher A. and Tao Zha. 2006b. "Were There Regime Switches in U.S. Monetary Policy?" *American Economic Review,* 96(1): 54-81.

Smets, Frank and Rafael Wouters. 2007. "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach." *American Economic Review,* 97(3): 586-606.

Taylor, John B. 1993. "Discretion versus Policy Rules in Practice." *Carnegie Rochester Conference Series on Public Policy*, 39(1): 195-214.

Uhlig, Harald. 2005. "What are the effects of monetary policy on output? Results from an agnostic identification procedure." *Journal of Monetary Economics,* 52(2): 381-419.

Walsh, Carl E. 1993. "What Caused the 1990-1991 Recession?" Federal Reserve Bank of San Francisco *Economic Review,* 2: 33-48.

Figure 1: The Contribution of Monetary Policy Shocks to Business Cycle Fluctuations
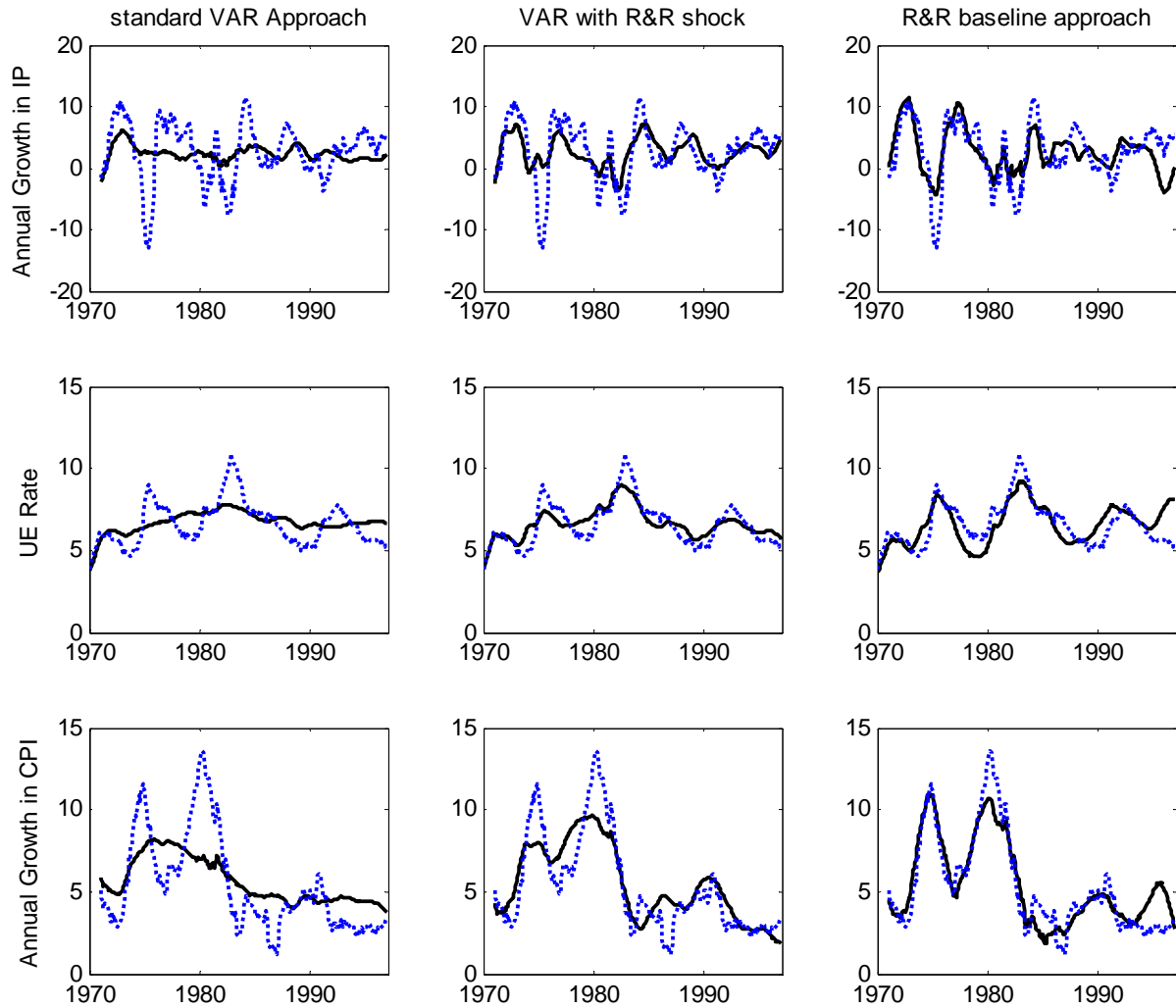

Note: The dashed blue lines are actual macroeconomic variables, while the solid black lines are the predicted time paths of each macroeconomic variable from identified monetary policy shocks from each approach. The *Standard VAR Approach* corresponds to the baseline VAR, the *VAR with R&R Shocks* corresponds to the baseline VAR using the cumulative measure of R&R shocks in place of the effective Federal Funds Rate, and *R&R Baseline Approach* corresponds to the estimation procedure of Romer and Romer (2004). See section 1 in the text for details.

Figure 2: Impulse Responses of Macroeconomic Variables to Monetary Policy Shocks
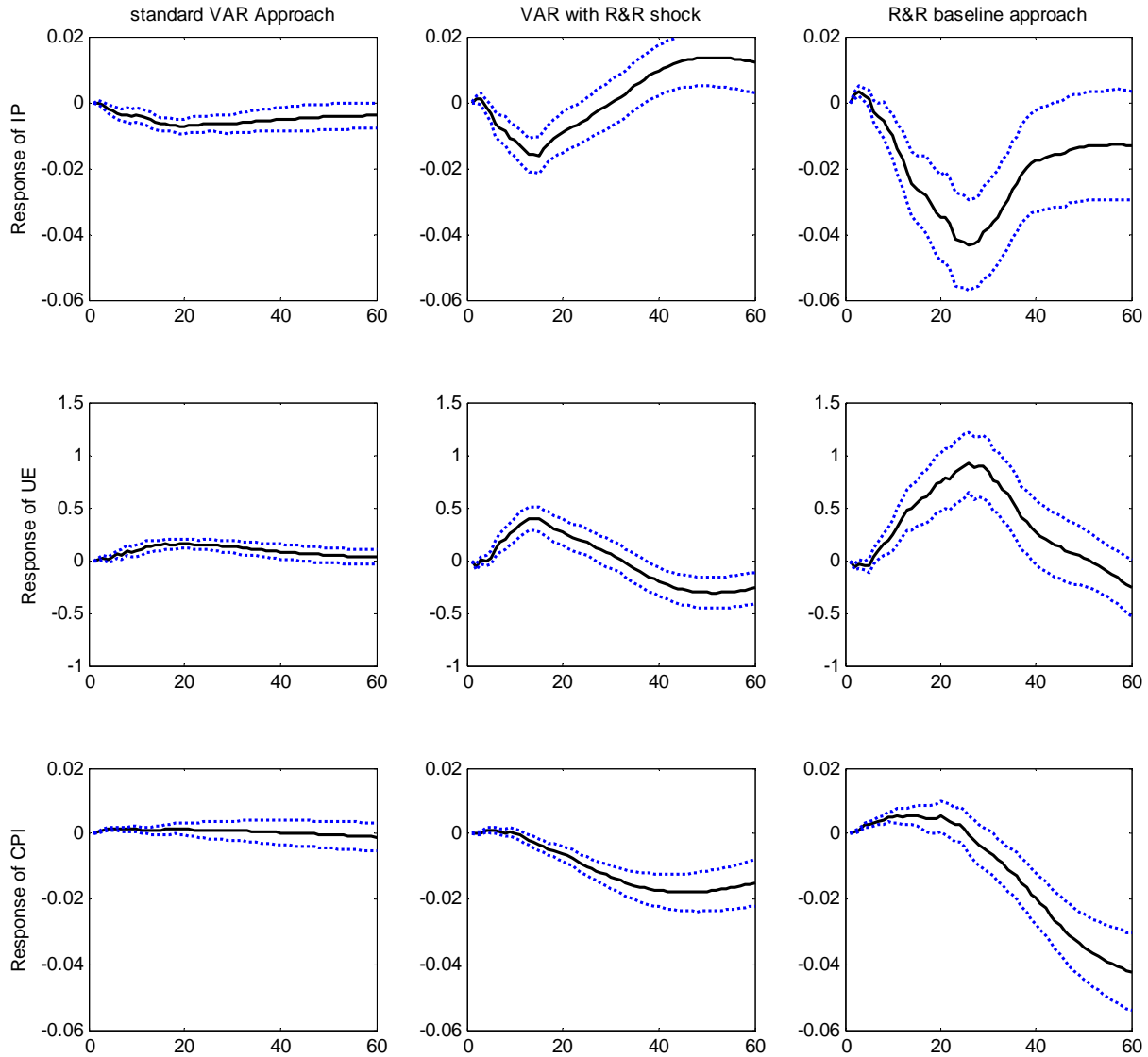
Note: The solid black lines are the impulse response functions of each macroeconomic variable to a 100 basis point monetary policy innovation. The dashed blue lines are one standard deviation intervals from a bootstrap. The *Standard VAR Approach* corresponds to the baseline VAR, the *VAR with R&R Shocks* corresponds to the baseline VAR using the cumulative measure of R&R shocks in place of the effective Federal Funds Rate, and *R&R Baseline Approach* corresponds to the estimation procedure of Romer and Romer (2004). The horizontal axis is in months. See section 1 in the text for details.

Figure 3: Identified Monetary Policy Shocks from each Approach
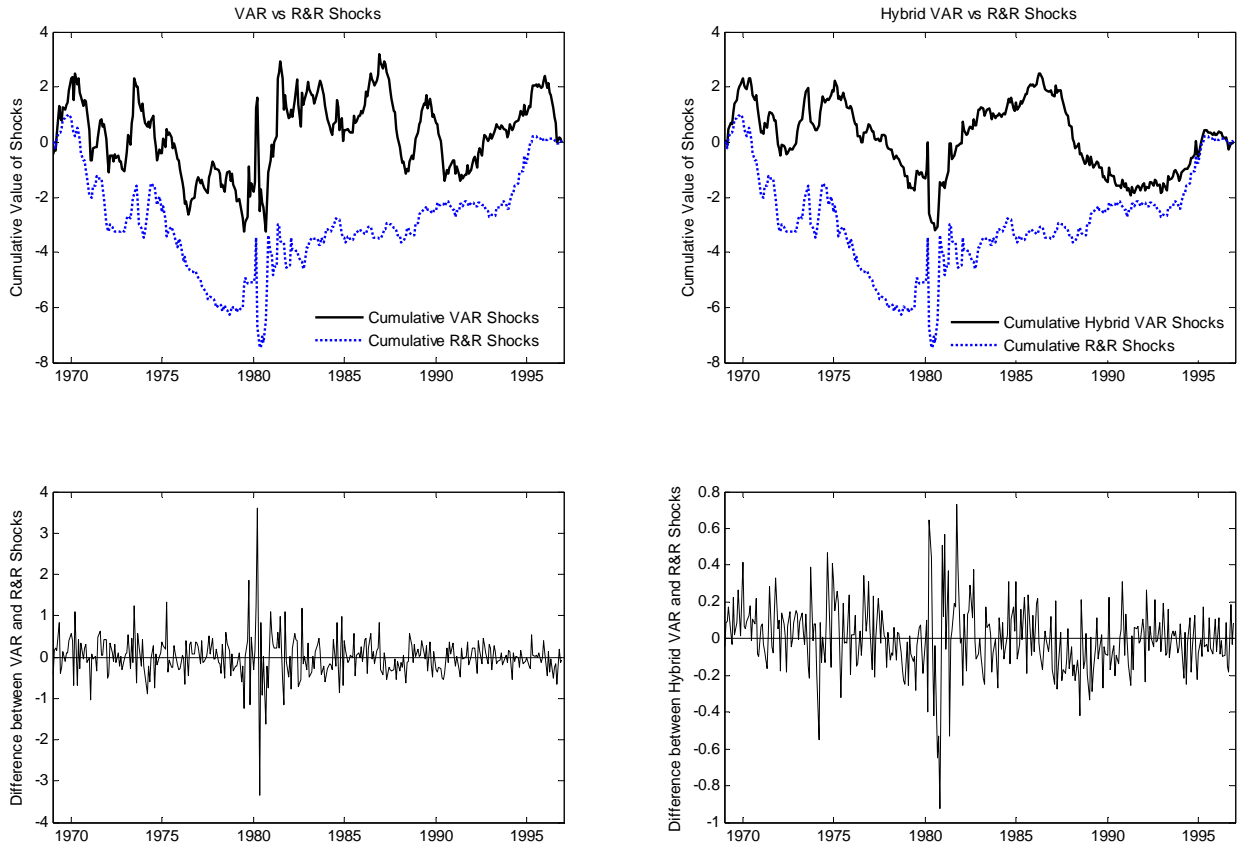
Note: The top two figures plot the cumulative sum of identified monetary policy innovations from each approach described in the text. The *Cumulative VAR Shocks* corresponds to the sum of Choleski-identified innovations to the effective FFR in the baseline VAR, the *Cumulative R&R Shocks* corresponds to the sum of the monetary policy shocks identified in Romer and Romer (2004), and the *Cumulative Hybrid VAR Shocks* corresponds to the sum of Choleski-identified innovations to the cumulative R&R shock measure added to the baseline VAR in place of the effective Federal Funds Rate. The bottom two panels plot the difference between the baseline VAR (left) and the hybrid VAR (right) shocks relative to the R&R shocks each month. See section 2.1 in the text for details.

Figure 4: Impulse Responses to Monetary Policy Shocks using Single Equation Approach
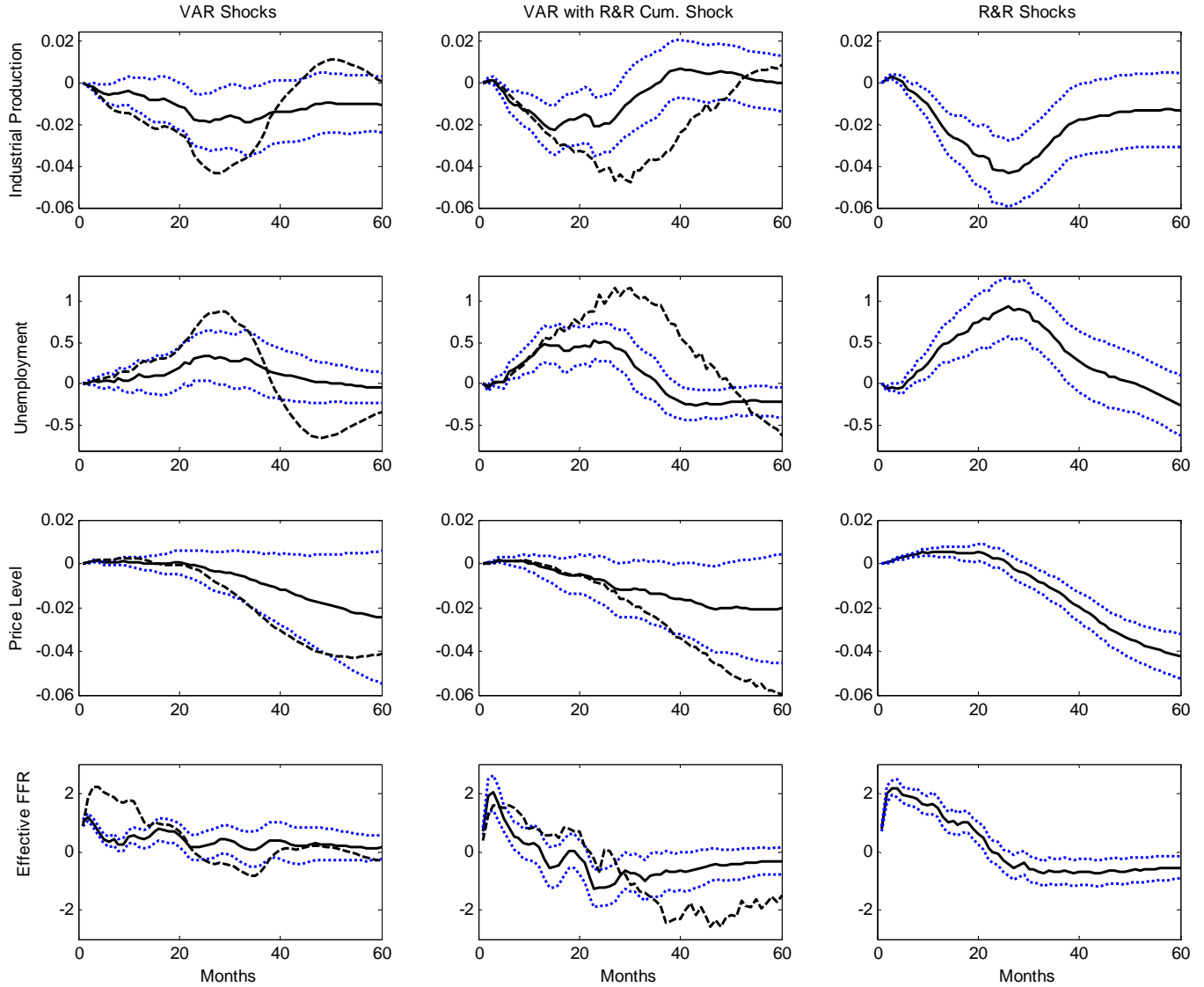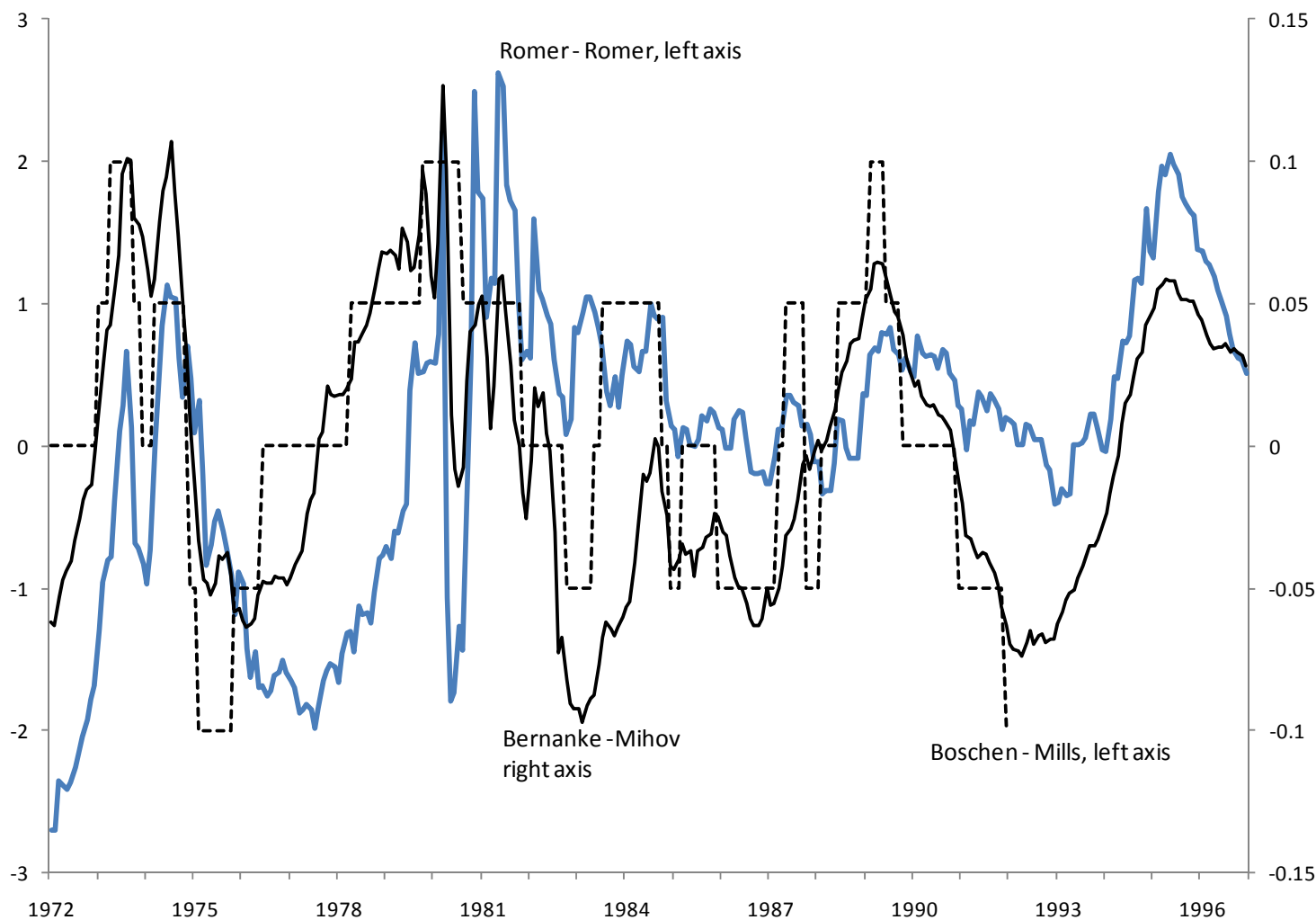

Note: The solid black lines are the impulse response functions of each macroeconomic variable to a 100 basis point monetary policy innovation from each identification procedure. Each impulse response is estimated following Romer and Romer (2004). The dotted blue lines are one standard deviation confidence intervals from a bootstrap. The *VAR Shocks* correspond to impulse responses using monetary policy shocks identified using the baseline VAR. The *VAR with R&R Cum. Shocks* correspond to impulse responses using monetary policy shocks identified using the VAR with R&R cumulative shocks in place of the effective Federal Funds Rate. The *R&R Shocks* correspond to impulse responses using monetary policy shocks identified in Romer and Romer (2004). The dashed black lines are the counterfactual impulse responses to a sequence of shocks designed to replicate the IRF of the Federal Funds Rate to R&R shocks. See section 2.2 for details.

## Figure 5: Sensitivity of Peak Effects of Monetary Policy Shocks to Individual Episodes
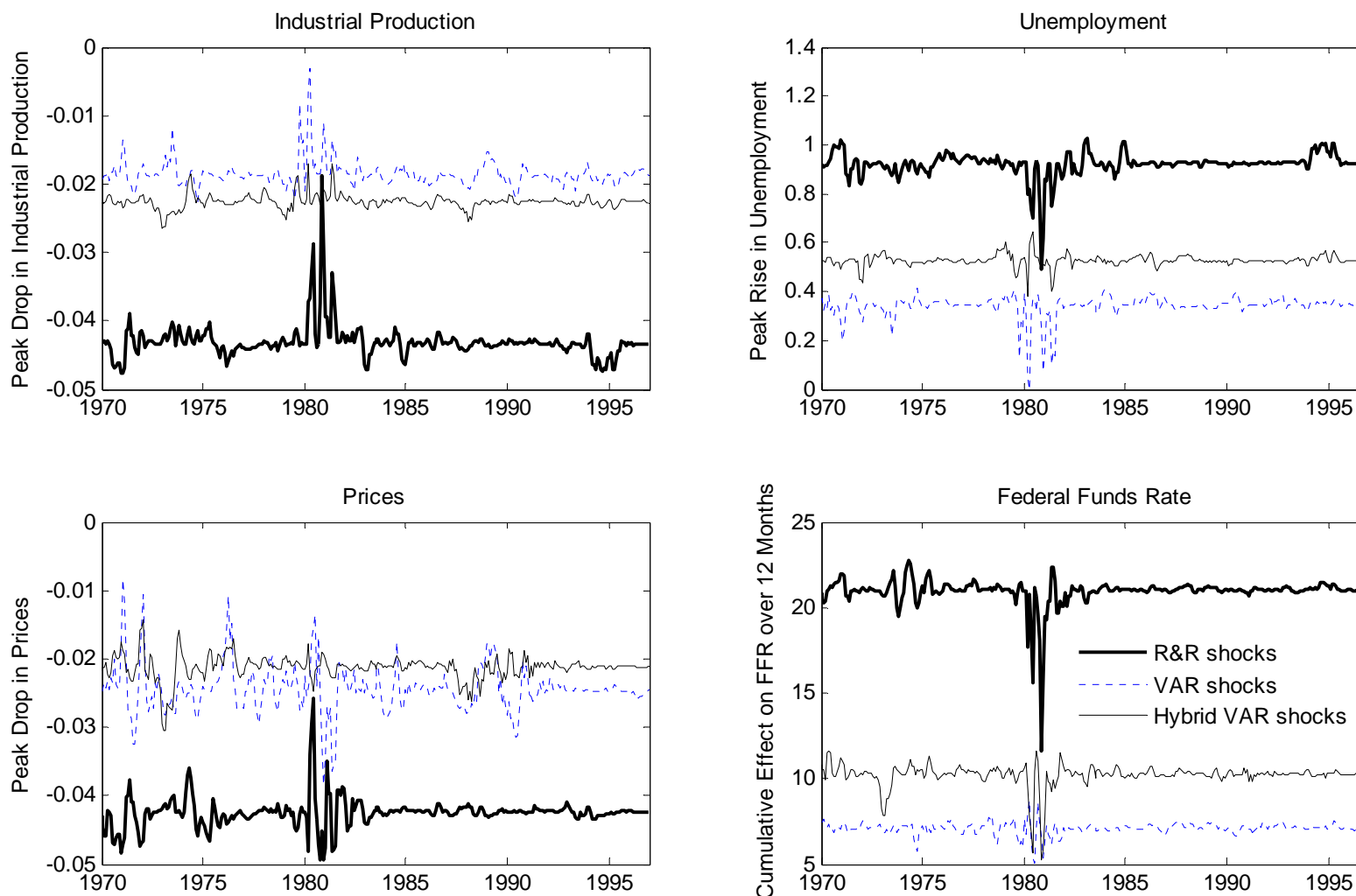
Note: Each line presents the peak effect of monetary policy shocks on macroeconomic variables from estimating equation (1) using alternative measures of monetary policy shocks. For each time period, the monetary policy shock for that month and the two previous months are set equal to zero, so the lines present the sensitivity of estimated peak effects to 3-month time intervals. The bottom right panel measures the sensitivity of the cumulative interest rate response to each shock measure over the first 12 months of impulse responses. The *VAR Shocks* correspond to monetary policy shocks identified using the baseline VAR. The *Hybrid VAR Shocks* correspond to monetary policy shocks identified using the VAR with R&R cumulative shocks in place of the effective Federal Funds Rate. The *R&R Shocks* correspond to monetary policy shocks identified in Romer and Romer (2004). See section 2.2 for details.

**Figure 6: Measures of the Stance of Monetary Policy**

Note: The solid blue line is the cumulative Romer and Romer (2004) shocks (minus the 36-month moving average of its lagged values), the thin black line is the Bernanke and Mihov (1998) measure of the stance of monetary policy, while the dashed black line is the Boschen and Mills (1995) indicator of the stance of monetary policy. For each series, a higher value is associated with tighter monetary policy. See section 2.3 in the text for details.

Figure 7: Impulse Responses to Monetary Policy Shocks Omitting the Early Volcker Period
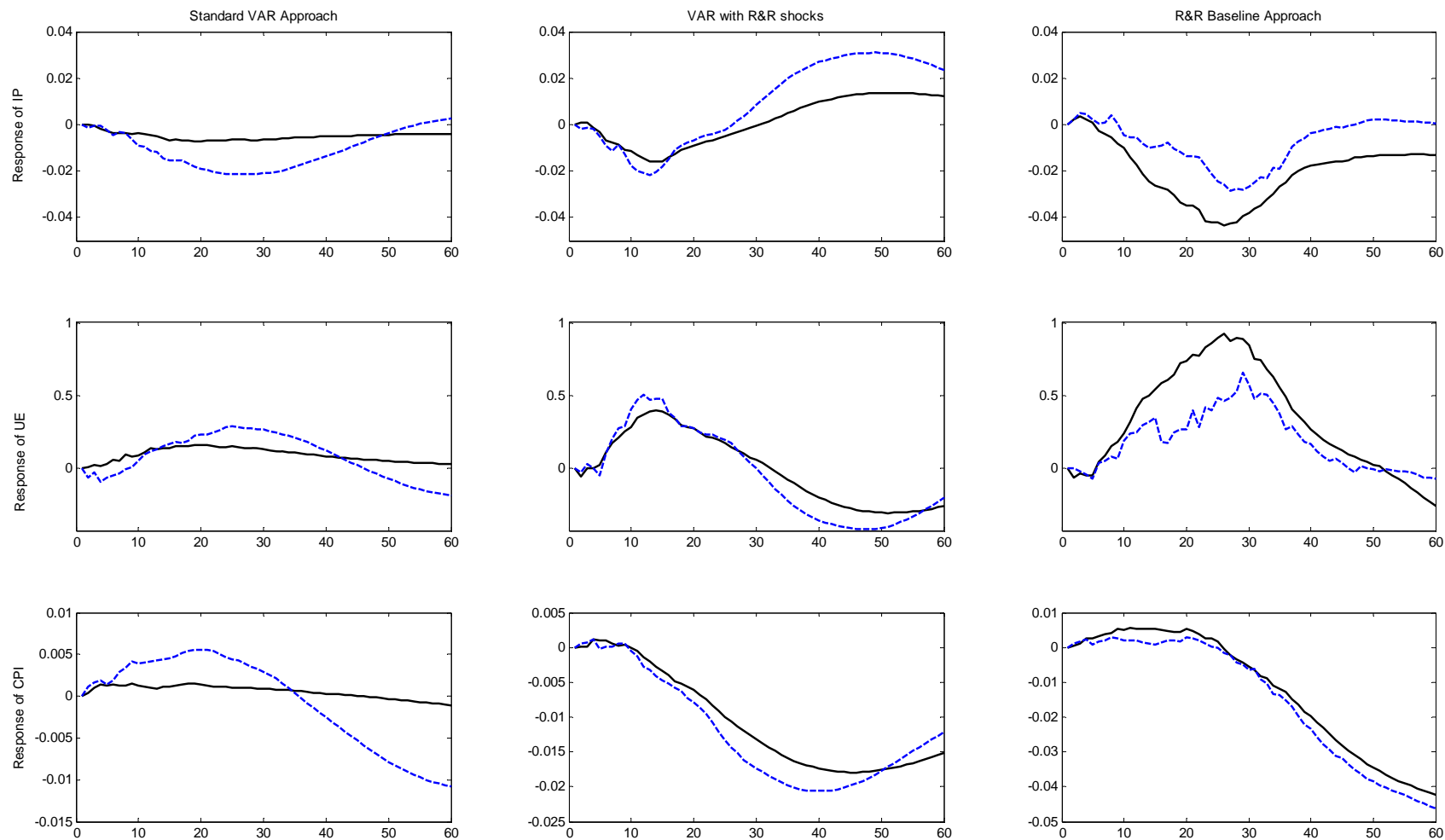

Note: The solid black lines are the impulse responses of each macroeconomic variable to a 100 basis point monetary policy innovation from each identification procedure using the entire time sample. The blue dashed lines are the impulse responses dropping the period from 1979:10-1983:12. The *Standard VAR Approach* corresponds to the baseline VAR, the *VAR with R&R Shocks* corresponds to the VAR using the cumulative measure of R&R shocks in place of the effective Federal Funds Rate, and *R&R Baseline Approach* corresponds to the estimation procedure of Romer and Romer (2004). The horizontal axis is in months. See section 2.3 in the text for details.

Figure 8: Sensitivity of Peak Effects of Monetary Policy Shocks to Lag Lengths

Panel A: Whole Sample


Panel B: Dropping Period of Non-FFR Targetting
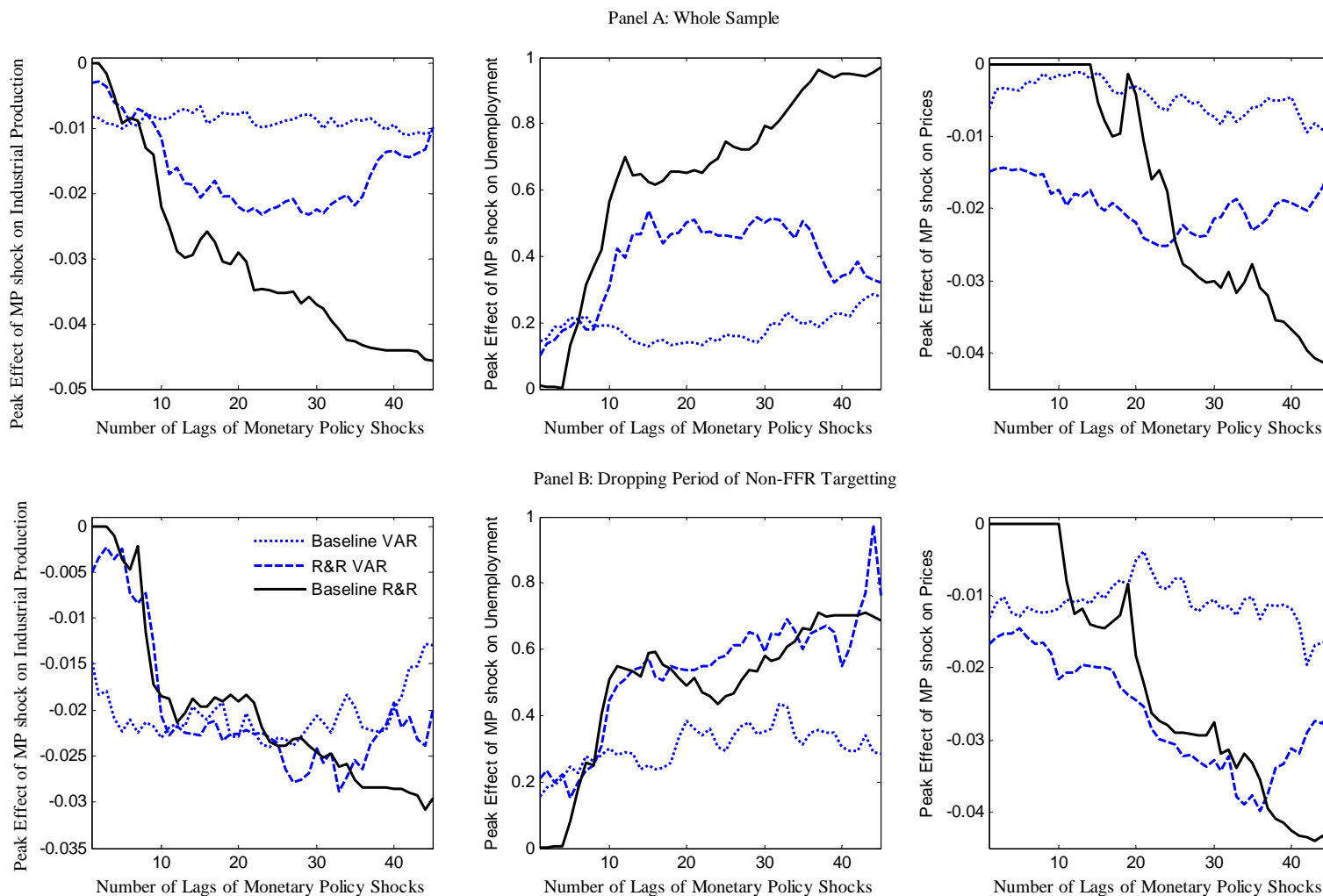
Note: Each line presents the estimated peak effect of monetary policy shocks on macroeconomic variables for different lag specifications. The *Baseline VAR* lines correspond to peak effects of monetary policy shocks identified using the baseline VAR with lag length given on the horizontal axis. The *R&R VAR* lines correspond to peak effects of monetary policy shocks identified using the VAR with R&R cumulative shocks in place of the effective Federal Funds Rate, with lag length given on the horizontal axis. The *Baseline R&R* lines correspond to peak effects of monetary policy shocks identified in Romer and Romer (2004), using an autoregressive lag of 24 months and lags of the monetary policy shocks given on the horizontal axis. See section 2.4 in the text for details.

Figure 9: Controlling for Heteroskedasticity in Estimation of Romer and Romer (2004) Shocks
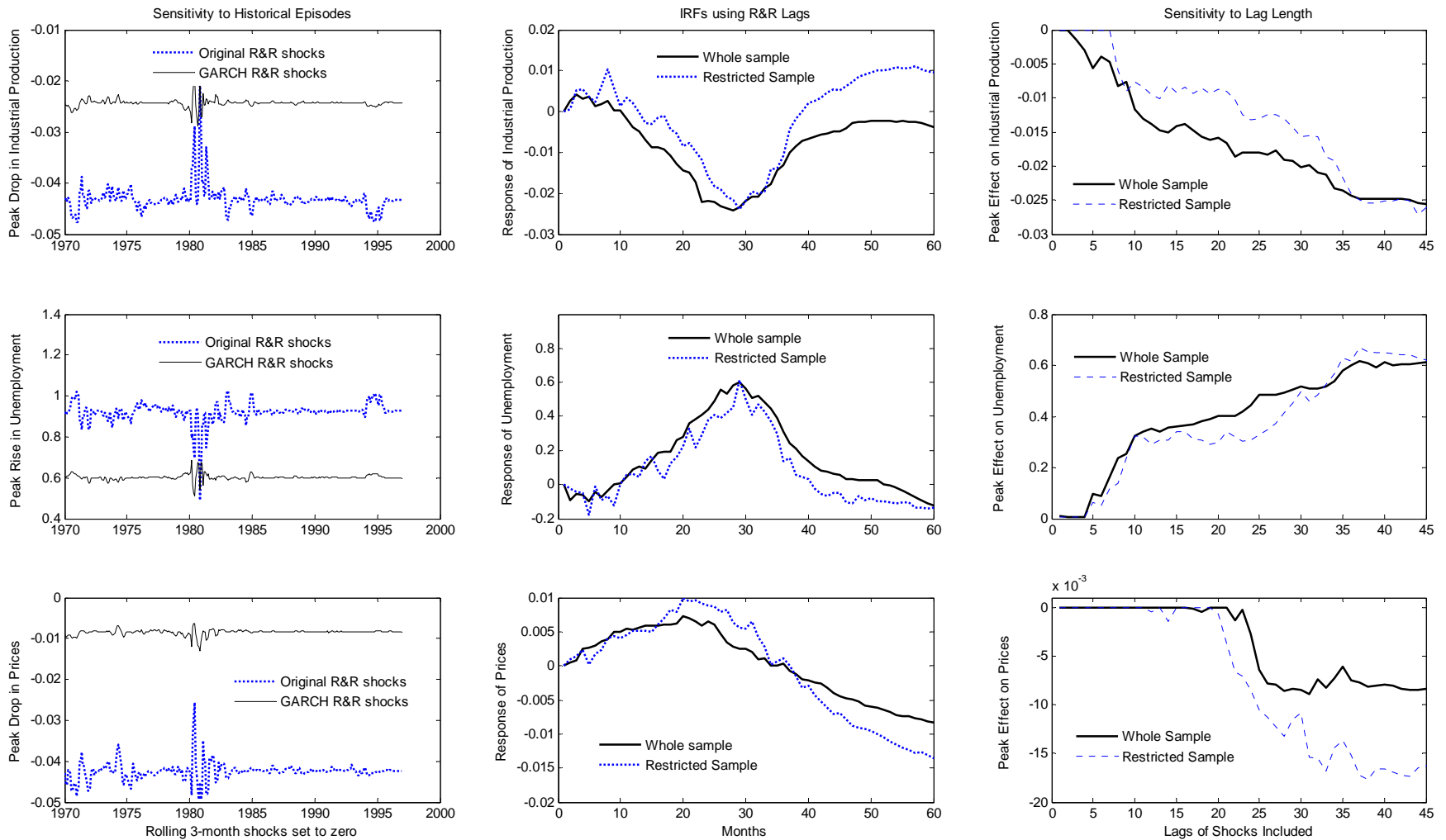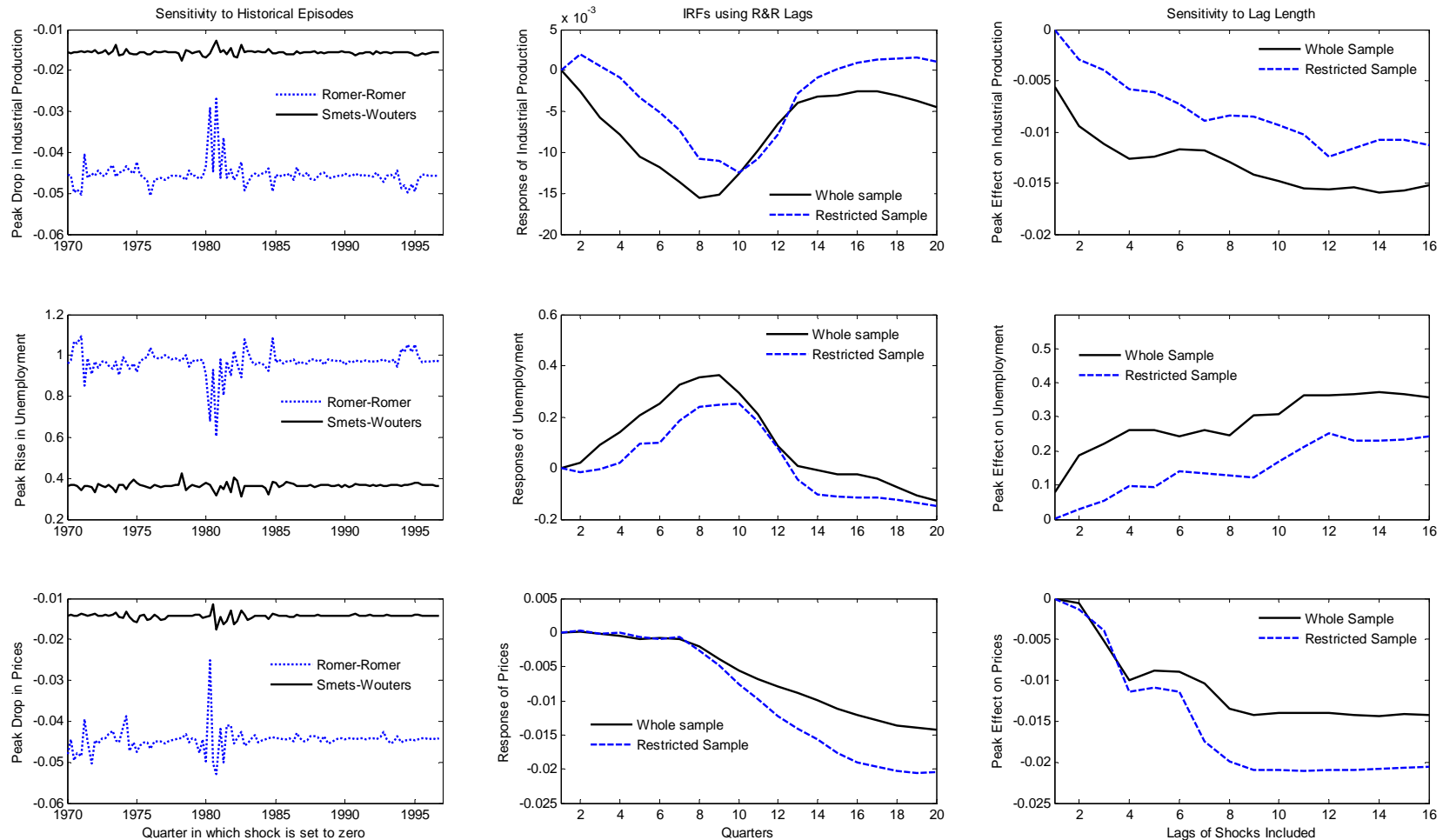

Note: The left column presents peak effects of monetary policy shocks estimated using the impulse response procedure of R&R with their lag structure, setting rolling 3-month shocks equal to zero, using the original R&R shocks as well as R&R shocks estimated by GARCH. The middle column presents impulse responses to R&R shocks estimated by GARCH. Impulse responses are estimated as in R&R, with their lag structure, over the whole sample as well as the restricted sample omitting 1979:10-1983:12. The third column presents sensitivity of peak effects of monetary policy shocks to the number of lags of shocks (R&R shocks estimated by GARCH) included in the empirical specification, over the whole sample and the restricted sample. See section 3.1 for details.

## Figure 10: Estimation of the Taylor Rule with Time-Varying Coefficients
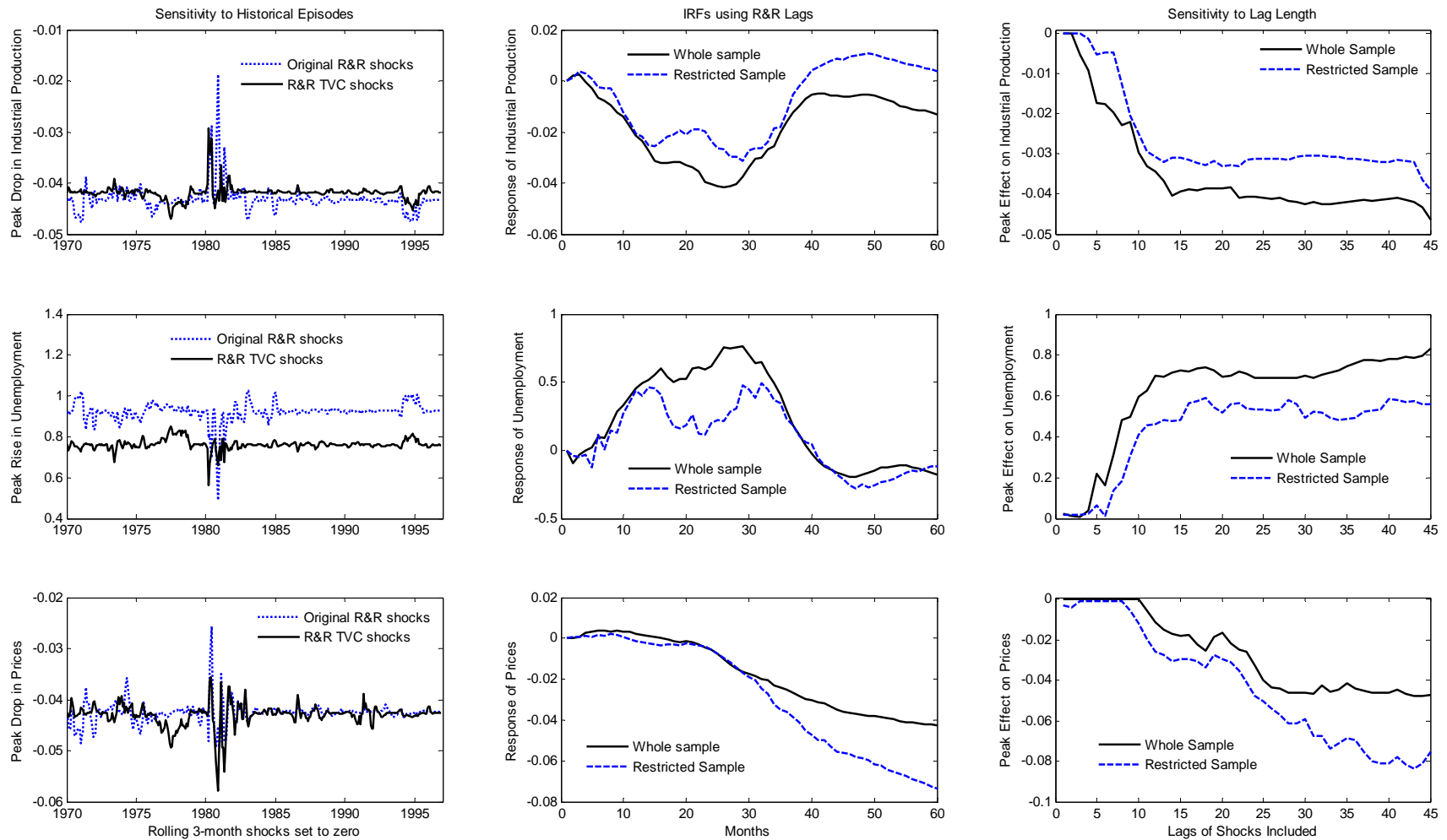
Note: The left column presents peak effects of monetary policy shocks estimated using the impulse response procedure of Romer and Romer (2004) with their lag structure, setting rolling 3-month shocks equal to zero, using the original R&R shocks as well as the shocks from estimating a restricted version of the R&R Taylor rule with time-varying coefficients (*R&R TVC shocks*). The middle column presents impulse responses to the TVC shocks. Impulse responses are estimated as in R&R, with their lag structure, over the whole sample as well as the restricted sample omitting 1979:10-1983:12. The third column presents sensitivity of peak effects of monetary policy shocks to the number of lags of the TVC shocks included in the empirical specification, over the whole sample and the restricted sample. See section 3.2 for details.

Figure 11: Results using the Smets-Wouters (2007) Measure of Monetary Policy Shocks

Note: The left column presents peak effects of monetary policy shocks estimated using the impulse response procedure of Romer and Romer (2004) with their lag structure, setting individual quarterly shocks equal to zero, using the original R&R shocks as well as the estimated monetary policy shocks from Smets and Wouters (2007). The middle column presents impulse responses to Smets-Wouters shocks. Impulse responses are estimated as in R&R, with their lag structure, over the whole sample as well as the restricted sample omitting 1979:4-1983:4. The third column presents sensitivity of peak effects of monetary policy shocks to the number of (quarterly) lags of the Smets-Wouters shocks included in the empirical specification, over the whole sample and the restricted sample. See section 3.3 for details.

Figure 12: The Contribution of Monetary Policy Shocks to Business Cycle Fluctuations
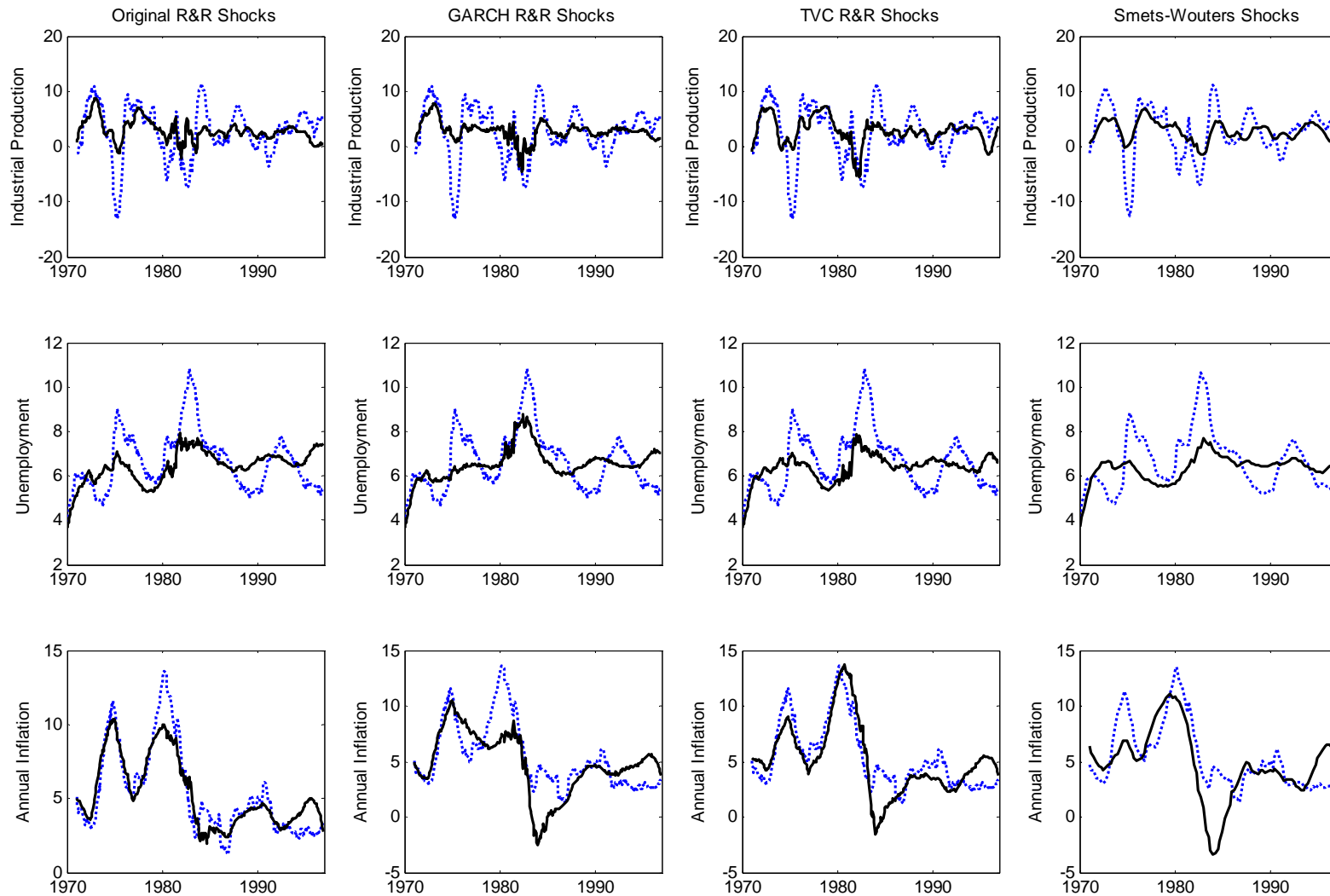
Note: The dashed blue lines are actual macroeconomic variables, while the solid black lines are the predicted time paths of each macroeconomic variable driven exclusively by identified monetary policy shocks. The first column uses the *Original R&R Shocks,* the second column uses the R&R shocks estimated using GARCH, the third column uses the shocks allowing for time-varying coefficients in the Taylor rule, while the fourth column uses Smets-Wouters (2007) shocks. Counterfactuals are based on estimating equation (1) in the text excluding the 1979:10-1983:12 period using 24 autoregressive lags and 24 lags of the shocks for industrial production and unemployment, and the model-averaging procedure for inflation. See section 3.4 in the text for details.

Table 1: Predictability of Romer and Romer (2004) Shocks

	1970:1-1996:12	1970:1-1979:9 1982:10-1996:12
Change in Industrial Production	2.28**	0.74
Monthly Inflation	2.45**	1.09
Unemployment Rate	1.62	0.96

Note: The table reports F-statistics for the null that all of the β's are equal to zero when estimating equation (2) in section 2.3 for different right-hand side variables. In each case, regressors include contemporaneous values and six monthly lags. ** denotes statistical significance at the 5% level.

Table 2: Lag Selection and Estimated Peak Effects of Monetary Policy Shocks

	Lag Length Selection			Estimated Peak Effects		
	Baseline	AIC	Averaging	Baseline	AIC	Averaging
Panel A: Whole Sample						
Industrial Production						
Standard VAR	12	3	3.1	-0.007	-0.009	-0.009
Hybrid VAR with R&R Shocks	12	3	3.0	-0.016	-0.003	-0.003
R&R Single Equation	24 36	12 12	13.5 16.9	-0.043	-0.034	-0.034
Unemployment						
Standard VAR	12	3	3.1	0.162	0.196	0.197
Hybrid VAR with R&R Shocks	12	3	3.0	0.398	0.148	0.148
R&R Single Equation	24 36	12 12	14.8 18.9	0.926	0.607	0.622
Prices						
Standard VAR	12	3	3.1	-0.001	-0.003	-0.003
Hybrid VAR with R&R Shocks	12	3	3.0	-0.018	-0.014	-0.014
R&R Single Equation	24 48	12 39	14.0 39.0	-0.042	-0.033	-0.034
Panel B: Restricted Sample						
Industrial Production						
Standard VAR	12	3	3.2	-0.022	-0.018	-0.018
Hybrid VAR with R&R Shocks	12	36	24.8	-0.022	-0.028	-0.026
R&R Single Equation	24 36	12 9	14.4 17.9	-0.028	-0.026	-0.028
Unemployment						
Standard VAR	12	3	3.2	0.398	0.191	0.193
Hybrid VAR with R&R Shocks	12	36	24.8	0.511	0.561	0.569
R&R Single Equation	24 36	24 39	23.2 40.0	0.657	0.704	0.679
Prices						
Standard VAR	12	3	3.2	-0.011	-0.010	-0.010
Hybrid VAR with R&R Shocks	12	36	24.8	-0.021	-0.035	-0.029
R&R Single Equation	24 48	12 39	14.1 38.7	-0.044	-0.042	-0.041

Note: The table presents optimal lag length selections and associated estimated peak effects for alternative measures of monetary policy shocks. *Standard VAR* corresponds to the baseline VAR approach described in section 1, *Hybrid VAR with R&R Shocks* corresponds to the baseline VAR using the cumulative measure of R&R shocks in place of the interest rate, and *R&R Single Equation* corresponds to the estimation procedure of Romer and Romer (2004). The restricted sample consists of 1970:1-1979:9 and 1984:1-1996:12. *Averaging* corresponds to the model-averaging procedure described in section 2.4. For the *R&R Single Equation* approach, the first lag corresponds to autoregressive lags and the second to lags of the shock. Lags in the Averaging column are weighted averages where weights come from the Averaging procedure. See section 2.4 for details.

Appendix 1: Description of the Model-Averaging Procedure

This section describes the model-averaging procedure used in the paper. For simplicity, I describe the algorithm used for the single-equation method, but the same procedure applies to the VAR framework. Assume that the DGP is given by

$$y_t = c + \sum_{i=1}^{J_1^*} \beta_i y_{t-1} + \sum_{j=1}^{J_2^*} \gamma_j \varepsilon_{t-j}^{mp} + v_t \qquad (A1)$$

where $J^* = [J_1^* \; J_2^*]$ is the true lag structure of the DGP, which is unobservable to the econometrician. We are interested in estimating some moment of the data $M(J^*)$ which can be recovered from equation (A1). These moments could include the impulse response of y to ε^{mp} shocks, counterfactual paths of y, etc…

Step 1: Let J denote the set of lag lengths considered by the econometrician. Use an information criterion to identify optimal lag length selection from J for equation (A1) given the data, denoted by J^{IC}.

Step 2: For each $J_i = [J_1^i \; J_2^i]$ in the set J,
1. Estimate equation (A1) with J_i lags and associated moments $\widehat{M}(J_i)$.
2. Given the estimated parameters of (A1) conditional on J_i, generate N artificial time series.
3. For each n of the N artificial time series, use the information criterion to identify the optimal lag length selection $J_{n,i}^{IC}$.
4. The bootstrapped estimate of the probability of observing J^{IC} when J_i is the true lag specification is

$$\hat{p}(J^{IC}|J_i) = \left(\frac{1}{N}\right) \sum_{n=1}^{N} I(J_{n,i}^{IC} = J^{IC})$$

where I is the indicator variable equal to one when the argument is true.

Step 3: Construct the model-averaging estimate of the moments of the data $\widehat{M}(J)$ as

$$\widehat{M}(J) = \sum_{J_i \in J} \left(\frac{\hat{p}(J^{IC}|J_i)}{\sum_{J_i \in J} \hat{p}(J^{IC}|J_i)}\right) \widehat{M}(J_i)$$

which follows Bayesian averaging with equal priors.

Estimation of Model-Averaging Procedure in Section 2.4
When applied to the single-equation approach, I consider $J_1 = [12 \; 24]$ and J_2 goes from 3 months to 45 months in 3 month intervals. For the VAR's, J goes from 3 months to 36 months in 3 month intervals. In each case, $N = 1,000$. The moments used are the impulse responses of each macroeconomic variable to monetary policy shocks. For the inflation counterfactuals in Figure 12, the moments are the historical contribution of monetary policy shocks to annual inflation.

Appendix Figure 1: Monte Carlo Simulation of Lag Selection by Information Criteria
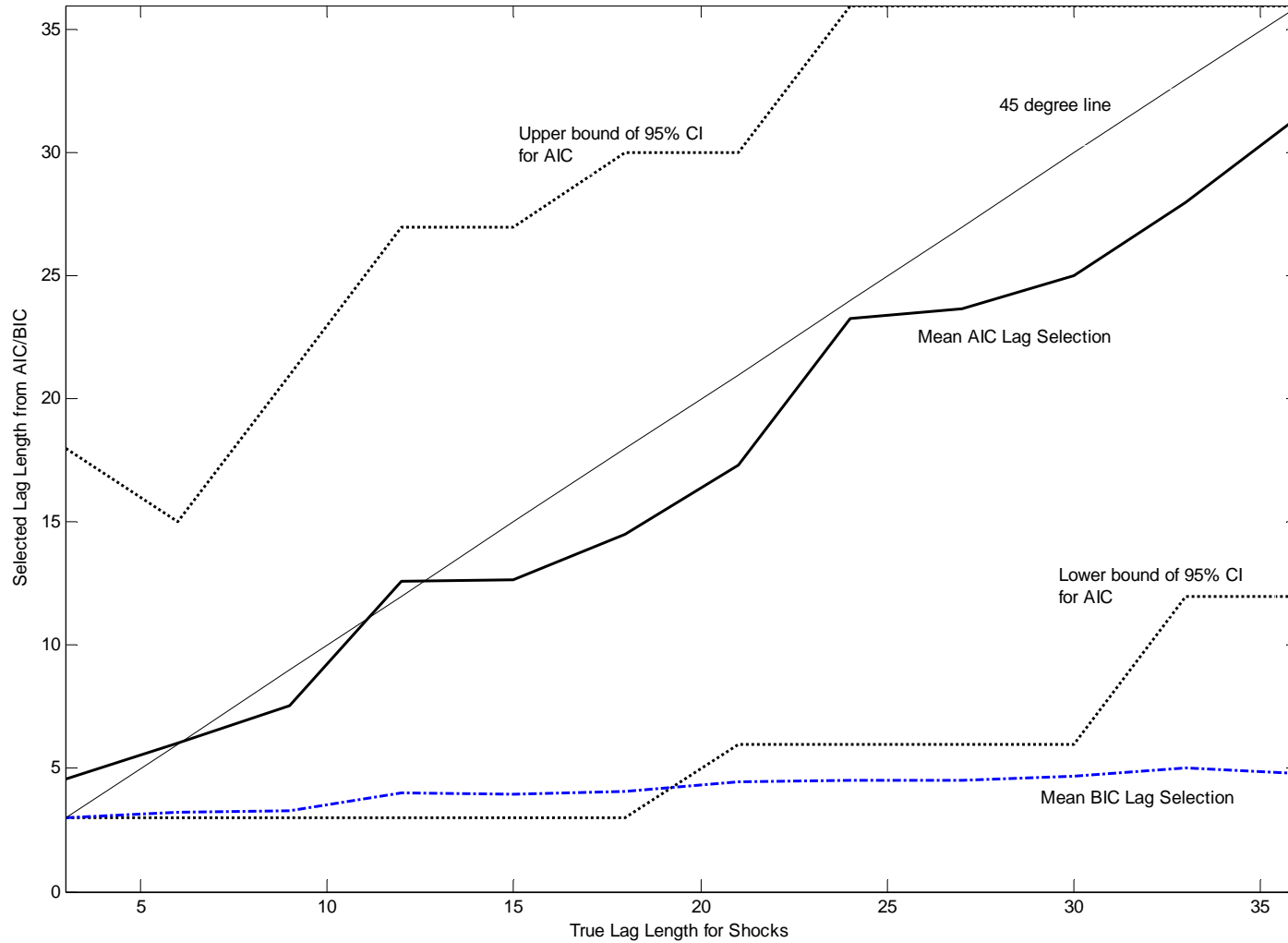

Note: The figure plots the results of a Monte Carlo simulation of equation (1) in the text for industrial production. The true autoregressive lag length is set to 12 months, while the number of lags of monetary policy shocks in the DGP varies, as indicated on the horizontal axis. For each value of the lag length for monetary policy shocks, 1,000 times artificial time series are generated. For each time series, the AIC and BIC are applied to select optimal lag lengths. The mean lag selection from each approach for each true lag length for shocks is plotted in the graph. In addition, the figure plots the 95% interval of lag selections from the AIC for each true lag length.

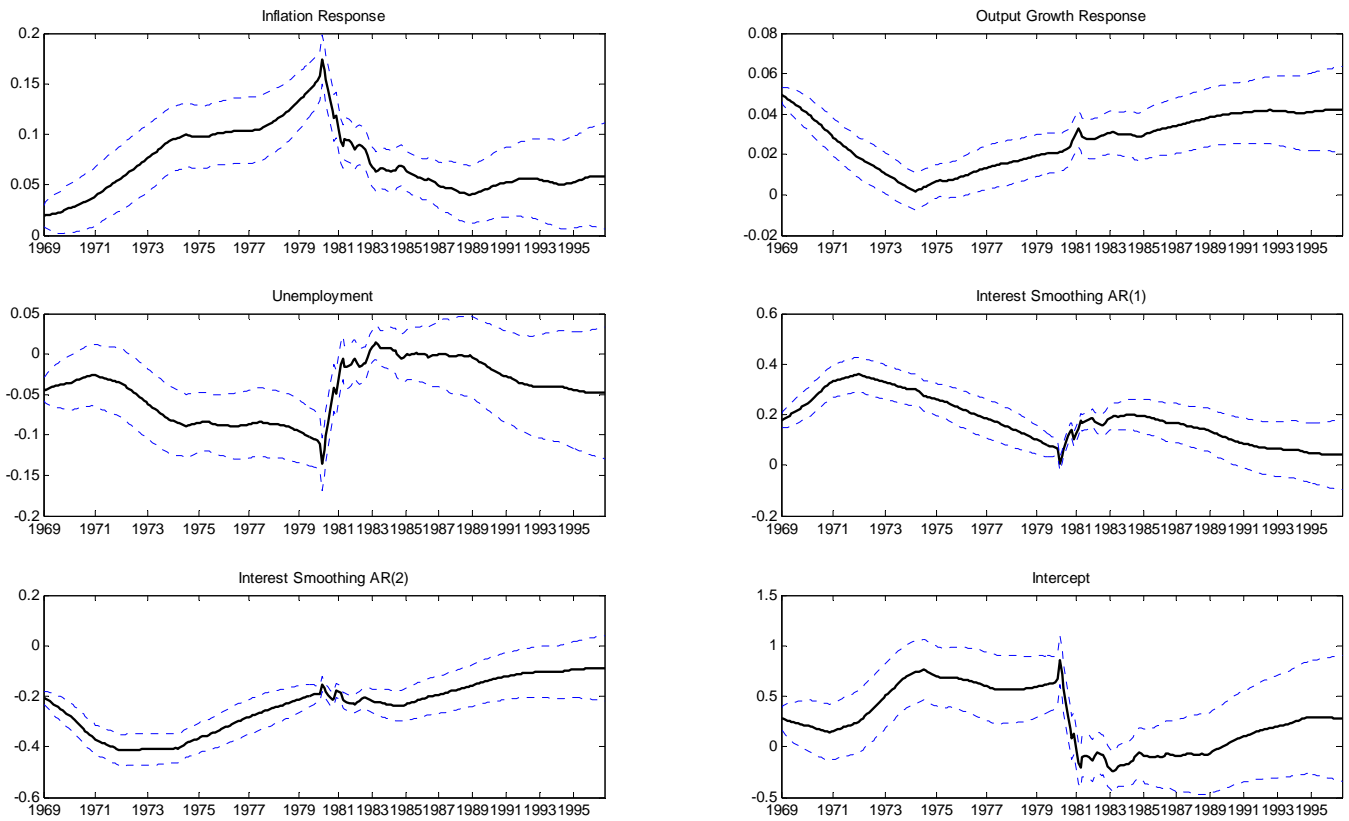**Appendix Figure 2: Original vs GARCH Romer and Romer Shocks**
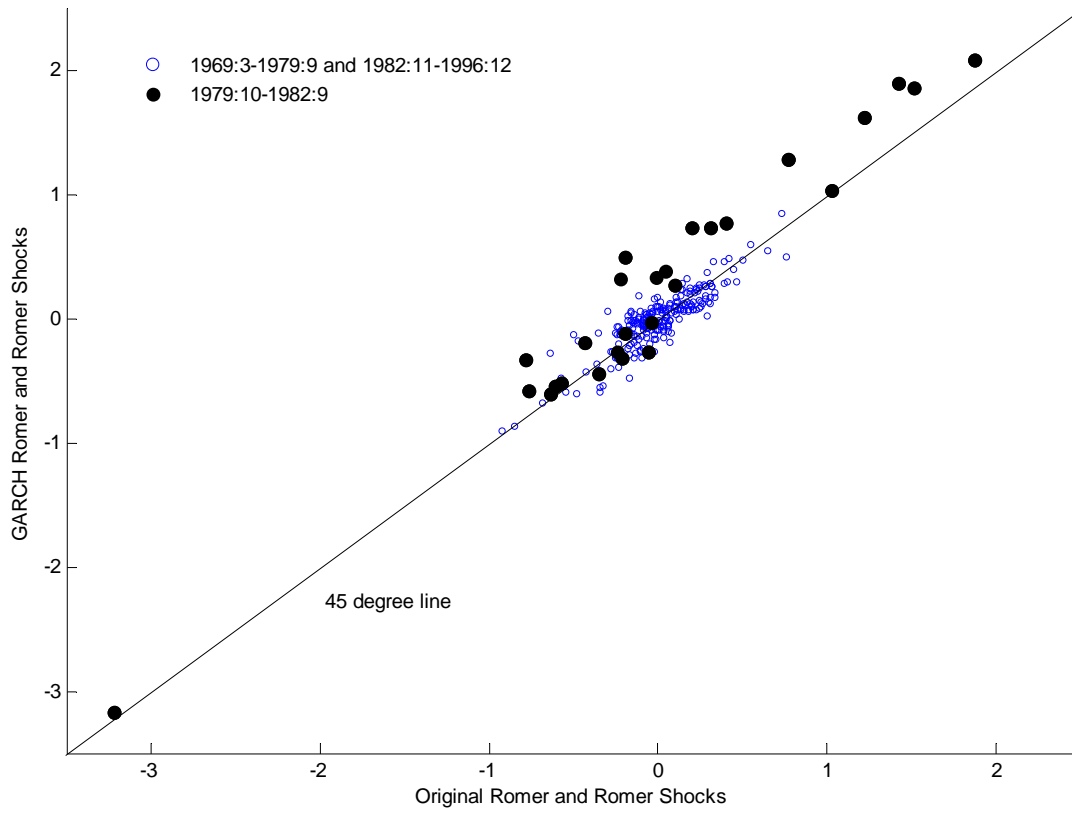
Note: The figure presents a scatter plot of the original Romer and Romer (2004) monetary policy shocks versus the residuals from estimating the same Taylor rule by GARCH(1,1). See section 3.1 for details.

Appendix Figure 3: Cumulative Shocks from Romer and Romer (2004) and GARCH
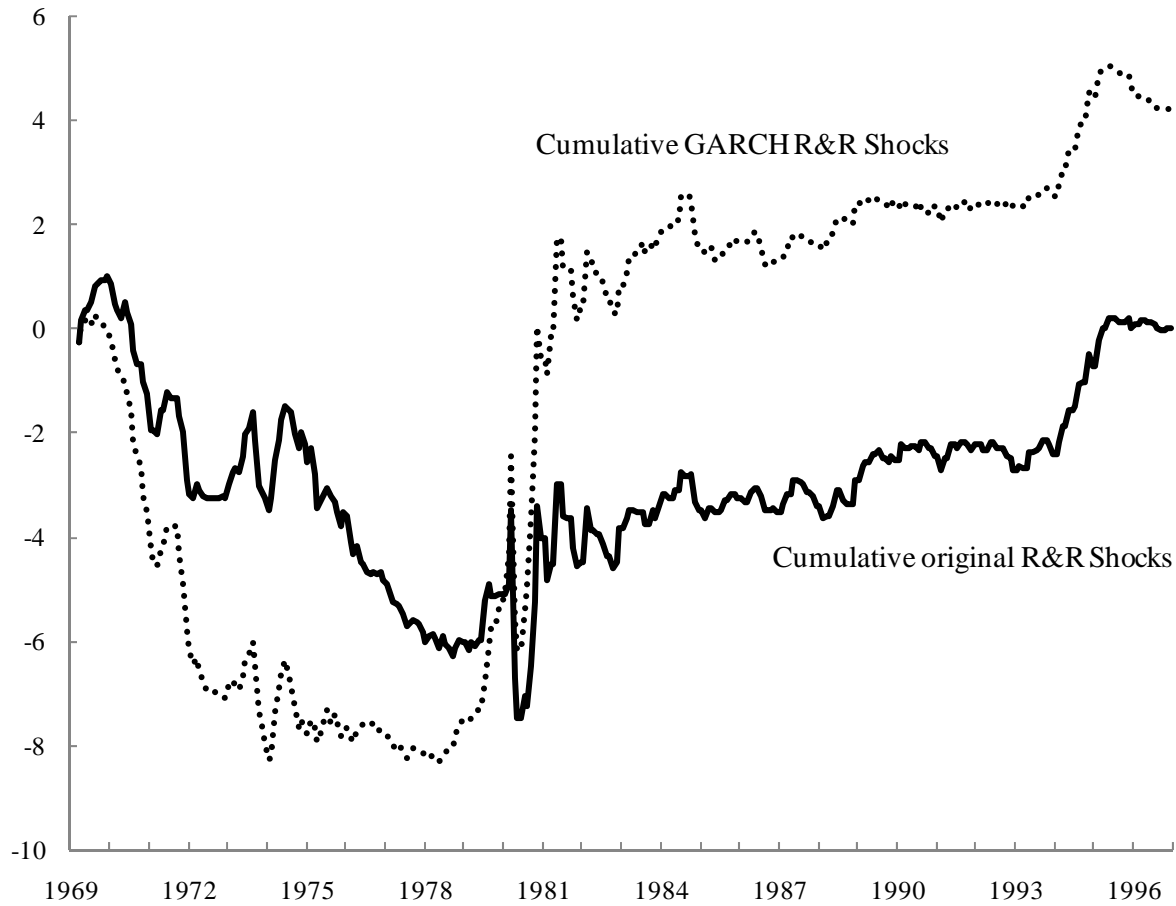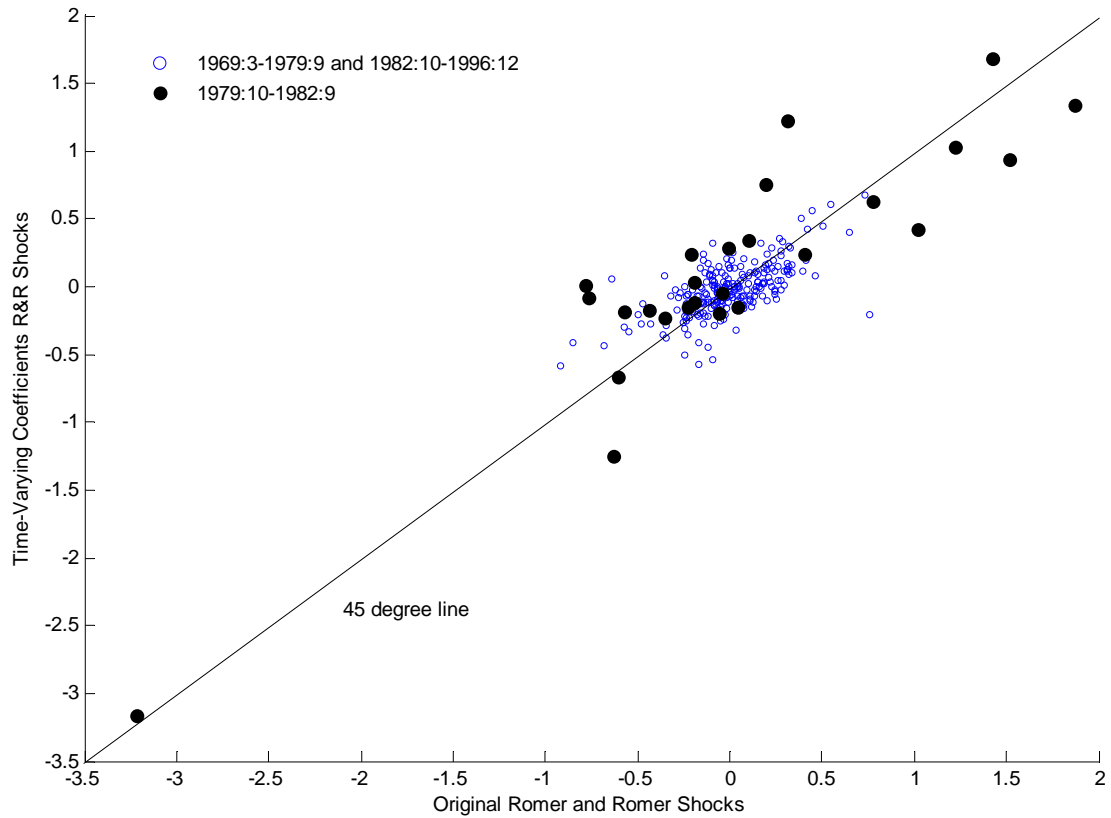

Note: The figure presents the cumulative sum of Romer and Romer (2004) monetary policy shocks and the cumulative sum of the residuals from estimating the same Taylor rule as in Romer and Romer (2004) by GARCH(1,1). See section 3.1 for details.

Appendix Figure 4: Estimated Coefficients from Time-Varying Parameters Approach


Note: The figure plots the estimated time-varying (smoothed) coefficients of the Taylor rule in section 3.2. One standard deviation confidence intervals are denoted by the dashed lines.

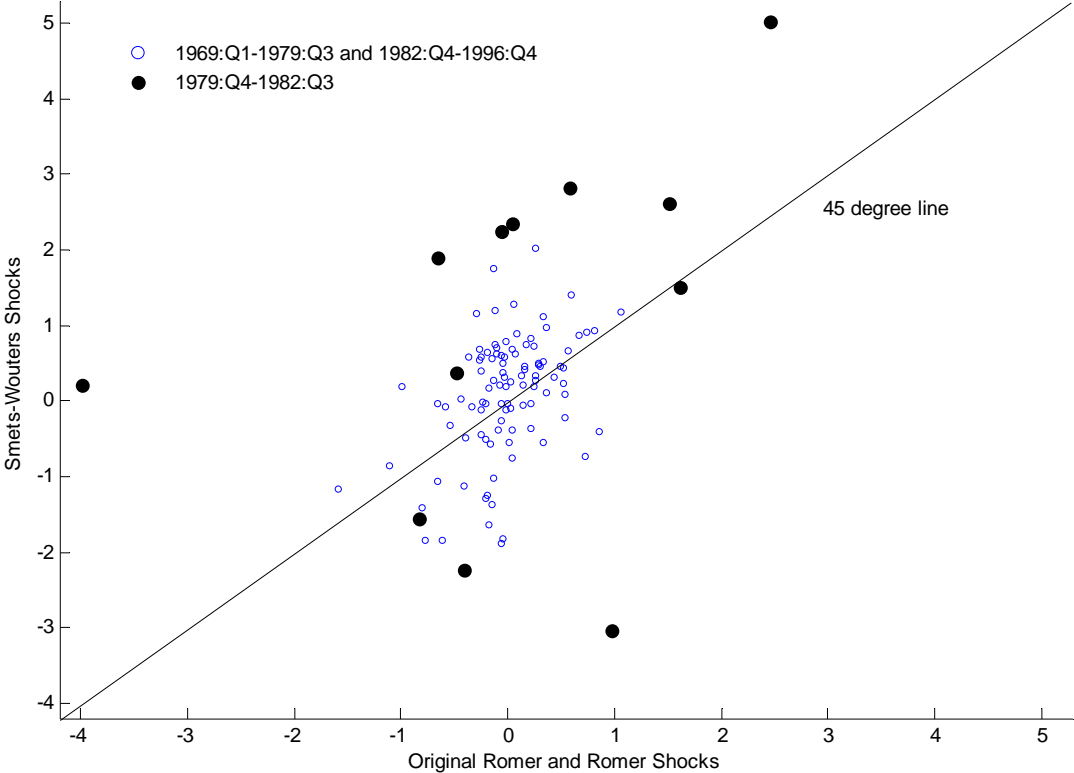**Appendix Figure 5: Original Romer and Romer Shocks vs TVC Shocks**

Note: The figure presents a scatter plot of the original Romer and Romer (2004) monetary policy shocks versus the shocks estimated from the restricted Romer and Romer Taylor rule with time-varying coefficients. See section 3.2 for details.

Appendix Figure 6: Cumulative Shocks from Original and TVC Romer-Romer Taylor Rule
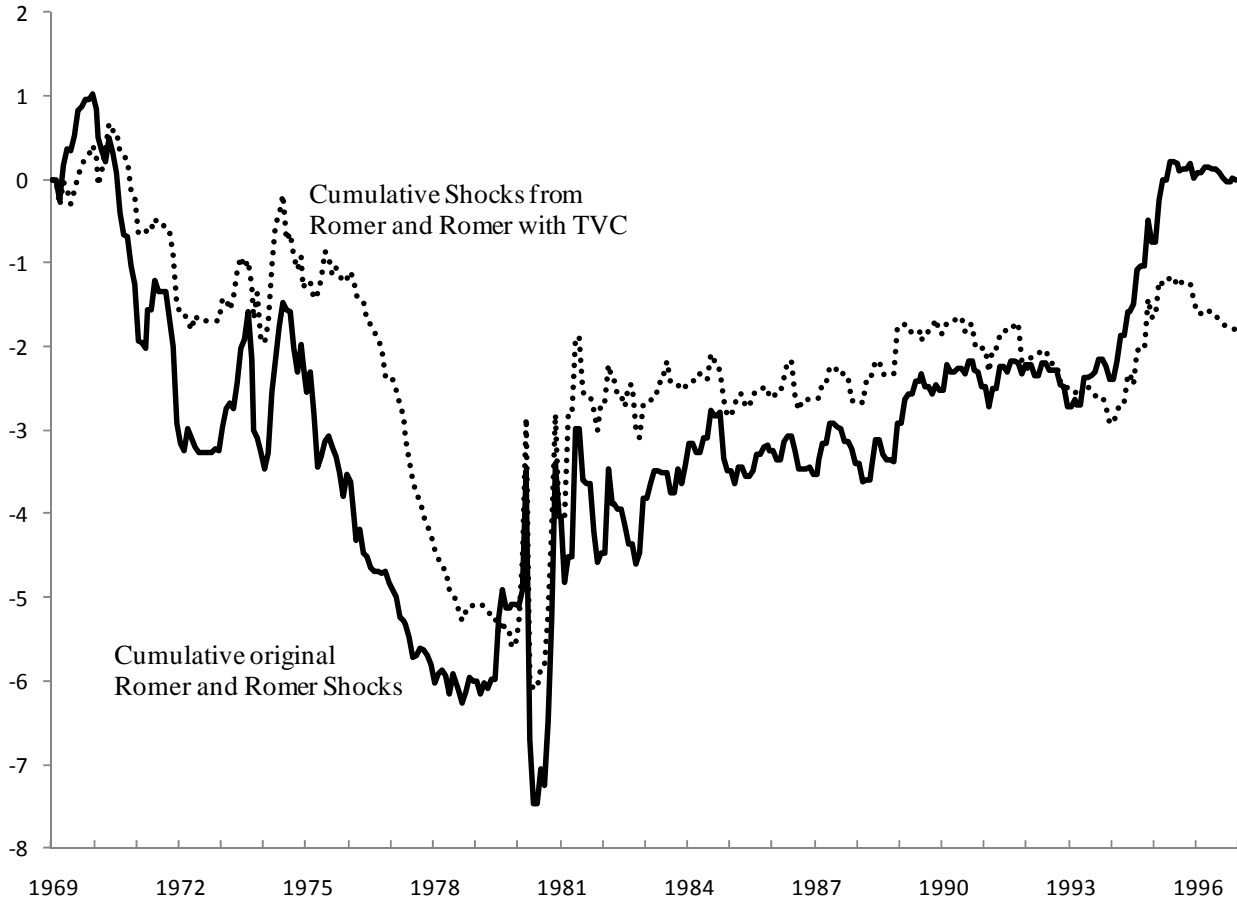
Note: The figure presents the cumulative sum of the original Romer and Romer (2004) monetary policy shocks and the cumulative sum of the shocks from the restricted Romer and Romer Taylor rule with time-varying coefficients. See section 3.2 for details.

Appendix Figure 7: Original Romer and Romer Shocks vs Smets-Wouters Shocks

Note: The figure presents a scatter plot of the original Romer and Romer (2004) monetary policy shocks versus Smets-Wouters (2007) shocks from an estimated DSGE model. See section 3.3 for details.

Appendix Figure 8: Cumulative Shocks from Romer-Romer and Smets-Wouters
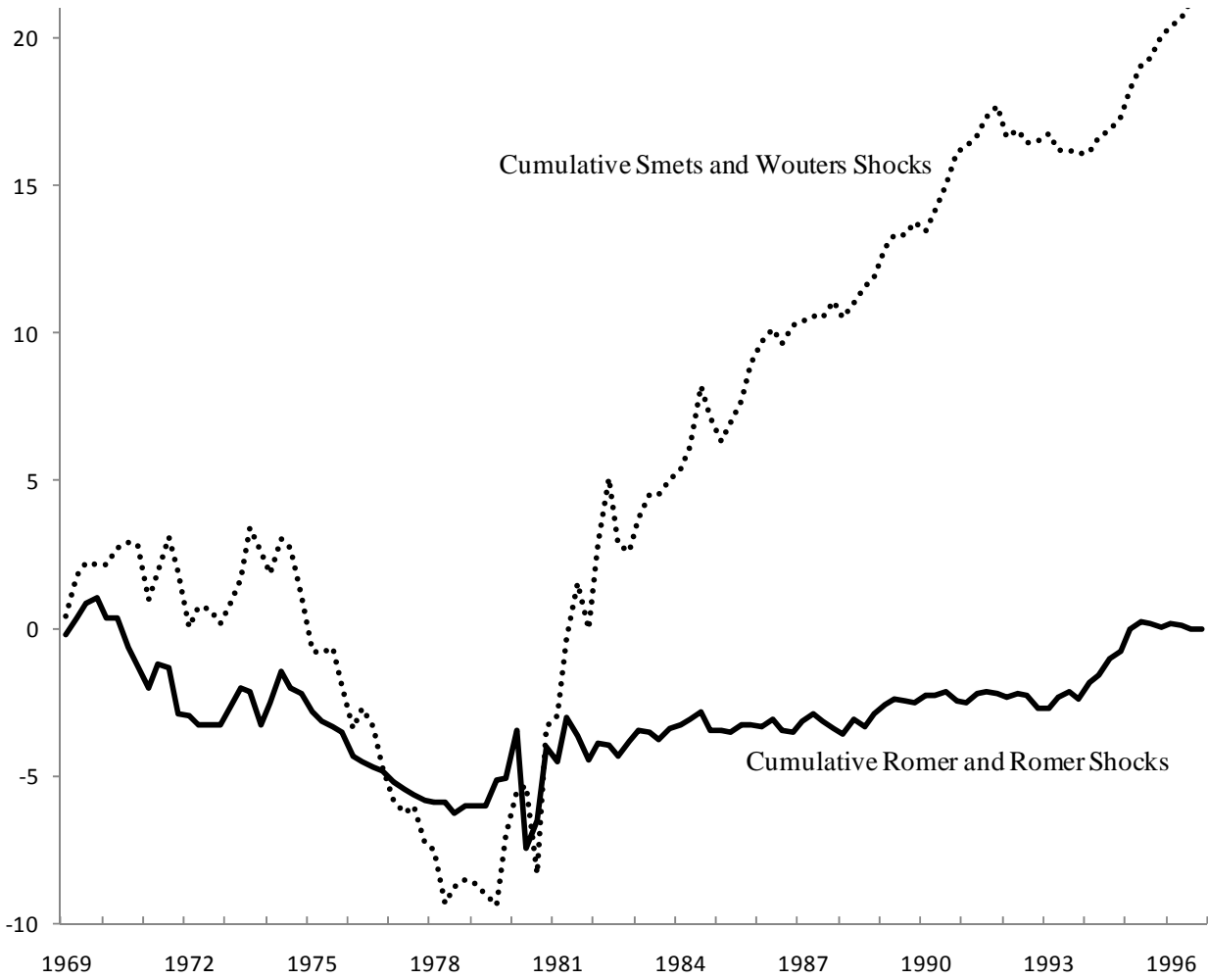
Note: The figure presents a scatter plot of the original Romer and Romer (2004) monetary policy shocks versus Smets and Wouters (2007) shocks from an estimated DSGE model. See section 3.3 for details.

Appendix Table 1: Monte Carlo Simulations of Information Criteria and Model-Averaging

	Imposed Lag Length			Information Criteria Lag Selections		Model-Averaging
	Correct Lags (12)	Insufficient Lags (6)	Excessive Lags (24)	AIC	BIC	AIC-based
Mean Lag Selection						
Autoregressive Lags	12	12	12	10.5	6.1	14.0
Monetary Policy Shock Lags	12	6	24	11.5	6.4	14.5
Fraction of Correct Lag Selections	na	na	na	55.6%	0.3%	na
Peak Effect of MP Shocks						
Mean % Error of Estimated Peak Effect	2.0%	-57.7%	8.7%	-7.0%	-56.9%	-6.1%
Standard Deviation of Estimated Peak Effects	0.0115	0.0085	0.0127	0.0134	0.0099	0.0122
MSE of Impulse Responses						
Mean	0.0028	0.0108	0.0038	0.0040	0.0106	0.0034
Ratio of mean MSE to when using correct lags	1.00	3.86	1.36	1.43	3.79	1.21

Note: The table presents results from Monte Carlo simulations of the ability of standard information criteria and a model-averaging approach to recover the true peak effects of monetary policy shocks. The Monte Carlo simulations are done as follows. First, estimate equation (1) in the text for industrial production and the original R&R shocks using 12 autoregressive lags and 12 lags of monetary policy shocks. This is assumed to be the true DGP. Second, generate 1,000 artificial time series from this specification, by feeding in random draws from the estimated residuals of (1) and the concurrent monetary policy shocks. Each time series has a burn-in sample of 100 periods then an additional 320 observations are generated to be used in subsequent steps. Third, for each simulated time series, apply the AIC and BIC criterion to select the optimal lag length from the set of $I=\{6,12,24\}$ and $J=\{6,12,24\}$. Fourth, apply the model-averaging procedure described in section 2.4 to calculate conditional probabilities to assign to each model. The first two rows of the table present the average lag selection of the AIC and BIC, as well as the mean of the weighted average of lags from the model-averaging procedure. The third row presents the frequency at which the AIC and BIC select the true lag specification. Fifth, for each simulated time series, I compute the estimated peak effect of monetary policy shocks on industrial production using a) the correct lag specification ($I=12$, $J=12$), b) using fewer lags of the shocks ($I=12$, $J=6$), c) using more lags of the shocks ($I=12$, $J=24$), d) the AIC-selected lag specification, e) the BIC-selected lag specification, and f) the model-averaging approach, which takes a weighted average of the impulse responses across possible lag specifications where the weights are determined by the estimated conditional probabilities assigned to each model. The fourth row of the table then displays the mean percentage error (relative to the true peak effect assumed) associated with each approach, while the fifth row shows the standard deviation of estimated peak effects across simulations. Sixth, for each simulation, I compute the Mean Squared Error (MSE) of each impulse response relative to the assumed true impulse response over 36 months. The last two rows display the average MSE across simulations for each approach, as well as the ratio of these means to that achieved when using the correct lag specification.

Appendix Table 2: Estimated Coefficients of the Taylor Rule in Romer and Romer (2004)

	Original R&R		Restricted R&R		GARCH R&R	
	coef	se	coef	se	coef	se
Constant	0.17	(0.12)	0.27	(0.16)	0.32	(0.09)
Initial Level of intended funds rate	-0.02	(0.01)	-0.02	(0.01)	-0.01	(0.01)
Forecasted inflation						
Quarters Ahead:						
-1	0.02	(0.03)	0.02	(0.04)	0.02	(0.01)
0	-0.04	(0.03)	-0.06	(0.04)	0.00	(0.02)
1	0.01	(0.07)	0.04	(0.08)	-0.01	(0.02)
2	0.05	(0.06)	0.03	(0.09)	-0.02	(0.03)
Change in forecasted inflation since previous meeting						
Quarters Ahead:						
-1	0.06	(0.05)	0.04	(0.05)	0.03	(0.03)
0	0.00	(0.04)	0.01	(0.05)	0.00	(0.03)
1	0.03	(0.07)	0.00	(0.10)	0.04	(0.04)
2	-0.06	(0.07)	-0.03	(0.09)	0.06	(0.05)
Forecasted output growth						
Quarters Ahead:						
-1	0.01	(0.01)	0.00	(0.01)	0.00	(0.01)
0	0.00	(0.01)	0.00	(0.02)	0.02	(0.01)
1	0.01	(0.02)	0.01	(0.03)	0.02	(0.02)
2	0.02	(0.03)	0.04	(0.04)	0.02	(0.02)
Change in forecasted output growth since previous meeting						
Quarters Ahead:						
-1	0.05	(0.03)	0.06	(0.04)	0.02	(0.02)
0	0.15	(0.04)	0.15	(0.04)	0.06	(0.02)
1	0.02	(0.04)	0.03	(0.05)	0.03	(0.03)
2	0.02	(0.04)	-0.01	(0.06)	-0.02	(0.04)
Forecasted UE rate, current quarter	-0.05	(0.02)	-0.07	(0.02)	-0.06	(0.01)
Conditional Variance Equation (for GARCH)						
Constant					0.00	(0.00)
Lag of Squared Residual					0.48	(0.12)
Lag of Forecast Variance					0.57	(0.08)
Sample	1969:3-1996:12		1972:11-1996:12		1972:11-1996:12	
N	263		221		221	

Note: The table presents estimates of the Taylor rule in Romer and Romer (2004). The first specification replicates the OLS estimation of Romer and Romer, the second estimates the same Taylor rule by OLS starting in 1972:11, and the third estimates the same Taylor rule by GARCH(1,1) over the restricted sample. See section 3.1 for details.