Center for European, Governance and Economic Development Research
cege

Discussion Papers

Number 295 – November 2016

MONETARY POLICY SHOCKS, SET-IDENTIFYING RESTRICTIONS, AND ASSET PRICES: A BRENCHMARKING APPROACH FOR ANALYZING SET-IDENTIFIED MODELS

Gábor B. Uhrin, Helmut Herwartz

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

ISSN: 1439-2305

Monetary policy shocks, set-identifying restrictions, and asset prices: A benchmarking approach for analyzing set-identified models*

Gábor B. Uhrin† Helmut Herwartz‡

First version: February 15, 2016.

This version: November 17, 2016.

Abstract

A central question for monetary policy is how asset prices respond to a monetary policy shock. We provide evidence on this issue by augmenting a monetary SVAR for US data with an asset price index, using set-identifying structural restrictions. The impulse responses show a positive asset price response to a contractionary monetary policy shock. The resulting monetary policy shocks correlate weakly with the Romer and Romer (2004) (RR) shocks, which matters greatly when analyzing impulse responses. Considering only models with shocks highly correlated with the RR series uncovers a negative, but near-zero response of asset prices.

JEL Classification: C32, E44, E52

Keywords: monetary policy shocks, asset prices, sign restrictions, zero restrictions, set identification, structural VAR models

*Financial support from the Deutsche Forschungsgemeinschaft (HE 2188/8-1) is gratefully acknowledged. We thank conference and seminar participants at the European Meeting of the Econometric Society 2016, IE-HAS Summer Workshop Budapest 2016, and DIW Berlin for comments and discussions, especially Kerstin Bernoth, Gábor Kőrösi, Helmut Lütkepohl, and Martin Wagner. We thank Emanuel Mönch for providing access to replication codes. The first version of this paper has been circulated under the title: "Monetary policy shocks, sign restrictions, and asset prices: A novel approach for analyzing sign restricted models".

†Corresponding author: Department of Economics, University of Göttingen, Humboldtallee 3, D-37073 Göttingen, Germany, e-mail: guhrin@uni-goettingen.de, tel: +49/551/3927175.

‡Department of Economics, University of Göttingen.

1 Introduction

The financial crisis of 2008–2009 has stirred up the debate on the conduct of monetary policy all around the world. One of the questions that came into the focus of the discussion is the extent to which monetary policy should react to developments in asset markets. Should central banks "lean against the wind" and try to mitigate turbulences in asset markets through raising interest rates, or should they rather concentrate solely on stabilizing the output gap and the inflation?[1] Arguing for either of these positions raises the need to quantify the (contemporaneous) effects of monetary policy actions on asset prices.

Starting with Sims (1980), such empirical questions have often been investigated by means of (structural) vector autoregressive (VAR) models. The crucial problem of identifying exogenous, unanticipated monetary policy shocks has been addressed in several studies that aimed to quantify the effects of monetary policy on, for instance, real output. Most of the classical procedures developed in these studies have been applied in a monetary policy – asset price context. As a particular exception, the agnostic sign-restriction approach exemplified by Uhlig (2005) has not been employed yet to explore the linkage between monetary policy shocks and asset prices.

Parallel to the SVAR literature, alternative approaches to identify monetary policy shocks have also been proposed. A major contribution has been put forth by Romer and Romer (2004), henceforth RR, who combined narrative evidence with statistical methods to construct a monetary policy shock series free of endogeneity and anticipation effects.

In the present paper we make several, related, contributions. First, we augment the VAR specification of Uhlig (2005) with the S&P 500 Composite Index, and estimate the model on monthly US data from 1959 January to 2007 December. We use two set identifying restrictions to identify monetary policy shocks and examine the effects of these shocks on asset prices. The first restrictions are the sign restrictions of Uhlig (2005) (Scheme I), and the second set of restrictions are the zero and sign restrictions on the structural matrix A_0 put forth recently by Arias, Caldara, and Rubio-Ramírez (2015) (Scheme II). According to our results, the SVAR impulse responses point towards a mildly positive asset price response to an increase in the monetary policy instrument. This result is puzzling in light of earlier literature. Second, we argue that the resulting identified monetary policy shocks correlate only weakly with the monetary policy shock series of RR. We show that this finding matters greatly when analyzing (structural) impulse responses. In particular, we make the following observations: *i.*) the

[1] A concise summary of these debates can be found, e.g., in Assenmacher-Wesche and Gerlach (2010).

majority of admissible models yield impulse responses that vary widely in their shapes and impact magnitudes; *ii.*) this ambiguity affects those variables most whose responses are left agnostic by the identification scheme; *iii.*) models that are highly correlated with the RR shocks yield clearly shaped and less ambiguous impulse responses. Thus, third, we propose to restrict attention to those specifications that yield monetary policy shocks highly correlated with the RR series. We show that impulse response analysis of these models leads to more robust and reliable conclusions. Ultimately, we find evidence of: 1.) asset prices responding mildly negatively (in Scheme I), or ambiguously (in Scheme II) to a positive monetary policy shock, 2.) a mildly positive output response to what is understood to be a "contractionary" monetary policy shock. The former findings are contrary to our first results, but in line with conclusions of earlier studies. The latter finding is contrary to the baseline results obtained recently by Arias, Caldara, and Rubio-Ramírez (2015). Thus, we also conclude that comparing structurally (set-) identified shocks to a benchmark series can uncover by default hidden, but relevant and robust empirical conclusions. As a result, our methodological contribution complements the concerns of Kilian and Murphy (2012) regarding the interpretation of results from set-identified SVARs, and can be a useful empirical strategy when the identified set is not sufficiently narrow for sharp empirical conclusions. In fact, the benchmarking approach that we put forth can be considered as a step towards a frequentist parallel of the most likely models of Inoue and Kilian (2013).

The paper proceeds as follows: In Section 2 we provide an overview of existing results in identifying monetary policy shocks and their effects on asset prices. In Section 3 we detail the econometric model and the structural identifying assumptions. In Section 4 we present our baseline results,. In Section 5 we analyze the identified monetary policy shock series and compare them with the Romer and Romer (2004) series. In Section 6 we re-investigate our baseline results concentrating only on a certain subset of admissible models. Section 7 provides a discussion, some further results and robustness checks. Finally, Section 8 concludes.

2 Monetary policy shocks and asset prices

While the crucial empirical problem in characterizing effects of monetary policy shocks is identifying exogenous, unanticipated changes in monetary policy, there seems to be no consensus in the literature on the identifying assumptions to use. Ramey (2016) provides a critical review of several identifying assumptions and argues that previous results based on distinct identifying assumptions cannot easily be reconciled, especially in longer, more recent

samples.

Since distinct identifying assumptions may generate distinct results, the lack of consensus also applies to the empirical question: what are the effects of monetary policy shocks on asset prices? Compared with the literature on quantifying the effects of monetary policy on real variables, the empirical literature on monetary policy and asset prices is relatively small-scale. While the literature generally concludes that asset prices react negatively to an exogenous increase in the monetary policy instrument, the magnitude, the timing and the persistence of this negative reaction varies greatly across studies.

Earlier papers that use a recursive identification scheme, including Patelis (1997), Thorbecke (1997), Neri (2004), find that an increase in the monetary policy instrument leads to a small decrease in the stock prices. Bjørnland and Leitemo (2009) criticize the use of recursive identification schemes. Applying short and long run restrictions, they find large and persistent negative effects. More recently, Lanne, Meitz, and Saikkonen (2015) assume a non-Gaussian SVAR and confirm the findings of Bjørnland and Leitemo (2009) in rejecting the recursive identification scheme, and finding a significant instantaneous negative effect that, however, dies out quickly. In contrast, utilizing changes in the heteroskedasticity structure of the error term, Rigobon and Sack (2004) and Lütkepohl and Netšunajev (2014) find smaller, but relatively persistent negative effects. In a time-varying SVAR, Galí and Gambetti (2015) find negative short run effects that quickly turn into positive after impact especially in the 1980s and 1990s. Following an event-study approach around the monetary policy decision changes, Bernanke and Kuttner (2005) uncover that a 25 basis point cut in the federal funds rate leads, on average, to a 1% increase in asset prices.

As the above list of contributions indicate, a wide variety of approaches to SVAR analysis have been applied in the monetary policy – asset prices context. Notable exceptions are the usage of sign restrictions as proposed by, e.g., Uhlig (2005), and sign and zero restrictions advocated by Arias, Caldara, and Rubio-Ramírez (2015). We aim to fill this gap in the present paper, and we argue in the next section that using these restrictions as identifying assumptions in the context of our empirical question has several advantages over other identification schemes.

3 Identifying monetary policy shocks with sign and zero restrictions

We consider the following K-dimensional structural VAR,

$$A_0 y_t = A_1 y_{t-1} + A_2 y_{t-2} + \cdots + A_p y_{t-p} + \varepsilon_t, \tag{1}$$

where $y_t \in \mathbb{R}^K$, $\varepsilon_t \sim WN(0, I_K)$, $A_0, \ldots, A_p \in \mathbb{R}^{K \times K}$, and A_0, what we call the *structural matrix*, is assumed to be non-singular. In order to define a unique lag length we assume that $A_p \neq 0$. In the above equation ε_t is the vector of *structural innovations*. The corresponding, estimable reduced form is

$$y_t = B_1 y_{t-1} + \cdots + B_p y_{t-p} + u_t, \tag{2}$$

with $B_i = A_0^{-1} A_i,\ i = 1, \ldots, p$. For u_t, the vector of *reduced form innovations* the following holds: $A_0^{-1}\varepsilon_t = u_t \sim WN(0, \Sigma_u)$. That is, the vector of structural innovations is a linear combination of the vector of reduced form innovations. Writing $B(z) = I_K - B_1 z - \cdots - B_p z^p$, we assume that the reduced form is causal, that is, $\det(B(z)) \neq 0\ \forall\, |z| \leq 1$. Then, the moving average representation of y_t exists and is given by (Brockwell and Davis, 1991, Th. 11.3.1, p. 418):

$$y_t = \sum_{j=0}^{\infty} \Phi_j u_{t-j} = \sum_{j=0}^{\infty} \Theta_j \varepsilon_{t-j}, \quad \Phi_0 = I_K, \tag{3}$$

where element (i, k) of the coefficient $\Theta_j = \Phi_j A_0^{-1}$ is interpreted as the reaction of the i-th variable on the k-th structural innovation at horizon j. We call A_0^{-1} the *structural impact matrix*, since $\Theta_0 = A_0^{-1}$.

In this paper we aim to identify only one particular structural form innovation, the monetary policy shock, ε_t^{mp}, that is an element of the vector ε_t. We use sign restrictions on the impulse responses and zero restrictions on the structural matrix A_0 as identifying assumptions. Zero and sign restrictions on the structural matrix A_0 are straightforward: $A_0^{(i,k)}$, the (i, k)-th element of A_0 is restricted to be zero, positive, or negative. Sign restrictions on the impulse responses can be formulated as follows: $\Theta_j^{(i,k)}$, the (i, k)-th element of Θ_j, is restricted to be either negative or positive for some a priori selected combinations of (i, k, j), $i, k \in \{1, \ldots, K\}$, $j \in \mathbb{N}_0$. Note, that an insufficient amount of zero restrictions on A_0, or sign restrictions in general cannot point identify the structural parameters $A_0, \ldots, A_p$.[2]

[2]For necessary and sufficient conditions for exact (point) identification see, for example, Rubio-Ramírez,

Using sign restrictions on impulse responses for several periods to identify monetary policy shocks has been first proposed by Uhlig (2005).[3] Somewhat surprisingly, we are not aware of any attempt to utilize sign restrictions in a monetary policy – asset prices context. The use of zero restrictions on the structural matrix to restrict the systematic component of monetary policy in the SVAR has been put forth recently by Arias, Caldara, and Rubio-Ramírez (2015), and we are not aware of any research employing this identification to our empirical question. While the employed sign and zero restrictions cannot, in general, point-identify a structural VAR model, or a structural shock, using set identification has two important advantages in our view.[4]

First, sign restrictions by construction avoid the problem of deciding upon the exact recursive ordering of shocks. As Bjørnland and Leitemo (2009) pointed out, it is important to allow for the possibility of the monetary policy shocks contemporaneously affecting asset prices *and vice versa* – a view supported by theoretical models of, e.g., Castelnuovo and Nisticò (2010). A simple recursive identification scheme necessarily excludes one of these possibilities. Further, this critique of the recursive identification schemes was also strengthened recently by Lanne, Meitz, and Saikkonen (2015), who assume non-Gaussian error terms, and statistically test and reject the adequacy of the recursive scheme.

Second, sign restrictions, and a small number of zero restrictions, on the other hand, are considered to be mild assumptions that are relatively easy to interpret, justify, and agree upon. If we are striving for exact identification together with allowing for non-recursivity, or contemporaneous interdependence, then we have to argue for at least one additional restriction to achieve it. If, however, one is willing to give up on exact identification, then restrictions that were used to achieve exact identification may become harder to argue for.

Thus, in the following we use sign and zero restrictions to identify the monetary policy shock and investigate the effect of this identified shock to a stock price index variable. To this end, we augment a VAR specification similar to Uhlig (2005) with the S&P500 Composite Index and use, as our first assumption, the same identifying assumption as Uhlig (2005):

Restriction SR1 A monetary policy shock's effects on the impulse responses of commodity prices, GDP deflator and non-borrowed reserves is non-positive, and on the impulse response of the federal funds rate is non-negative for the impact period and four periods

Waggoner, and Zha (2010).

[3]Similar contributions are Faust (1998) and Canova and De Nicoló (2002).

[4]We refer to a shock being set identified if there are at least two parameter points in the structural parameter space that are observationally equivalent, i.e., lead to the same reduced form parameters. This terminology is also used by Arias, Caldara, and Rubio-Ramírez (2015).

after impact.

We call this restriction *Scheme I*. Besides being intuitively reasonable, these sign restrictions are also supported by New-Keynesian DSGE models under a wide set of parameter calibrations (Carlstrom, Fuerst, and Paustian, 2009). Note, that the original formulation by Uhlig (2005) requires the impulse responses of "prices" in general to be non-positive. While this assumption may also include asset prices, we prefer to remain agnostic about the signs of effects of monetary policy on asset prices, hence we do not constrain the response of asset prices to monetary policy shocks.

Arias, Caldara, and Rubio-Ramírez (2015) argue, however, that the sign restrictions of Uhlig (2005) imply parameter estimates that are incompatible with theoretical considerations about and empirical evidence on the systematic component of the monetary policy, the Taylor rule. Since monetary policy shocks are innovations to the Taylor rule, the identification of monetary policy shocks should be coupled with identifying the corresponding systematic monetary policy equation in the SVAR. This can be achieved by means of zero restrictions on the structural matrix. Following Arias, Caldara, and Rubio-Ramírez (2015), we use the following zero and sign restrictions on A_0 to identify monetary policy shocks:

Restriction ZR The federal funds rate only reacts contemporaneously to GDP, GDP deflator, commodity prices, and asset prices.

Restriction SR2 The federal funds rate's contemporaneous reaction to GDP, and to the GDP deflator is positive.

We call these restrictions jointly *Scheme II*. These restrictions explicitly impose a Taylor-type rule on the federal funds rate equation of the SVAR consistent with empirical and theoretical evidence about the systematic component of monetary policy. In particular, restriction ZR implies that the contemporaneous reaction of the federal funds rate to non-borrowed reserves and total reserves is zero. The monetary policy shock is identified as the innovation corresponding to this correctly specified equation in the SVAR. Since we are interested in the response of the monetary policy instrument to asset prices, we allow the federal funds rate to react contemporaneously to asset prices. This leaves our identification agnostic in the asset price – monetary policy context. It is important to note that Restrictions ZR and SR2 restrict the structural matrix A_0. Thus, in contrast to Restriction SR1, the impact period impulse response coefficient A_0^{-1} is restricted only indirectly.[5]

[5]Note, that a zero restriction in A_0 in general does not imply a zero restriction in A_0^{-1}.

4 Monetary policy shocks and asset prices

First we investigate the effects of monetary policy on asset prices in a structural VAR similar to Uhlig (2005). The VAR is estimated with monthly US data from 1959:01 to 2007:12. The seven variables used in the specification are: Real GDP, GDP deflator, commodity price index, stock price index, federal funds rate, non-borrowed reserves, total reserves. Monthly series for real GDP and the GDP deflator were interpolated as in Mönch and Uhlig (2005). Real GDP was interpolated using the industrial production index, while the GDP deflator was interpolated by means of consumer and producer price indices. The commodity price index is the Commodity Research Bureau's BLS spot index obtained from Thomson Reuters' Datastream and is determined as the monthly average of daily data. Monthly observations of the S&P 500 Composite Index were obtained from the FRED MD project website maintained by Michael W. McCracken (McCracken and Ng, 2016). For the empirical analysis, the values were deflated by the GDP deflator. The remaining variables were obtained from the St. Louis FRED database under the following names: GDPC1 (real GDP), INDPRO (industrial production), GDPDEF (GDP deflator), CPIAUSL (consumer price index), PPIFGS (producer price index), FEDFUNDS (federal funds rate), TOTRESNS (total reserves), and BOGNONBR (non-borrowed reserves).[6]

To facilitate comparability, we employ the same VAR specification as Uhlig (2005): the VAR contains $p = 12$ lags and does not include a constant or deterministic trend. The federal funds rate is considered in levels. All other variables are in logarithms and multiplied by 100. We estimate the VAR by OLS. In order to simulate the set of sign and zero restricted impulse responses, we use the algorithms proposed by Rubio-Ramírez, Waggoner, and Zha (2010) and Arias, Rubio-Ramírez, and Waggoner (2014). A detailed description of these algorithms can be found in Appendix B. In short, we draw random orthogonal matrices, Q, to rotate the lower-triangular Cholesky decomposition $\widehat{A}_u^c$ of $\widehat{\Sigma}_u$, the estimated reduced form variance-covariance matrix. The rotation matrices Q are constructed in a systematic way so that the structural form parameters estimated using $\widehat{A}_0^{-1} = \widehat{A}_u^c Q$ satisfy the sign and zero restrictions (i.e., they are *admissible*). We repeat the random drawing procedure until we have 65000 admissible impulse responses. Each of these 65000 impulse responses corresponds to a distinct *admissible model*, $\widehat{A}_{0,s}^{-1} = \widehat{A}_u^c Q_s$, where $s = 1, \ldots, 65000$ is the simulation index.[7] In the plots below we also report the median target (MT) impulse responses as advocated by Fry

[6] Further details on the data and sources can be found in Appendix A.

[7] While the number 65000 is based on computational constraints, our results are robust to several dozen runs of the same procedure.

and Pagan (2011). The MT impulse response is the impulse response that is closest in terms of a standardized squared distance to the median of the set of admissible impulse responses.

Sign restrictions on impulse responses

We identify a monetary policy shock first solely via Scheme I (Restriction SR1), our baseline restriction on the impulse responses. In Figure 1 we display the impulse response of the asset price index to a one per cent increase in the federal funds rate. The sub-figure on the left contains the pointwise median, as well as the pointwise 0.3 and 0.7 quantiles of the set of admissible impulse responses. In the sub-figure on the right we report the median target impulse response joint with a 90% bootstrap confidence band.[8]

Figure 2 visualizes the set of admissible impulse responses of the rest of the variables to a one per cent positive monetary policy shock (increase in the federal funds rate). The results are very similar to those presented in Uhlig (2005, Fig. 6., p. 397), thus we do not discuss them in detail. The only difference is the slight rising trend of the GDP deflator after the impact period. Figure 3 contains the MT impulse responses for the same variables as in Figure 2.

[FIGURES 1, 2, 3 ABOUT HERE]

The conclusion of our baseline analysis is that, for a one per cent increase in the federal funds rate, the set of admissible impulse responses of asset prices have more mass above the zero line. A similar result holds also for the response of GDP. While one can argue that these conclusions are ambiguous, as there are also admissible models that yield negative asset price responses, the MT impulse responses further point towards positive responses to a positive monetary policy shock.

Zero and sign restrictions on the structural matrix

The alternative identification scheme we employ is described by Scheme II (Restrictions ZR and SR2). These are restrictions on the structural matrix, A_0, and they are imposed jointly. The structure and interpretation of the following figures is similar to those in the previous subsection.

Figure 4 paints a more ambiguous picture than the baseline sign restriction specification: the median of the admissible impulse responses for asset prices starts at zero. While it turns positive in the short and medium run, and remains so later on, the admissible set does not

[8] Details of the bootstrap procedure can be found in Appendix B.

leave the neighborhood of zero markedly. In the right-hand-side subfigure, the median target impulse response for asset prices, we can observe a similarly ambiguous pattern, even though the impulse response is significantly poisitive after the 10th month.

[FIGURES 4, 5, 6 ABOUT HERE]

Figure 5 mostly corroborates the results of Arias, Caldara, and Rubio-Ramírez (2015, Figure 1., p. 12). Under the zero and sign restriction on the structural matrix the identified monetary policy shock has mostly a contractionary effect on the output. This is in contrast to our baseline results, and results by Uhlig (2005). Since the responses of the GDP deflator are mostly negative, the often discovered price puzzle does not seem to appear in this setup. This picture is further strengthened by the median target impulse responses in Figure 6.

5 Examining the monetary policy shock

As we noted in the previous subsection each of the $s = 1, 2, \ldots, 65000$ impulse responses corresponds to a different admissible model $\widehat{A}_{0,s}^{-1}$. Similarly, to each $\widehat{A}_{0,s}^{-1}$ corresponds an identified monetary policy shock series $\{\hat{\varepsilon}_{ts}^{mp}\}_{t=1,\ldots,T}$ that is obtained from the reduced form residuals by the relation $\hat{\varepsilon}_t = \widehat{A}_0 \hat{u}_t$. In this section we investigate the identified monetary policy shocks by comparing the obtained series $\{\hat{\varepsilon}_{ts}^{mp}\}_{t=1,\ldots,T}$ for each s with the Romer and Romer (2004) series. In order to keep the argumentation compact, in the present section we report results using only the baseline identification restrictions, Scheme I. The following arguments, however, hold similarly for monetary policy shocks identified with the restrictions of Scheme II.[9]

Romer and Romer (2004) develop a monthly measure of monetary policy shocks for the period 1969–1996 that is based on the following methodology: the authors 1.) identify intended federal funds rate changes around meetings of the Federal Open Market Committee (FOMC) by combining narrative accounts of the FOMC meetings and the report of the manager of open market operations; 2.) regress the intended changes on the Fed's internal (so called "Greenbook") forecasts of inflation, real output growth and unemployment in order to control for information about future developments in the economy. Specifically, the regression that

[9] Further results on the latter case can be found in Appendix C.

they estimate is the following (Romer and Romer, 2004, Eq. 1, p. 1061):

$$\Delta ff_m = \alpha + \beta ffb_m + \sum_{i=-1}^{2} \gamma_i \widetilde{\Delta y}_{mi} + \\ + \sum_{i=-1}^{2} \lambda_i (\widetilde{\Delta y}_{mi} - \widetilde{\Delta y}_{m-1,i}) + \sum_{i=-1}^{2} \phi_i \tilde{\pi}_{mi} + \sum_{i=-1}^{2} \theta_i (\tilde{\pi}_{mi} - \tilde{\pi}_{m-1,i}) + \rho \tilde{u}_{m0} + \nu_m, \quad (4)$$

where Δff_m is the change in the intended federal funds rate at the FOMC meeting m, ffb_m is the intended federal funds rate before any changes decided on meeting m, and $\widetilde{\Delta y}_{mi}$, $\tilde{\pi}_{mi}$, $\tilde{u}_{m0}$ are the forecasts of real output growth, inflation and unemployment, respectively, for quarter i at the time of meeting m. The estimated residuals $\hat{\nu}_m$ represent unanticipated monetary policy shocks, and they are averaged over months to obtain the monthly series $\hat{\varepsilon}_t^{rr}$, the RR monetary policy shock series that runs from January 1969 to December 1996.[10] Figure 7 shows the RR series.

[FIGURE 7 ABOUT HERE]

Romer and Romer (2004) carefully argue about the validity of the interpretation of their measure as monetary policy shocks. To our knowledge, only Coibion (2012) provides a critical examination of the implications of the RR shocks. The main objective of Coibion (2012) is to try to reconcile the surprisingly large influence of monetary policy shocks on, for example, output, with earlier similar studies. While he argues that the implications of Romer and Romer (2004) are not robust to, for example, excluding certain episodes in US central banking history, we do not read Coibion (2012) as an argument against the validity of interpreting the RR series as a "pure" monetary policy shock series. We, in fact, go further and argue that any identified structural monetary policy shock series obtained from, e.g., a SVAR analysis should resemble the RR series, $\hat{\varepsilon}_t^{rr}$. Further, Coibion (2012)'s analysis is based partly on alternative monetary policy shock series proposed in Coibion and Gorodnichenko (2011). These alternative shock series allow for *i.)* heteroskedasticity in the error term ν_m, *ii.)* time-varying coefficients in Equation (4). These modifications seem a priori sensible, however, using the monetary policy shock series from Coibion and Gorodnichenko (2011) does not lead to conclusions different from what we describe below.[11] Thus, in the following we view the RR series, as a benchmark monetary policy shock series.

How strongly do the identified monetary policy shocks from our baseline analysis resemble the benchmark monetary policy shocks? In order to answer this question, for each admissible

[10] Note, that in any particular month there can be several FOMC meetings m or no meetings at all.
[11] Estimation results are available upon request.

monetary policy shock series $\hat{\varepsilon}_{ts}^{mp}$, with $s = 1, \ldots, 65000$, we calculate the correlation $\rho_s = \widehat{\mathsf{Corr}}(\hat{\varepsilon}_{ts}^{mp}, \hat{\varepsilon}_t^{rr})$ on the subsample running from January 1969 to December 1996.[12] For simplicity of notation, and without loss of generality, assume that for $s < s'$ it holds that $\rho_s \leq \rho_{s'}$, that is, we index admissible models by their corresponding correlation coefficients.

[FIGURE 8 ABOUT HERE]

Figure 8 contains the histogram of the 65000 obtained correlation coefficients ρ_s. As visual inspection immediately suggests, the correlations are mostly quite weak. Indeed, the average correlation is 0.1863, and the median is 0.1899. In the previous section we reported the median target impulse responses, and used these as further evidence for our results. However, the correlation corresponding to the median target model is 0.1621, i.e., lower than both the average and the median. This implies that at least half of the models have larger ρ_s values than the median target model.

What does it imply for the impulse response analysis if the identified monetary policy shocks are weakly correlated with the Romer and Romer shock series? In Figure 11 we report the impulse responses for each variable of the models $s = 1, \ldots, 100$, i.e., those 100 models where ρ_s is the weakest. Figure 12 contains the impulse responses of models $s = 37450, \ldots, 37550$, i.e., the 100 median models, and Figure 13 contains the impulse responses from models $s = 64900, \ldots, 65000$ – the 100 models with the highest correlations ρ_s.

[FIGURES 11, 12, 13 ABOUT HERE]

The figures suggest that by concentrating on impulse responses from models with low correlations ρ_s, we might be led to notably different qualitative and quantitative conclusions in comparison with the impulse responses of models with high correlations ρ_s. In particular, Figure 11 supports the conclusion that asset prices react positively to a positive monetary policy shock, and they persistently remain so for several periods after impact. Similarly, GDP reacts rather ambiguously on impact, but there is a clear hump-shape suggesting a sluggish response to a monetary policy shock. In contrast, Figure 13 suggests that GDP reacts positively to an increase in the monetary policy instrument followed by a steady gradual decrease. Asset prices, on the other hand, respond mildly negatively on impact.

More surprisingly, according to Figure 12, models that yield shocks featuring an average correlation with the RR series hardly support any unambiguous empirical conclusion. Anything can happen as regards the shapes and magnitudes of impulse responses for *all* variables, and,

[12] Note, that, while for $\hat{\varepsilon}_{ts}^{mp}$ the t index runs from January 1960 to December 2007, the RR series, $\hat{\varepsilon}_t^{rr}$, is available only between January 1969 and December 1996.

especially, for those responses that are left unrestricted by the identification scheme. This finding is worth emphasizing for two reasons. First, the large majority of models are close to an average (or median) correlation level, cf., Figure 8. Second, note, that a central idea of sign restrictions, emphasized by Uhlig (2005), is to leave those variables' responses agnostic whose behavior is of key interest to the analysis at hand. Hence, if one randomly selects two admissible models, then they might, with high probability, lead to distinct conclusions especially with respect to key variables. Further, since the median target impulse response is inevitably tied to some "average" model, analyzing the median target alone offers inconclusive, or even misleading results.

In sum, we have argued in the previous paragraphs that the qualitative and quantitative features of the impulse response functions of models $s = 1, \ldots, 65000$ are closely linked to their implied correlation with the RR shocks, ρ_s. Models with similar *high* (or *low*) ρ_s values imply similar impulse responses. Models with notably different ρ_s values imply notably different impulse responses.

Thus, if one accepts that the identified monetary policy shocks should ideally be closely correlated with the RR shocks, one should put particular emphasis on analyzing the impulse responses displayed in Figure 13. Indeed, such highly correlated models grasp best what is implied by a "true" monetary policy shock. Thus, in the next section we reconsider our baseline results concentrating on those 100 models that yield monetary policy shocks that are closest to the RR series. In addition, we give special attention to model number 65000 showing the highest correlation with $\hat{\varepsilon}_t^{rr}$.

6 Analyzing models with the highest correlation

We have argued that most of the monetary policy shocks identified in our baseline model correlate only weakly with the Romer and Romer (2004) monetary policy shock series that we view as a benchmark series. In this section we reconsider the impulse responses concentrating only on $\mathcal{S}^{high} = \{\, s : s \in [64900, 65000] \cap \mathbb{N} \,\}$, those 100 models that have the highest ρ_s, and $s^* = 65000$, the model with the highest ρ_s.

Sign restrictions on impulse responses

Figure 14 displays the minimal envelope, the maximal envelope and the median of impulse responses of models in $\mathcal{S}^{high}$ identified through Scheme I (Restriction SR1). In response to a positive monetary policy shock we find evidence for a positive, rising, but then quickly

falling and in the end negative GDP response. This finding strengthens the result from the baseline analysis. Contrary to the baseline analysis, however, we also see evidence for a zero or mildly negative asset price response, which seems to reconcile results from the baseline analysis with existing previous studies. Note, that the hump shapes around months 7 and 10, respectively, are typical in the considered set $\mathcal{S}^{high}$, and this shape is also in line with impulse responses obtained earlier in the literature (e.g., Galí and Gambetti (2015)). The remaining variables behave similarly to the baseline analysis, but we get a notably sharper picture on how the GDP deflator, non-borrowed reserves and the federal funds rate behave in reaction to a positive monetary policy shock.

[FIGURES 14, 15 ABOUT HERE]

We also show in Figure 15 the impulse responses corresponding to model s^*, i.e., the model that yields identified monetary policy shock series that has the highest correlation with the RR series. Model s^* strengthens the results in the previous paragraph: in particular, asset prices react sizably negatively and they return only slowly to their starting value.

In addition to the impulse responses of model s^*, Figure 9 displays the (standardized) identified monetary policy shock series corresponding to model s^*, together with the (standardized) RR series. As visual inspection suggests, $\hat{\varepsilon}^{mp}_{ts^*}$ matches quite well with $\hat{\varepsilon}^{rr}_t$, especially in those times when $\hat{\varepsilon}^{rr}_t$ shows sizable swings strengthening the argument to restrict attention to models that are highly correlated with the RR shocks.[13]

[FIGURE 9 ABOUT HERE]

Zero and sign restrictions on the structural matrix

Figure 16 shows the impulse response of models in $\mathcal{S}^{high}$ determined from Scheme II (Restrictions ZR and SR2). A remarkable feature of the GDP response to a positive monetary policy shock is a qualitative similarity to the GDP response of Figure 14. This finding is worth noting especially in light of our Figure 5, and the arguments of Arias, Caldara, and Rubio-Ramírez (2015), where the main tendency of the GDP response is markedly negative.

[FIGURE 16 ABOUT HERE]

Further, while in the baseline analysis of Scheme II we didn't find strong evidence for the price puzzle, based on Figure 16 we cannot claim that the existence of the price puzzle is not

[13]Note, that an important caveat in interpreting results from model s^*, however, is that the model and the precise form of the impulse responses depend on the particular draw of the orthogonal matrix.

plausible. The asset price response is centered around zero. Thus, contrary to the results in the previous paragraph, we find no evidence of exogenous monetary policy shocks affecting asset prices. The rest of the responses exhibit high similarity to the baseline analysis, but concentrating only on hundred models leads to sharper conclusions.

In comparison to the figures describing $\mathcal{S}^{high}$ the impulse responses of model s^* in Scheme II are quite sensitive to the particular draw of the rotation matrix, therefore, we do not attempt to analyze the impulse responses this particular scheme. The (standardized) monetary policy shock series implied by Scheme II is shown in Figure 10, jointly with the (standardized) RR series. Similarly to the conclusions of the previous subsection, we can observe that the monetary policy shock series matches the RR series quite well especially in the high volatility phases.

[FIGURE 10 ABOUT HERE]

7 Discussion and robustness

Our proposal of combining structural identifying assumptions with some benchmark series is similar in spirit to two different approaches proposed earlier in the literature. First, Faust, Swanson, and Wright (2004) identify monetary policy shocks by requiring the federal funds rate response to the policy shock to be equal to a certain benchmark response directly measured from federal funds futures data. Second, Mertens and Ravn (2013, 2014) utilize a narrative series as a proxy for the identified policy shock series and thereby provide additional identifying moment conditions. While both approaches solve the identification problem in a data-oriented way, their explicit aim is exact identification. In contrast, we believe that there can be several economic (structural) models compatible with the data. Sign restrictions, or non-exact identification in general, are adequate assumptions in line with this view as long as one carefully interprets the results.

From a methodological point of view, in the previous sections we argued that one should evaluate the effects of structural monetary policy shocks only within a subset of the set of *admissible* models. That is, in our analysis we started out by constraining the set of admissible models with set-identifying restrictions. An other possibility would be to simply generate several $\widehat{A}_0^{-1}$ matrices that do not satisfy any a priori structural assumptions and select the one that generates an estimated shock that has the highest correlation with the Romer and Romer (2004) series. However, there are several theoretical and empirical caveats for this approach. First, in the comparison exercise we might simply discover a high but

spurious correlation between the benchmark and the estimated series. In particular, if the estimated series does not have a structural interpretation, then we cannot claim that the selected estimated series is, in fact, not a completely different shock. Second, the selected estimated series will highly depend on the particular random draw of $\widehat{A}_0^{-1}$. In our analysis it turns out that by not restricting the response of GDP deflator to be negative in the first several periods we immediately encounter the so-called price puzzle in a particularly severe form: the GDP deflator reacts positively to a monetary policy shock and its response remains positive in the long run. Further, by considering unrestricted models we could not significantly improve the "fit" to the RR shocks: the maximal correlation is around 0.42 compared with the approximately 0.39 of the set-identified specifications.

Our approach in this paper is explicitly frequentist. Thus, we cannot discriminate statistically between competing admissible models, and claiming that a particular admissible model is "most likely" (Inoue and Kilian, 2013) is not feasible. Nevertheless, our benchmarking approach extends the possibilities of empirical SVAR analysis in two directions. First, it restricts the set of models beyond what is achievable by the structural assumptions alone. This gives the possibility to sharply focus the evaluation of empirical and economic implications of the identified structural models. The empirical analysis then avoids the point made by Kilian and Murphy (2012) about the potential perils of analyzing summary statistics of the identified set when this set is too large. Second, while speaking of most likely models is not possible in a frequentist setting, if one accepts the postulated structural assumptions and the validity of the benchmark series, the models in $\mathcal{S}^{high}$ can be argued to be frequentist counterparts to the most likely models of Inoue and Kilian (2013).

We argued earlier that using the series proposed by Coibion and Gorodnichenko (2011) as the benchmark series instead of the RR series does not change any of our conclusions. This is despite the fact that Coibion (2012) arrives at distinct conclusions regarding the contribution of monetary policy shocks to fluctuations of real variables using these alternative narrative-based series. This might imply that our results are not driven by the "identifying power" of the benchmark shocks. However, if we use a completely uninformative simulated white noise series as benchmark series in place of the RR series, then the ordering of the models according to their ρ_s coefficients yield completely uninformative results: the impulse responses for models in $\mathcal{S}^{high}$ are similarly unstructured as the impulse responses for models in any other subsets of the ordering. This finding strengthens our results: investigating models relative to a benchmark indeed provides additional information.

Up until now we were silent about how more conventionally identified monetary policy shocks compare with the RR shock. As an example, using our baseline variable ordering[14] with a straight-forward recursive identification we obtained a correlation of $\rho = 0.3838$, which is very close to our best set-identified models. Values of similar magnitude for exact identification were reported also by Coibion (2012). These correlations are higher than the majority of correlations that we uncover in our analysis. On the one hand, this result might indicate that the applied sign restrictions need not be very successful assumptions to identify monetary policy shocks. However, the fact that models with the highest ρ_s have correlation around 0.40 implies that by a careful analysis of admissible models we can *improve* on other, especially exact identification procedures while at the same time settling on less restrictive identifying assumptions. We leave a comprehensive comparative evaluation of other exact identification schemes and empirical specifications for future research.[15]

8 Conclusions

How do asset prices respond to exogenous monetary policy shocks? We provide empirical results on this question. To this end, we augment the VAR specification of Uhlig (2005) with the S&P 500 Composite Index, and estimate the model on monthly US data from 1959 January to 2007 December. We use two identification schemes that result in set identification of the monetary policy shock. First, we use the sign restrictions put forth in Uhlig (2005) (Scheme I). Second, we utilize zero and sign restrictions on the structural matrix A_0 proposed by Arias, Caldara, and Rubio-Ramírez (2015) (Scheme II).

The SVAR impulse responses identified via Scheme I and Scheme II both point towards a mild positive asset price response to an increase in the monetary policy instrument. We argue that the resulting identified monetary policy shocks correlate only weakly with the monetary policy shock series of Romer and Romer (2004) that we view as a benchmark series for monetary policy shocks. We show that this finding matters greatly when analyzing (structural) impulse responses. In particular, the majority of admissible models yield impulse responses that vary widely in their shapes and impact magnitudes. Thus, we propose to restrict attention to those specifications that yield monetary policy shocks highly correlated with the RR series, and we show that the impulse response analysis of these models leads to more robust and reliable conclusions.

[14] Real GDP, GDP deflator, commodity price index, stock price index, federal funds rate, non-borrowed reserves, total reserves.

[15] Estimation results regarding the experiments in the above section are available upon request.

Ultimately, we find evidence of asset prices responding mildly negatively (Scheme I) or with ambiguous sign (Scheme II) to a positive monetary policy shock. Besides the asset price response, we also uncover a mildly positive output response under both identification schemes when concentrating on models with the highest correlation with the Romer and Romer (2004) series. The expansionary effect of a "contractionary" monetary policy shock on output may seem surprising, Ramey (2016), however, points out that the consensus on "contractionary" monetary policy shocks indeed having contractionary effects easily disappears once one lifts the recursiveness identification assumption.

While the theme of comparing identified monetary policy shocks with the Romer and Romer (2004) series is quite specific to monetary policy applications on US data, our proposed approach of evaluating (set-) identified shocks against some benchmark series is more general and can be applied to a wide variety of empirical questions. As Kilian and Murphy (2012, p. 1186) point out: "If the set of admissible models remains large, the most useful exercise will be to search for the admissible model most favorable to each of the competing economic interpretations (...)". We believe that our approach complements and extends this advice, and is, thus, beneficial for empirical research. We have shown that comparing structurally (set-) identified shocks to a benchmark series can uncover by default hidden, but relevant and robust empirical conclusions.

References

Arias, J. E., D. Caldara, and J. F. Rubio-Ramírez (2015): "The Systematic Component of Monetary Policy in SVARs: An Agnostic Identification Procedure," International Finance Discussion Papers 1131, Board of Governors of the Federal Reserve System (U.S.).

Arias, J. E., J. F. Rubio-Ramírez, and D. F. Waggoner (2014): "Inference Based on SVARs Identified with Sign and Zero Restrictions: Theory and Applications," Working paper.

Assenmacher-Wesche, K., and S. Gerlach (2010): "Monetary Policy and Financial Imbalances: Facts and Fiction," *Economic Policy*, 25, 437–482.

Bernanke, B. S., and K. N. Kuttner (2005): "What Explains the Stock Market's Reaction to Federal Reserve Policy?," *The Journal of Finance*, 60(3), 1221–1257.

Bjørnland, H. C., and K. Leitemo (2009): "Identifying the Interdependence Between US Monetary Policy and the Stock Market," *Journal of Monetary Economics*, 56(2), 275–282.

Brockwell, P. J., and R. A. Davis (1991): *Time Series: Theory and Methods*. Springer, New York, second edn.

Canova, F., and G. De Nicoló (2002): "Monetary Disturbances Matter for Business Fluctuations in the G-7," *Journal of Monetary Economics*, 49(6), 1131–1159.

Carlstrom, C. T., T. S. Fuerst, and M. Paustian (2009): "Monetary Policy Shocks, Choleski Identification, and DNK Models," *Journal of Monetary Economics*, 56(7), 1014–1021.

Castelnuovo, E., and S. Nisticò (2010): "Stock Market Conditions and Monetary Policy in a DSGE Model for the U.S.," *Journal of Economic Dynamics and Control*, 34(9), 1700–1731.

Coibion, O. (2012): "Are the Effects of Monetary Policy Shocks Big or Small?," *American Economic Journal: Macroeconomics*, 4(2), 1–32.

Coibion, O., and Y. Gorodnichenko (2011): "Monetary Policy, Trend Inflation, and the Great Moderation: An Alternative Interpretation," *American Economic Review*, 101(1), 341–70.

Faust, J. (1998): "The Robustness of Identified VAR Conclusions About Money," *Carnegie-Rochester Conference Series on Public Policy*, 49, 207 – 244.

Faust, J., E. T. Swanson, and J. H. Wright (2004): "Identifying VARs Based on High Frequency Futures Data," *Journal of Monetary Economics*, 51(6), 1107–1131.

Fry, R., and A. Pagan (2011): "Sign Restrictions in Structural Vector Autoregressions: A Critical Review," *Journal of Economic Literature*, 49(4), 938–60.

Galí, J., and L. Gambetti (2015): "The Effects of Monetary Policy on Stock Market Bubbles: Some Evidence," *American Economic Journal: Macroeconomics*, 7(1), 233–57.

Inoue, A., and L. Kilian (2013): "Inference on impulse response functions in structural VAR models," *Journal of Econometrics*, 177(1), 1–13.

Kilian, L. (1998): "Small-Sample Confidence Intervals For Impulse Response Functions," *The Review of Economics and Statistics*, 80(2), 218–230.

Kilian, L., and D. P. Murphy (2012): "Why Agnostic Sign Restrictions Are Not Enough: Understanding The Dynamics Of Oil Market Var Models," *Journal of the European Economic Association*, 10(5), 1166–1188.

Lanne, M., M. Meitz, and P. Saikkonen (2015): "Identification and estimation of non-Gaussian structural vector autoregressions," CREATES Research Papers 2015-16, School of Economics and Management, University of Aarhus.

Linnemann, L., G. B. Uhrin, and M. Wagner (2016): "Government Spending Shocks and Labor Productivity," Discussion paper Nr. 9/2016, SFB 823.

Lütkepohl, H., and A. Netšunajev (2014): "Structural Vector Autoregressions with Smooth Transition in Variances: The Interaction between U.S. Monetary Policy and the Stock Market," Discussion Papers of DIW Berlin 1388, DIW Berlin, German Institute for Economic Research.

McCracken, M. W., and S. Ng (2016): "FRED-MD: A Monthly Database for Macroeconomic Research," *Journal of Business & Economic Statistics*, 34(4), 574–589.

Mertens, K., and M. O. Ravn (2013): "The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States," *American Economic Review*, 103(4), 1212–47.

——— (2014): "A Reconciliation of SVAR and Narrative Estimates of Tax Multipliers," *Journal of Monetary Economics*, 68(S), S1–S19.

Meyer, C. D. (2000): *Matrix Analysis and Applied Linear Algebra*. Society for Industrial and Applied Mathematics, Philadelphia.

Mönch, E., and H. Uhlig (2005): "Towards a Monthly Business Cycle Chronology for the Euro Area," *Journal of Business Cycle Measurement and Analysis*, 2005(1), 43–69.

Neri, S. (2004): "Monetary Policy and Stock Prices: Theory and Evidence," Temi di discussione (Economic working papers) 513, Bank of Italy, Economic Research and International Relations Area.

Patelis, A. D. (1997): "Stock Return Predictability and The Role of Monetary Policy," *The Journal of Finance*, 52(5), 1951–1972.

Ramey, V. A. (2016): "Macroeconomic Shocks and Their Propagation," article prepared for Handbook of Macroeconomics, vol 2., UCSD.

Rigobon, R., and B. Sack (2004): "The Impact of Monetary Policy on Asset Prices," *Journal of Monetary Economics*, 51(8), 1553–1575.

Romer, C. D., and D. H. Romer (2004): "A New Measure of Monetary Shocks: Derivation and Implications," *American Economic Review*, 94(4), 1055–1084.

Rubio-Ramírez, J. F., D. F. Waggoner, and T. Zha (2010): "Structural Vector Autoregressions: Theory of Identification and Algorithms for Inference," *Review of Economic Studies*, 77(2), 665–696.

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

Thorbecke, W. (1997): "On Stock Market Returns and Monetary Policy," *The Journal of Finance*, 52(2), 635–654.

Uhlig, H. (2005): "What Are the Effects of Monetary Policy on Output? Results from an Agnostic Identification Procedure," *Journal of Monetary Economics*, 52(2), 381–419.

A Data

All data, except for the Romer and Romer (2004) series, is fully available to us from 1959 January to 2007 December. The data was gathered on 16.12.2015.

- **Romer and Romer monetary policy shocks:** The monthly series from 1/1/1969 to 12/1/1996 was obtained from Christina Romer's website: `http://eml.berkeley.edu/~cromer/#data`.

- **Real GDP:** The monthly GDP was approximated with state-space methods using the quarterly GDP series GDPC1 and the monthly industrial production series INDPRO obtained from the FRED database. The interpolation method is described in Mönch and Uhlig (2005).

- **GDP deflator:** The monthly GDP deflator was approximated with state-space methods using the quarterly GDP deflator series (GDPDEF) and the monthly series CPIAUCSL (consumer price index for all urban consumers) and PPIFGS (producer price index for finished goods). All series were downloaded from the FRED database, and the interpolation method is described in Mönch and Uhlig (2005).

- **Commodity price index:** Daily data from the Commodity Research Bureau BLS spot index was obtained from Thomson Reuters' Datastream. Monthly observations were calculated as the averages of daily observations for each month.

- **Stock price index:** Monthly observations of the S&P 500 composite index was obtained from the FRED MD project website (`https://research.stlouisfed.org/econ/mccracken/fred-databases/`) maintained by Michael W. McCracken. For the empirical analysis, the values were deflated by means of the GDP deflator.

- **Measures of monetary policy:** Monthly series of the federal funds rate (FEDFUNDS), total reserves (TOTRESNS), and non-borrowed reserves (BOGNONBR) were obtained from the FRED database.

B Econometric details

VAR model and impulse responses

Recall from the main text that we consider the following K-dimensional structural VAR,

$$A_0 y_t = A_1 y_{t-1} + A_2 y_{t-2} + \cdots + A_p y_{t-p} + \varepsilon_t, \tag{5}$$

where $y_t \in \mathbb{R}^K$, $\varepsilon_t \sim WN(0, I_K)$, $A_0, \ldots, A_p \in \mathbb{R}^{K \times K}$, and A_0, what we call the *structural matrix*, is assumed to be non-singular. In order to define a unique lag length we assume that $A_p \neq 0$. In the empirical application $K = 7$, and $p = 12$. In the above equation ε_t is the vector of *structural innovations*. The corresponding, estimable reduced form is

$$y_t = B_1 y_{t-1} + \cdots + B_p y_{t-p} + u_t, \tag{6}$$

with $B_i = A_0^{-1} A_i,\ i = 1, \ldots, p$. For u_t, the vector of *reduced form innovations* the following holds: $A_0^{-1}\varepsilon_t = u_t \sim WN(0, \Sigma_u)$. Writing $B(z) = I_K - B_1 z - \cdots - B_p z^p$ we assume, that the reduced form is causal, that is, $\det(B(z)) \neq 0\ \forall |z| \leq 1$. Then the moving average representation of y_t exists and is given by (Brockwell and Davis, 1991, Th. 11.3.1, p. 418)

$$y_t = \sum_{j=0}^{\infty} \Phi_j u_{t-j} = \sum_{j=0}^{\infty} \Theta_j \varepsilon_{t-j}, \quad \Phi_0 = I_K, \tag{7}$$

where element (i, k) of the coefficient $\Theta_j = \Phi_j A_0^{-1}$ is interpreted as the reaction of the i-th variable on the k-th structural innovation at horizon j.

We estimate the parameters with ordinary least squares, obtaining $\widehat{B}_1, \ldots, \widehat{B}_p$ and $\widehat{\Sigma}_u$. The corresponding estimate for the reduced form impulse response sequence $(\hat{\Phi}_j)_{j \geq 0}$ follows immediately. The starting point for estimating sign restricted impulse responses is the lower triangular Cholesky decomposition of $\widehat{\Sigma}_u = \widehat{A}_u^c \widehat{A}_u^{c\prime}$. Since $\widehat{\Sigma}_u$ is symmetric and positive definite, its Cholesky decomposition is unique (Meyer, 2000, p. 154). Note, that for any $K \times K$ orthogonal matrix Q with $Q'Q = QQ' = I_K$ it holds that

$$\widehat{\Sigma}_u = \widehat{A}_u^c Q Q' \widehat{A}_u^{c\prime}. \tag{8}$$

We are interested in finding those $\widehat{A}_u^c Q = \widehat{A}_0^{-1}(Q)$ matrices that imply structural form impulse response sequences $(\hat{\Theta}_j)_{j \geq 0} = (\hat{\Phi}_j \widehat{A}_0^{-1}(Q))_{j \geq 0}$ that satisfy the sign restrictions maintained in the main text. To this end we use the method proposed by Rubio-Ramírez, Waggoner, and

Zha (2010). Let H be the length of the impulse response horizon that we wish to estimate, and let $J \leq H$ be the length of the horizon on which sign restrictions are imposed. Then the procedure can be described as follows:

1. Draw a matrix M with i.i.d. standard normal entries and perform the QR-decomposition of the matrix $M = QR$. The resulting matrix Q is orthogonal and has the uniform (or Haar) distribution on the group of orthogonal matrices.

2. Calculate the corresponding structural impulse response function $\{\widehat{\Theta}_j^Q\}_{j=0,\ldots,H} = \{\widehat{\Phi}_j \widehat{A}_0^{-1}(Q)\}_{j=0,\ldots,H}$ and verify whether the formulated sign restrictions are fulfilled for $j = 1, \ldots, J$. If so, keep $\left\{\widehat{\Theta}_j^Q\right\}_{j=0,\ldots,H}$, otherwise discard it.

3. Repeat steps 1–2 until the set of retained structural impulse responses contains $S = 65000$ elements.

In the empirical application, we set $H = 60$, and $J = 4$ according to Restriction SR in the main text.

In order to simulate the set of impulse responses resulting from the sign and zero restrictions on A_0 we use the method proposed by Arias, Rubio-Ramírez, and Waggoner (2014). Let $\widehat{A}$ be $(\widehat{A}_u)^{-1}$, that is, the inverse of the Cholesky decomposition of $\widehat{\Sigma}_u$. Then, for our particular application, the algorithm can be described by the following steps:

1. Find a matrix $N_1 \in \mathbb{R}^{K\times(K-2)}$ with $N_1' N_1 = I_{K-2}$ such that $\widehat{A}_{[(K-1:\,K),\bullet]} N_1 = 0$, with $\widehat{A}_{[(K-1:\,K),\bullet]}$ denoting the $2 \times K$ matrix formed by the $(K-1)$-th and K-th rows of $\widehat{A}$.

2. Generate a vector $z \in \mathbb{R}^K$ with i.i.d. standard normally distributed entries and form the vector:

$$q = \frac{1}{||[N_1\ 0_{K\times 2}]z||}[N_1\ 0_{K\times 2}]z, \tag{9}$$

   i.e., project the vector z on the space spanned by N_1 and normalize it to unit length.

3. Find a matrix $N_2 \in \mathbb{R}^{K\times(K-2)}$ with $N_2' N_2 = I_{K-2}$ such that $q' N_2 = 0$.

4. Draw a matrix $M \in \mathbb{R}^{(K-2)\times(K-2)}$ with i.i.d. standard normal entries and calculate the QR decomposition of $N_2 M$, i.e.,

$$N_2 M = [\tilde{Q}_1\ \ \tilde{Q}_2] \begin{bmatrix} R_1 \\ 0 \end{bmatrix}, \tag{10}$$

with $\tilde{Q}_1 \in \mathbb{R}^{K\times(K-2)}$.

5. Form the matrix $Q^+ = [q\ \tilde{Q}_1]$, calculate the corresponding structural matrix $\widehat{A}_0^{Q^+} = Q^{+\prime}\widehat{A}$, and verify whether the formulated sign restrictions are fulfilled. If so, keep $\widehat{A}_0^{Q^+}$, and the implied structural parameters, otherwise discard it. Note that by construction, the zero restrictions on the structural matrix hold for all draws.

6. Repeat these steps until the set of retained structural parameters contains $S = 65000$ elements.

Inference on the median target impulse response

The median target (MT) impulse response is the impulse response that is (element-wise) closest in terms of standardized squared distance to the pointwise median of the set of sign restricted impulse responses (termed here the *median curve*). Our implementation of the MT impulse response follows Fry and Pagan (2011). In particular, let $\widehat{\Theta}^s = \{\widehat{\Theta}_j^s\}_{j=0,\ldots,H}$ be the set of structural impulse responses for $s = 1,\ldots,S$ with $S = 65000$, estimated on horizons $0,\ldots,H$. Denote the (element-wise) median curve as $\widehat{\Theta}_{med} = \{\widehat{\Theta}_{j,med}\}_{j=0,\ldots,H}$. The median target impulse response is defined as:

$$\widehat{\Theta}^{MT} = \text{argmin}_{s=1,\ldots,S} \sum_{r\in\mathcal{R}}\sum_{k\in\mathcal{K}}\sum_{j\in\mathcal{J}} \left(\frac{\widehat{\Theta}_{j,s}^{(r,k)} - \widehat{\Theta}_{j,med}^{(r,k)}}{\widehat{SD}_{r,k,j}}\right)^2, \tag{11}$$

with $\mathcal{R},\mathcal{K} \subseteq \{1,\ldots,K\}$ and $\mathcal{J} \subseteq \{0,\ldots,H\}$. Starting with $\overline{\widehat{\Theta}_{j,s}^{(r,k)}} = \frac{1}{S}\sum_{s=1}^{S}\widehat{\Theta}_{j,s}^{(r,k)}$, we can write $\widehat{SD}_{r,k,j} = \sqrt{\left(\frac{1}{S}\sum_{s=1}^{S}\widehat{\Theta}_{j,s}^{(r,k)} - \overline{\widehat{\Theta}_{j,s}^{(r,k)}}\right)^2}$. That is, $\widehat{SD}_{r,k,j}$ is, for each impulse response horizon j, each shock k, and each variable r the (pointwise) empirical standard deviation of the set of admissible impulse responses. In the empirical analysis we consider in Equation (11) only responses to the monetary policy shock, i.e., $\mathcal{K} = \{5\}$, and all the impulse responses, $\mathcal{R} = \{1,\ldots,7\}$. The length of the estimated impulse response horizon is $H = 60$, and $\mathcal{J} = \{0,\ldots,12\}$, i.e., impact period plus one year.

We denote by Q^{MT} the rotation that yields $\widehat{A}_0^{-1}(Q^{MT})$ corresponding to the median target model. While the median curve does not correspond to any particular structural model, it is possible to provide inference on the MT impulse response: an impulse response that corresponds to a well-defined, unique structural model.

Our bootstrap procedure for inference on the MT response follows Linnemann, Uhrin, and Wagner (2016). The algorithm can be described by the following steps:

1. Generate a bootstrap sample, $y_1^*, \ldots, y_T^*$ using the Kilian (1998) bootstrap.

2. Estimate the parameters of the VAR model using y_t^*, resulting in parameter estimates $\widehat{B}_1^*, \ldots, \widehat{B}_p^*$. Calculate the structural impulse response function using these parameter estimates and the *original* $\widehat{A}_0^{-1}(Q^{MT})$.

3. Verify whether the impulse response function from the previous item, $\{\widehat{\Theta}_j^{Q^{MT}*}\}_{j=0,\ldots,J}$, satisfies the formulated sign restrictions. If it does, keep it, otherwise discard it.

4. Repeat the above steps until 1000 impulse responses are retained and calculate pointwise bootstrap confidence bands as usual from these 1000 impulse responses.

C Additional results to comparison with the Romer and Romer (2004) shocks.

In this Appendix we report results on comparison between the monetary policy shock series, and the Romer and Romer (2004) (RR) series , $\hat{\varepsilon}_t^{rr}$, where the monetary policy shocks are identified with Scheme II (Restrictions ZR and SR2). With a slight abuse of notation we denote the identified monetary policy shock series as $\hat{\varepsilon}_{ts}^{mp}$, similarly to the series obtained with identification Scheme I. In the context of the present Appendix no confusion should arise from this shorthand. The arguments of Section 5 remain valid, and are further strengthened by the evidence below.

[FIGURE 17 ABOUT HERE]

Figure 17 displays the histogram of the 65000 obtained correlation coefficients $\rho_s = \widehat{\mathsf{Corr}}(\hat{\varepsilon}_{ts}^{mp}, \hat{\varepsilon}_t^{rr})$. As visual inspection suggests, the large majority of models are only midly correlated with the RR series. The average level of correlation is 0.1692, and the median is 0.1654. These values suggest, that the achievable correlation level using Scheme II are on average lower than those attained using Scheme I. However, the maximal correlation level (0.4) is very similar to the one obtained using Scheme I and the simple recusrive identification scheme, cf., Section 7). These three experiments suggest that there may be a cap on the achievable correlation level that is most likely influenced by the data and the model specification, but not the identification scheme.

Figures 18 – 20 display models with low ρ_s, medium ρ_s, and high ρ_s coefficients, respectively. As discussed in the main text, these figures show models that imply quite distinct impulse reponses both qualitatively and quantitatively. The lack of information content of Figure

19, displaying models with medium ρ_s coefficients, is even more pronounced than that of the corresponding figure in the main text, Figure 12.

[FIGURES 18, 19, 20 ABOUT HERE]

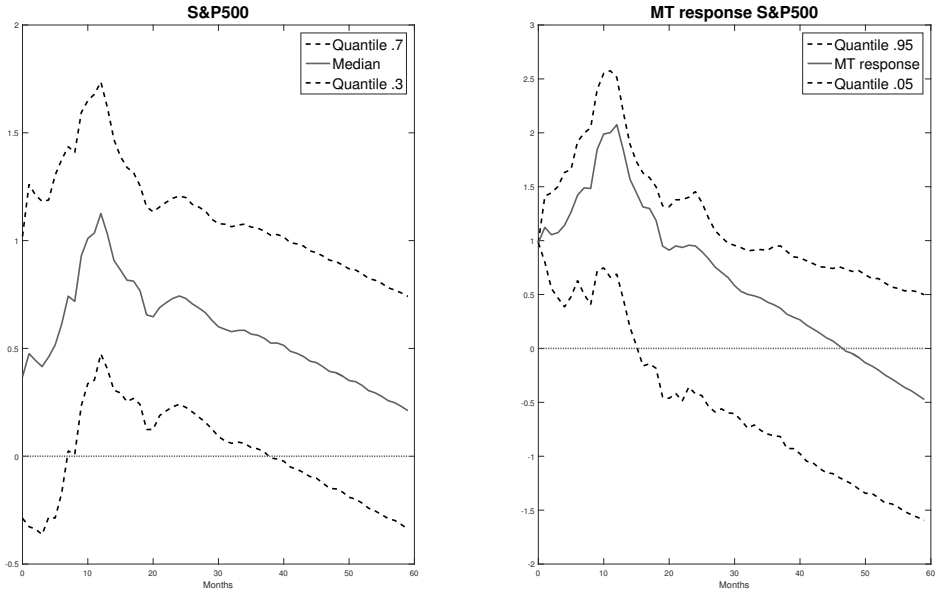

Figure 1: Sign restricted (left) and median target (right) impulse responses of the S&P 500 index for a one per cent increase in the federal funds rate. Identification Scheme I.

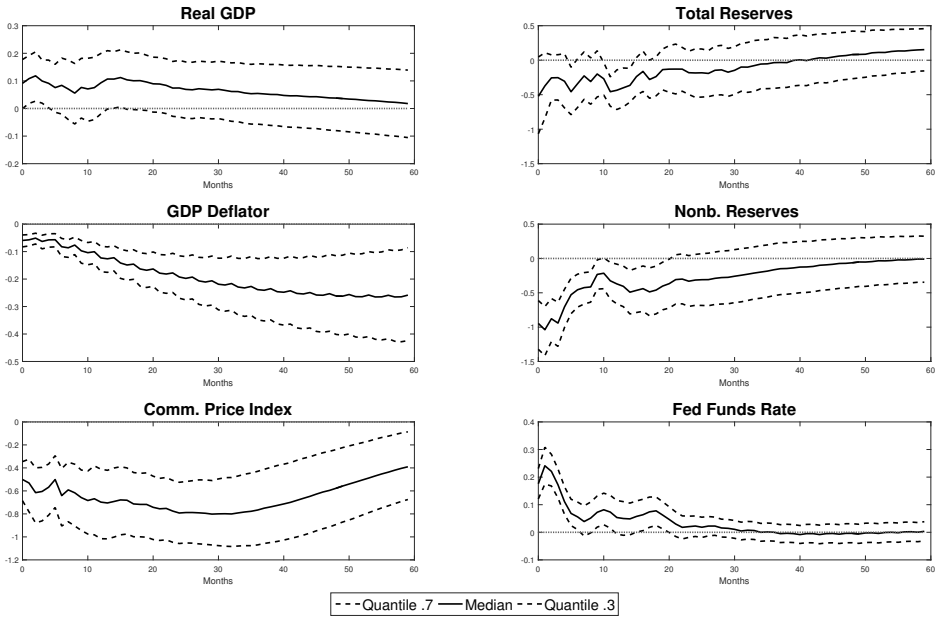

Figure 2: Sign restricted impulse responses for a one per cent increase in the federal funds rate. Identification Scheme I.

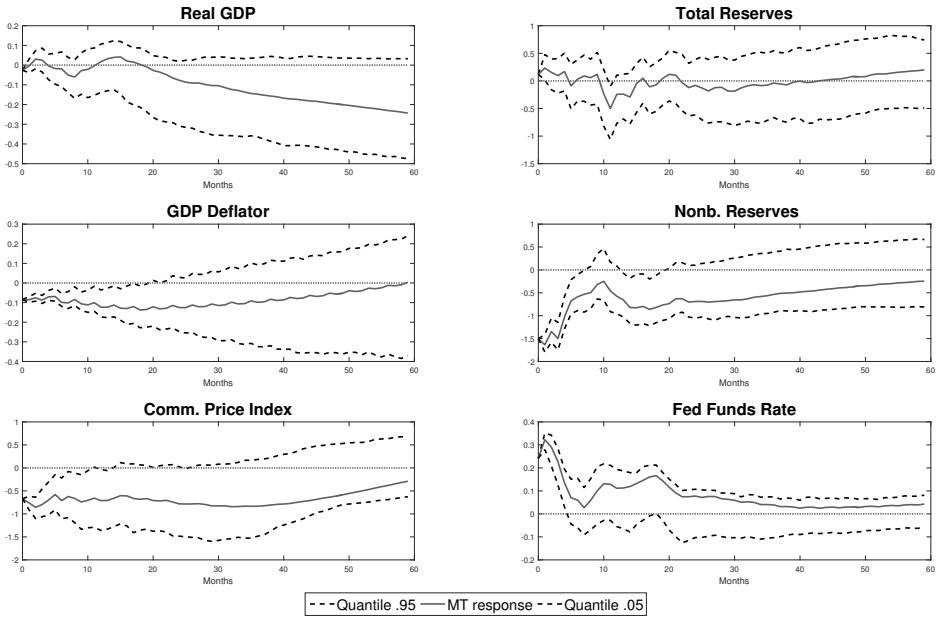

Figure 3: Median target impulse responses for a one per cent increase in the federal funds rate. Identification Scheme I.

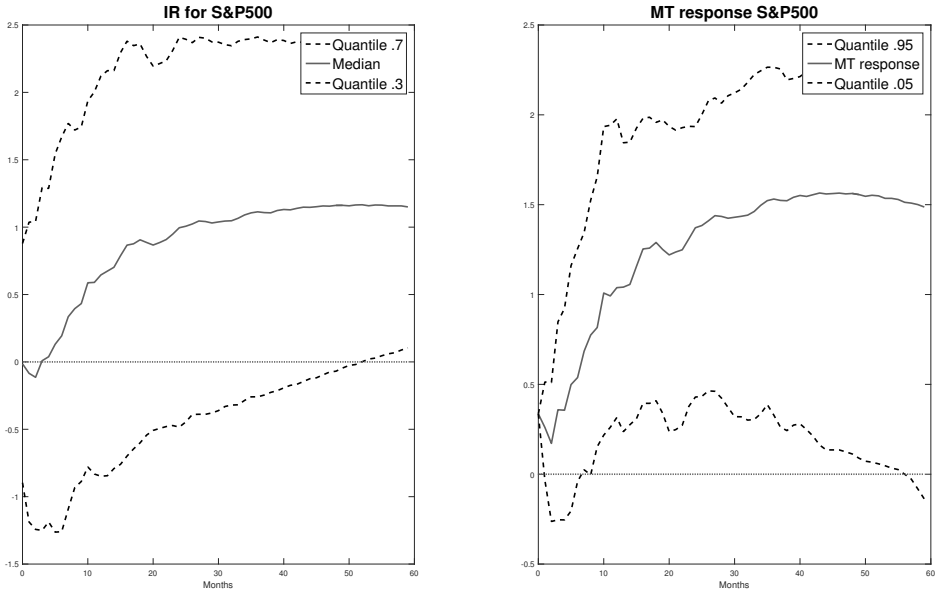

Figure 4: Sign restricted (left) and median target (right) impulse responses of the S&P 500 index for a one per cent increase in the federal funds rate. Identification Scheme II.

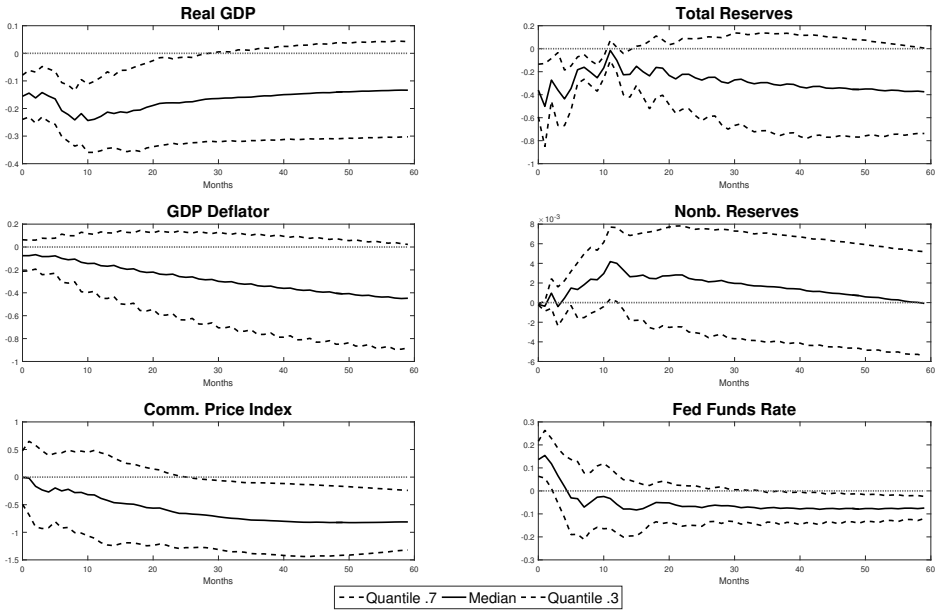

Figure 5: Sign restricted impulse responses for a one per cent increase in the federal funds rate. Identification Scheme II.

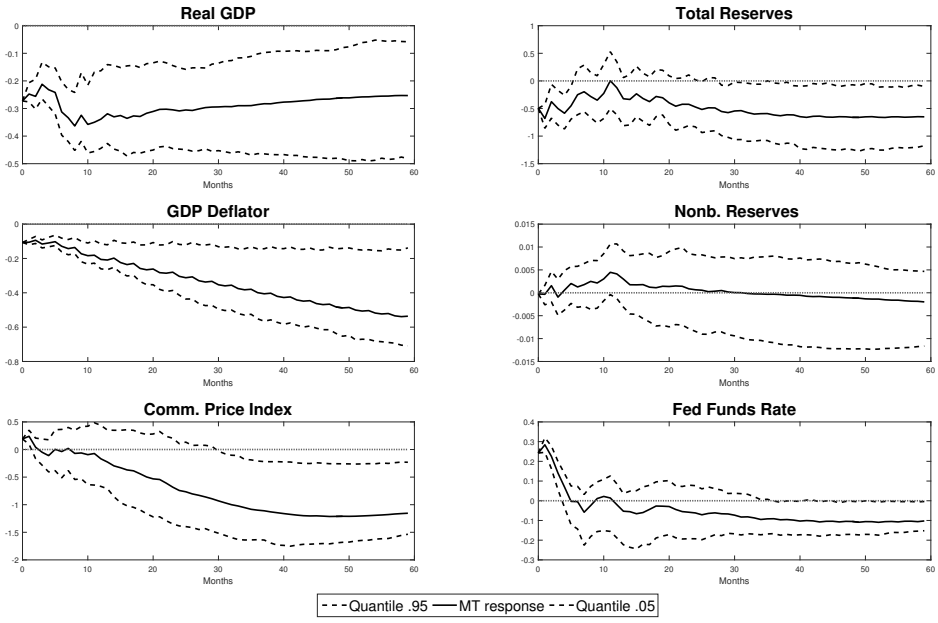

Figure 6: Median target impulse responses for a one per cent increase in the federal funds rate. Identification Scheme II.

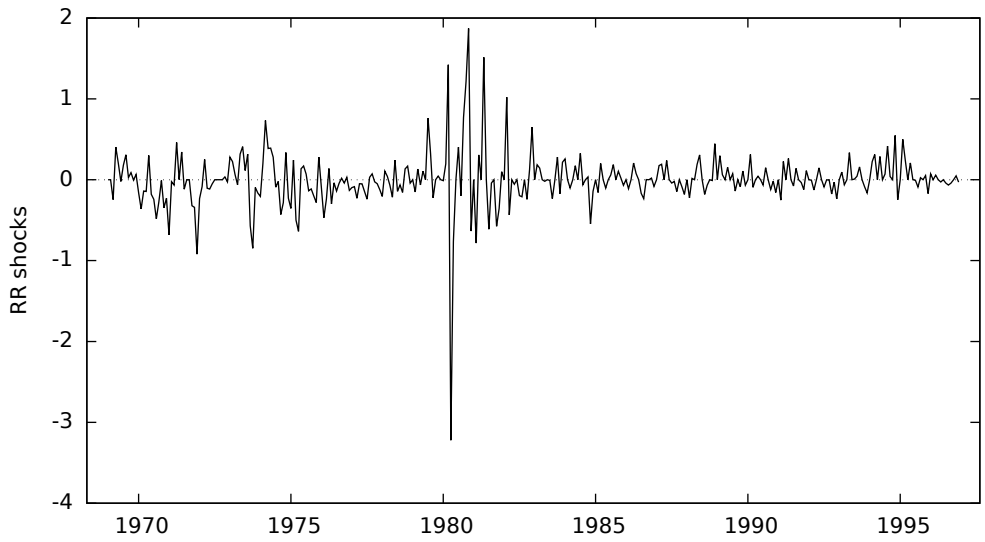

Figure 7: Romer and Romer (2004) shock series.

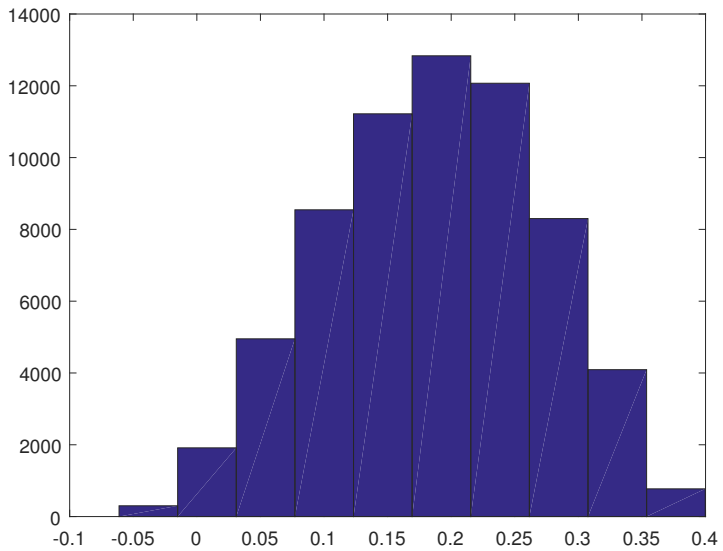

Figure 8: Histogram of correlations between $\hat{\varepsilon}_t^{mp}$ and $\hat{\varepsilon}_t^{rr}$. Scheme I.

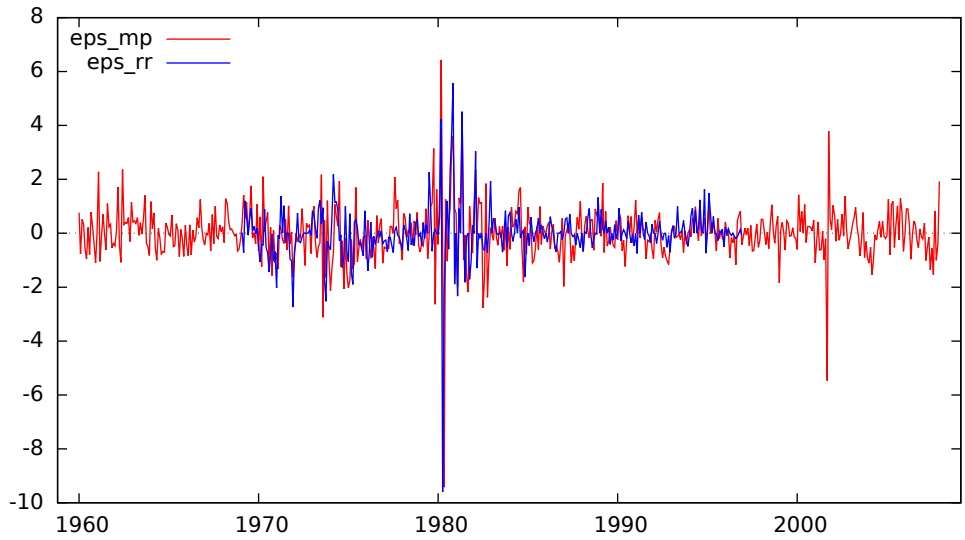

Figure 9: The series $\hat{\varepsilon}_{ts^*}^{mp}$ and $\hat{\varepsilon}_t^{rr}$, divided by their respective standard deviations. Scheme I.

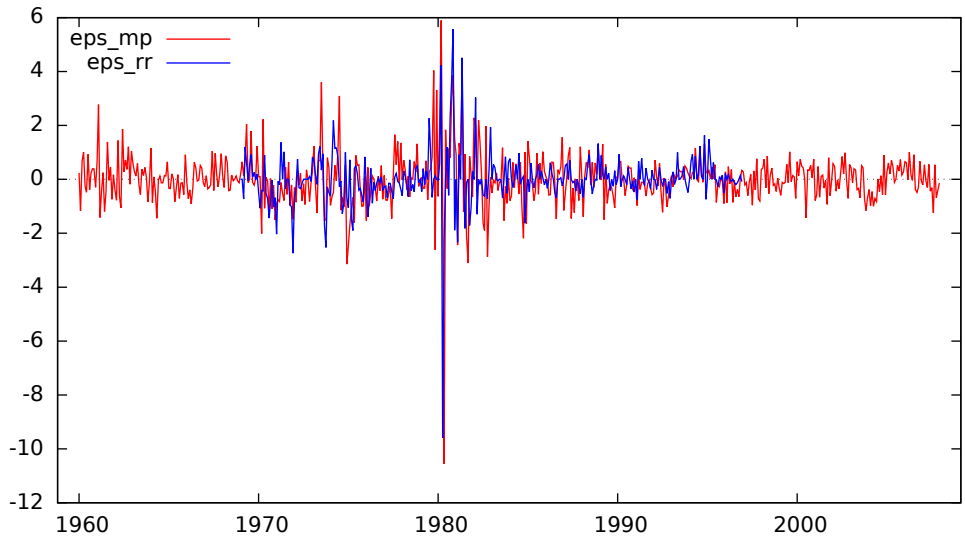

Figure 10: The series $\hat{\varepsilon}_{ts^*}^{mp}$ and $\hat{\varepsilon}_t^{rr}$, divided by their respective standard deviations. Scheme II.

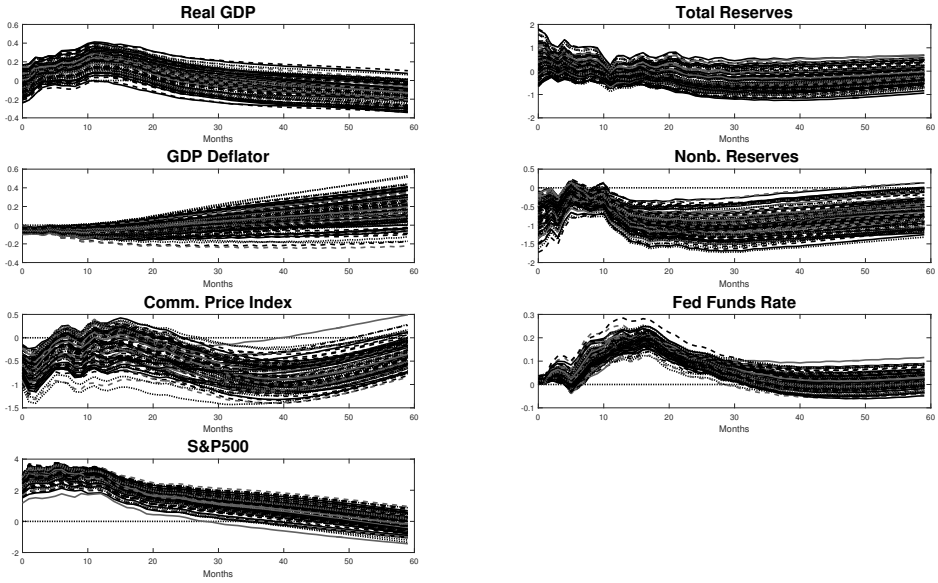

Figure 11: Impulse responses from models $s = 1, \ldots, 100$. Identification Scheme I.

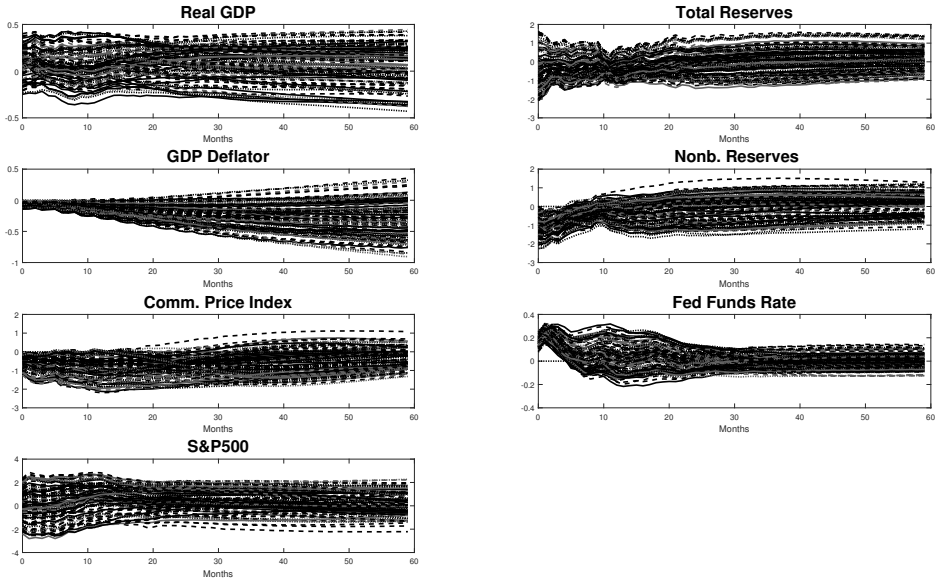

Figure 12: Impulse responses from models $s = 37450, \ldots, 37550$. Identification Scheme I.

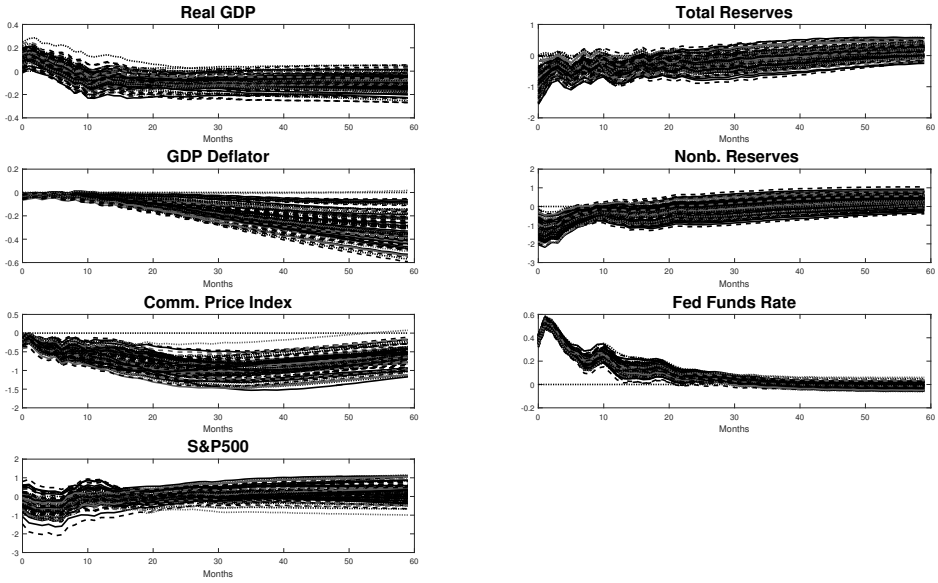

Figure 13: Impulse responses from models $s = 64900, \ldots, 65000$. Identification Scheme I.

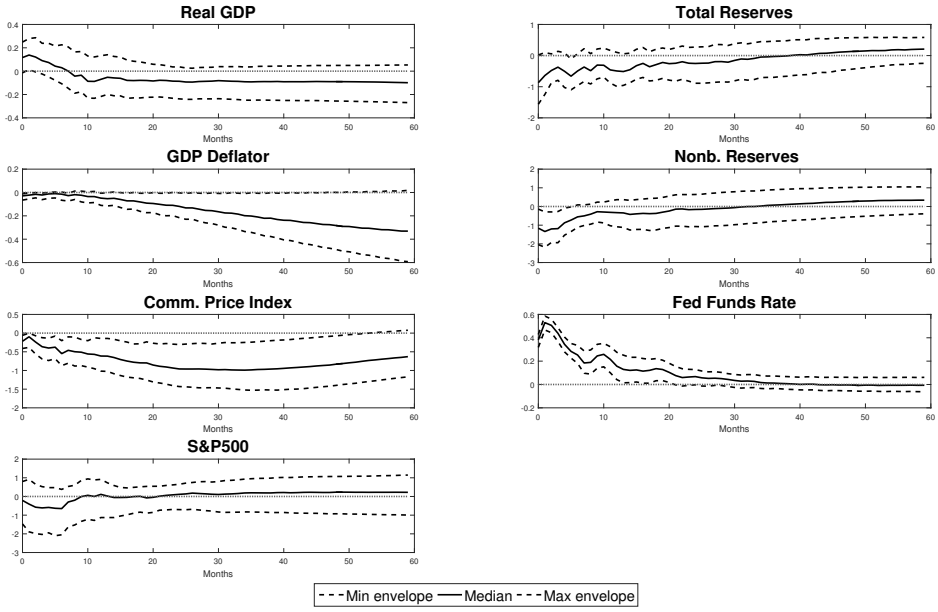

Figure 14: Range and median of models from $\mathcal{S}^{high}$. Identification Scheme I.

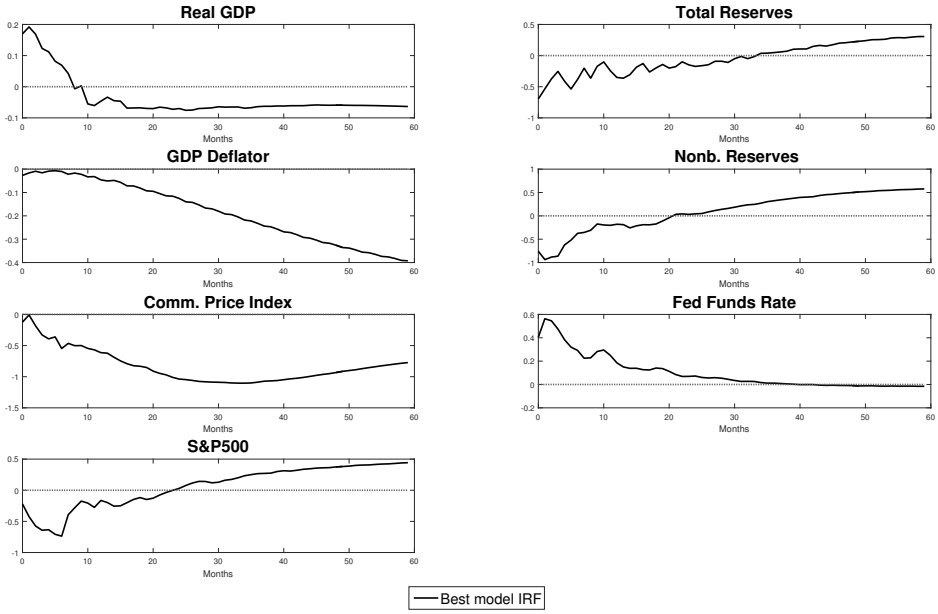

Figure 15: Impulse responses from model s^*. Identification Scheme I.

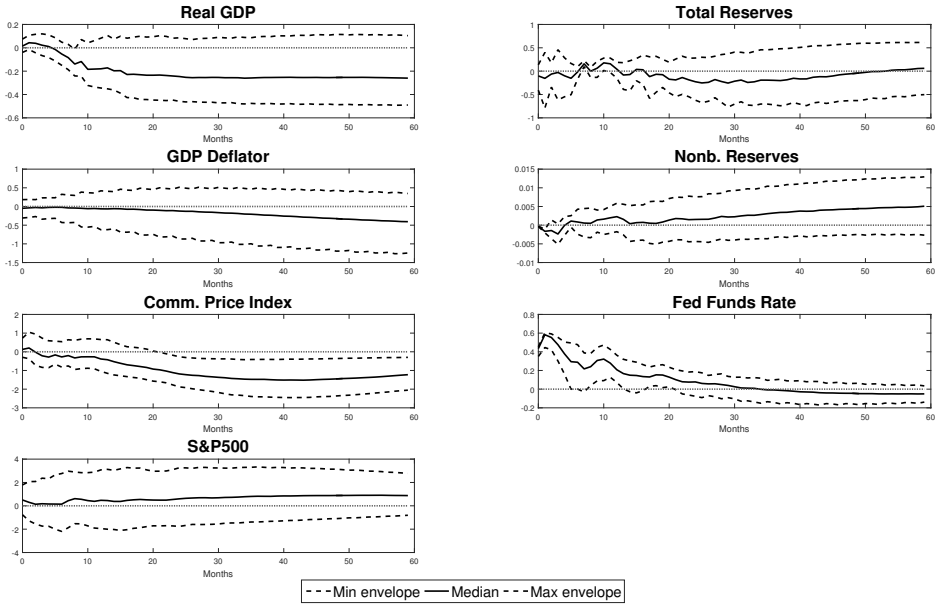

Figure 16: Range and median of models from $\mathcal{S}^{high}$. Identification Scheme II.

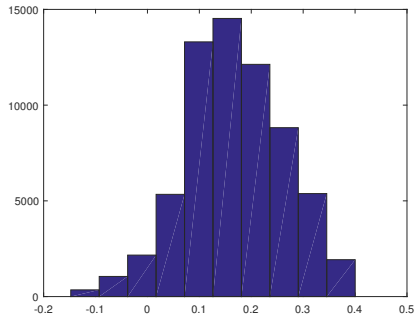

Figure 17: Histogram of correlations between $\hat{\varepsilon}_t^{mp}$ and $\hat{\varepsilon}_t^{rr}$. Scheme II.

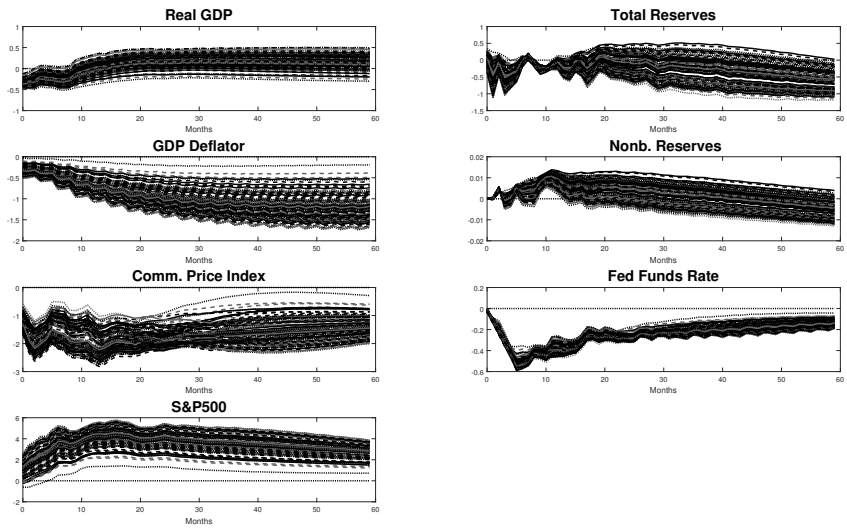

Figure 18: Impulse responses from models $s = 1, \ldots, 100$. Identification Scheme II.

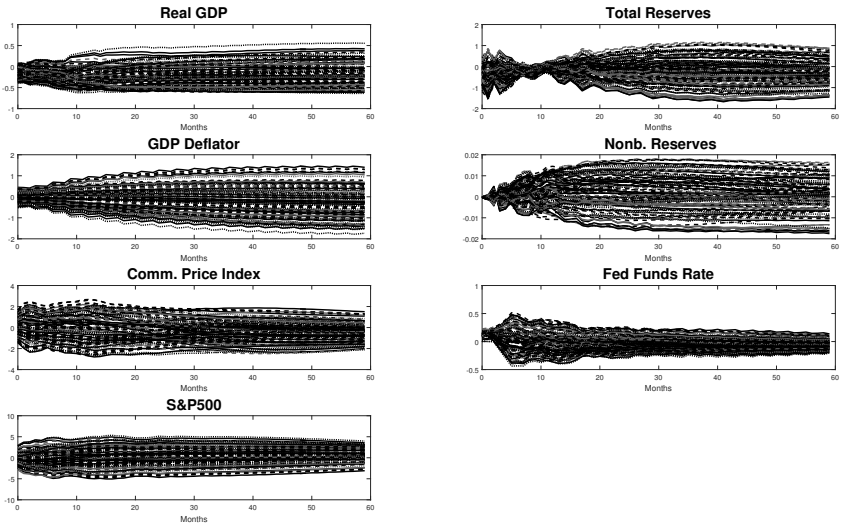

Figure 19: Impulse responses from models $s = 37450, \ldots, 37550$. Identification Scheme II.

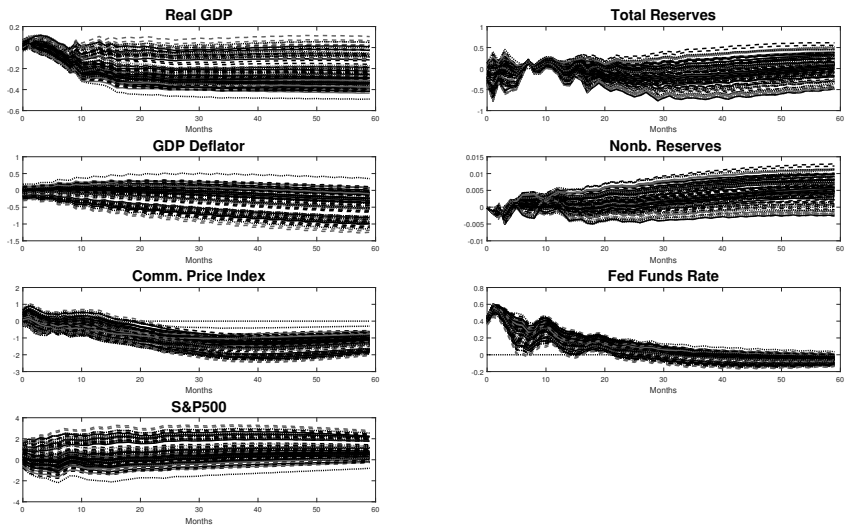

Figure 20: Impulse responses from models $s = 64900, \ldots, 65000$. Identification Scheme II.