Same, but different? Testing monetary policy shock measures

Stephanie Ettmeier[a] Alexander Kriwoluzky[b]

.

16th July 2019

Abstract

In this study, we determine the reliability and exogeneity of four popular monetary policy shock measures, namely the narrative series of Romer and Romer (2004), the high-frequency series of Barakchian and Crowe (2013), the high-frequency series of Gertler and Karadi (2015), and the hybrid series of Miranda-Agrippino and Ricco (2018b). To this end, we employ the Proxy-SVAR model and different empirical diagnostic tools to determine the shock measures' information content. We find that the measure of Miranda-Agrippino and Ricco (2018b), combining the insights from the narrative approach and high-frequency identification, outperforms the other three series.

JEL classification: C12, C32, E32, E52

Keywords: Identification with external instruments, Monetary policy shock measures, Proxy-SVAR

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

[a]DIW Berlin, Mohrenstr. 58, 10117 Berlin, Germany, and Freie Universität Berlin, Garystr. 21, 14195 Berlin, Germany.

[b]Corresponding author: DIW Berlin, Mohrenstr. 58, 10117 Berlin, Germany, and Freie Universität Berlin, Garystr. 21, 14195 Berlin, Germany, akriwoluzky@diw.de, +49 30 89789-430.

1 Introduction

What effect does a monetary policy shock have on the economy? This question is addressed by, among others, Romer and Romer (2004), Barakchian and Crowe (2013), Gertler and Karadi (2015), and Miranda-Agrippino and Ricco (2018b). The common underlying feature of these four studies is that they directly measure monetary policy shocks. All these measures are popular and many studies apply them synonymously. However, as Figure 1 shows, the four measures of these studies differ greatly from one another. One potential reason for this discrepancy highlighted in the literature is that directly measured shock series are not only prone to measurement error, but also incompatible with the same economic concept of a monetary policy shock.

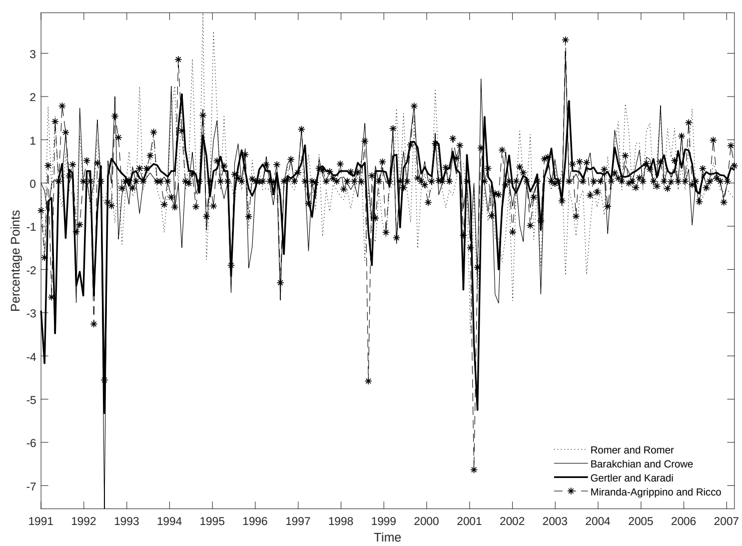

Figure 1. Comparison of standardized monetary policy shock measures for their overlapping period 1991:M1 to 2008:M6. The dotted line depicts Romer and Romer (2004) updated by Barakchian and Crowe (2013), the solid line depicts Barakchian and Crowe (2013), the bold solid line depicts Gertler and Karadi (2015), and the dashed line with stars Miranda-Agrippino and Ricco (2018b). To compute the standardized series, we demeaned the proxies and normalized their variance to 1.

Romer and Romer (2004) choose a narrative approach and study the archives of the Federal Reserve System to control for the information set of the central bank. On the other hand, the measures of Barakchian and Crowe (2013) and Gertler and Karadi (2015) rely on the high-frequency approach popularized by Kuttner (2001) and Gürkaynak et al. (2005), which controls for the information set of market participants, and measure the changes in interest-rate futures on days when the Federal Open Market Committee (FOMC) meets. While Barakchian and Crowe (2013) employ a factor model to condense the shock measure from different futures contracts, Gertler and Karadi (2015) calculate the first differences of a single contract, namely the 3-month-ahead Fed funds futures. The more recent informationally-robust measure of Miranda-Agrippino and Ricco (2018b) can be described as a blend of the narrative approach and high-frequency identification. Combining the central bank's and market participants' information set, it uses a high-frequency measure that is orthogonal to the Greenbook variables employed in Romer and Romer (2004).

For an applied researcher faced with choosing between competing monetary policy shock variables, certainly the two most important decision criteria are sharp inference and correctly identified dynamic causal effects. The first criterion, reliable inference, is linked to the *relevance* of the shock measure, i.e. its degree of correlation with the structural monetary policy shock. The second criterion, well-identified causal effects, is associated with the measure's *exogeneity*, meaning that it should not contain any information apart from the structural monetary policy shock. While it is econometrically impossible to test if a shock measure is orthogonal to other contemporaneous structural shocks, its orthogonality to leads and lags of these other shocks is testable. In this study, we examine the relevance and the lead-lag exogeneity for the monetary policy shock measures of Romer and Romer (2004), Barakchian and Crowe (2013), Gertler and Karadi (2015), and Miranda-Agrippino and Ricco (2018b) in a Proxy-SVAR model. The Proxy-SVAR model of Stock and Watson (2012) and Mertens and Ravn (2013) is an empirical framework for causal inference that applies the shock measures as proxies for the structural shock. We carry out impulse response analysis and employ different empirical diagnostic tools to determine the shock measures' information content. We find that only the shock measure that combines the insights from the narrative and the high-frequency approach passes the relevance and the lead-lag exogeneity test.

2 Testing monetary policy shock measures

In this section, we describe how the monetary policy shock measures identify the structural monetary policy shocks. As the structural shocks are unobserved, assessing their properties is not directly possible. However, studying the properties of the monetary policy shock measures it is possible to have some indirect inference on them.

2.1 The empirical model and identification

For our analysis, we employ the Proxy-SVAR model developed by Stock and Watson (2012) and Mertens and Ravn (2013). Generally, the vector autoregression (VAR) model expresses the observables y_t as projection on its past values and a reduced-form innovation u_t:

$$B(L)y_t = u_t, \tag{1}$$

where B denotes a coefficient matrix and L stands for the lag operator. y_t includes log industrial production (IP), the log consumer price index (CPI), the excess bond premium (EBP) of Gilchrist and Zakrajšek (2012), and as monetary policy indicator the federal funds shadow rate (FF^*) of Wu and Xia (2016).[1] We choose a lag order of 12. Similar to Stock and Watson (2018), we estimate the model over the sample time span from 1979:M7 to 2012:M6, but for identification, we consider the time span for which the four monetary policy shock series are jointly available; that is, from 1991:M1 to 2008:M6. In the Proxy-SVAR approach, the structural monetary policy shock can be identified by regressing the proxy variable s_t^* (here s_t^{RR}, s_t^{BC}, s_t^{GK}, or s_t^{MAR}) on the set of reduced-form shocks u_t.

[1] Appendix A provides a data overview. Like Stock and Watson (2018), we verified that our system is invertible. Appendix B contains the test results. Alternative policy indicators used in the literature are the one-year, or the two-year nominal government bond rate, respectively. These VAR model specifications do not pass the invertibility test. In Appendix C we show the results.

2.2 Relevance

In a first step, we test the strength of the proxy variables by computing the F-statistic from the linear projection of the reduced-form shock related to monetary policy u_t^{MP} on s_t^*. Table 1 presents the test results. The measures of Romer and Romer (2004), Gertler and Karadi (2015) and Miranda-Agrippino and Ricco (2018b) allow for rejection of the null hypothesis of weak relevance.

Table 1
Relevance test

	RR	**BC**	**GK**	**MAR**
F	14.3*	9.38	31.7*	18.47*

Note: * indicates significance according to the rule of thumb $F > 10$.

The test of relevance is instructive, but deserves a qualifying note. In case the measures are contaminated by other macroeconomic shocks, the F-statistics may increase. In the next step, we therefore analyze the four shock measures' exogeneity.

2.3 Exogeneity

Next, we examine the four shock measures' lead-lag exogeneity. Miranda-Agrippino and Ricco (2018a) show that a proxy that is orthogonal to leads and lags of the other shocks in the VAR system delivers stable impact responses, independent of whether the model is misspecified or not. On the contrary, a proxy that does not satisfy the lead-lag exogeneity condition gives unstable impact responses. We use this finding and compare the four shock measures' impact responses of our invertible system specified above with the responses of a trivariate VAR with two lags that includes FF^*, log IP, and log CPI.[2]

Figure 2 compares the impact responses. We benchmark our results against the impact responses of the monetary policy shock (SW) implied by the DSGE model of Smets and

[2]Caldara and Herbst (2019) show that monetary policy VARs without a measure for credit costs are misspecified.

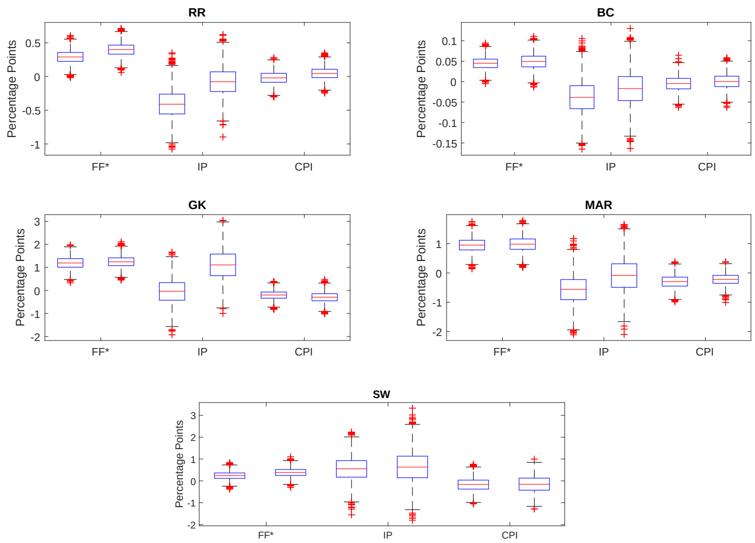

Figure 2. Impact effects of monetary policy proxies. For each variable, the boxplot on the left shows the results of the baseline VAR model with 12 lags. Confidence bands were obtained by a wild bootstrap.

Wouters (2007). It is by definition uncorrelated with other structural shocks and hence suitable to clarify the notion of stability in this context. The series of Barakchian and Crowe (2013), and Miranda-Agrippino and Ricco (2018b) produce impact responses over both models which resemble the responses of the DSGE model shock. The results obtained by the series of Romer and Romer (2004) and Gertler and Karadi (2015) vary considerably, suggesting that the lead-lag exogeneity condition is not satisfied, a point already raised by Ramey (2016) and Miranda-Agrippino and Ricco (2018b).

2.4 Dynamic effects of monetary policy

Last, we compare the impulse responses of a contractionary one-unit monetary policy shock as identified by the four shock measures (Figure 3). Although subtle, differences do exist. The measure of Miranda-Agrippino and Ricco (2018b) is the only series that produces significant impact effects that are in line with the theoretical predictions.

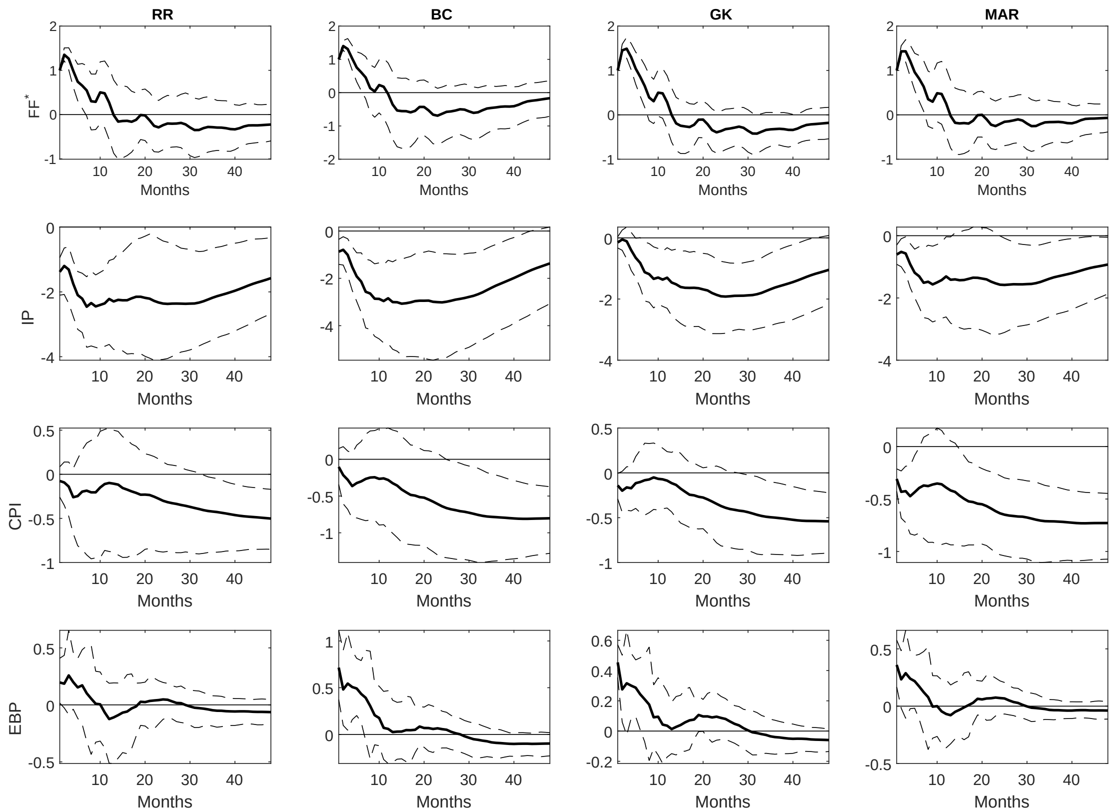

Figure 3. Dynamic effects of a monetary policy shock identified by the shock measures of Romer and Romer (2004), Barakchian and Crowe (2013), Gertler and Karadi (2015), and Miranda-Agrippino and Ricco (2018b). Inference was made using the method of Montiel Olea et al. (2016) that is robust to weak proxy variables. The bold line depicts the mean estimate and the dashed lines show 68% confidence bands.

3 Conclusion

In this study, we examine the relevance and lead-lag exogeneity of the popular monetary policy shock series of Romer and Romer (2004), Barakchian and Crowe (2013), Gertler and Karadi (2015), and Miranda-Agrippino and Ricco (2018b). Our results show that only the shock proxy of Miranda-Agrippino and Ricco (2018b) satisfies the relevance *and* the lead-lag exogeneity conditions. Moreover, it is the only series that produces significant impact effects of monetary policy. This draws us to the conclusion that combining the insights from the narrative approach and the high-frequency identification, as an increasing number of studies in the literature do, is a promising route to proceed.

Acknowledgements

We thank the Editor Pierre-Daniel Sarte, and the Associate Editor Eric Young. We also thank Martin Altemeyer-Bartscher, Georgios Georgiadis, Oliver Holtemöller, Kurt Lunsford, Chi Hyun Kim, Martin Kliem, Gernot Müller, Giovanni Ricco, and Moritz Schularick, and conference participants at the Nordic Econometric Meeting 2017, the annual conference of the International Association for Applied Econometrics 2017, and the Theories and Methods in Macroeconomics conference 2019.

References

BARAKCHIAN, S. M. AND C. CROWE (2013): 'Monetary policy matters: Evidence from new shocks data,' *Journal of Monetary Economics*, 60, 950–966.

CALDARA, D. AND E. HERBST (2019): 'Monetary Policy, Real Activity, and Credit Spreads: Evidence from Bayesian Proxy SVARs,' *American Economic Journal: Macroeconomics*, 11, 157–92.

GERTLER, M. AND P. KARADI (2015): 'Monetary policy surprises, credit costs, and economic activity,' *American Economic Journal: Macroeconomics*, 7, 44–76.

GILCHRIST, S. AND E. ZAKRAJŠEK (2012): 'Credit Spreads and Business Cycle Fluctuations,' *American Economic Review*, 102, 1692–1720.

GÜRKAYNAK, R. S., B. P. SACK, AND E. T. SWANSON (2005): 'Do actions speak louder than words? The response of asset prices to monetary policy actions and statements,' *International Journal of Central Banking*, 1, 55–93.

JORDA, O., M. SCHULARICK, AND A. M. TAYLOR (2015): 'Betting the house,' *Journal of International Economics*, 96, 2–18.

KUTTNER, K. N. (2001): 'Monetary policy surprises and interest rates: Evidence from the Fed funds futures market,' *Journal of Monetary Economics*, 47, 523–544.

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, 1212–1247.

MIRANDA-AGRIPPINO, S. AND G. RICCO (2018a): 'Identification with external instruments in structural VARs under partial invertibility,' Sciences Po publications 24, Sciences Po.

——— (2018b): 'The Transmission of Monetary Policy Shocks,' CEPR Discussion Papers 13396, C.E.P.R. Discussion Papers.

MONTIEL OLEA, J. L., J. H. STOCK, AND M. W. WATSON (2016): 'Inference in SVARs identified with External Instruments,' Manuscript.

RAMEY, V. A. (2016): 'Macroeconomic Shocks and Their Propagation,' in *Handbook of Macroeconomics*, ed. by J. B. Taylor and H. Uhlig, Elsevier, vol. 2, chap. 2, 71 – 162.

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

SMETS, F. AND R. WOUTERS (2007): 'Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach,' *American Economic Review*, 97, 586–606.

STOCK, J. H. AND M. W. WATSON (2012): 'Disentangling the Channels of the 2007-09 Recession,' *Brookings Papers on Economic Activity*, 44, 81–156.

——— (2018): 'Identification and Estimation of Dynamic Causal Effects in Macroeconomics Using External Instruments,' *The Economic Journal*, 128, 917–948.

Wu, J. C. and F. D. Xia (2016): 'Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound,' *Journal of Money, Credit and Banking*, 48, 253–291.

Appendix A Data description

The frequency of all data is monthly.

Barakchian and Crowe (2013) Shock Measure: The high-frequency monetary policy shock measure is downloadable as *shock*.
http://dx.doi.org/10.1016/j.jmoneco.2013.09.006 (01/30/2017).

Consumer Price Index: The consumer price index for all urban consumers and all items is seasonally adjusted, chained (reference year 1982–1984), and log transformed. This series is downloadable from the Federal Reserve Bank of St. Louis as *CPIAUCSL*.
https://fred.stlouisfed.org/series/CPIAUCSL (01/30/2017).

Excess Bond Premium: The excess bond premium is originally provided by Gilchrist and Zakrajŝek (2012). Monthly updates are made available by the Board of Governors of the Federal Reserve System. The series is downloadable as *ebp*.
https://www.federalreserve.gov/econresdata/notes/feds-notes/2016/updating-the-recession-risk-and-the-excess-bond-premium-20161006.html (01/10/2017).

Federal Funds Shadow Rate: The federal funds shadow rate is originally provided by Wu and Xia (2016).
https://sites.google.com/view/jingcynthiawu/shadow-rates (01/30/2017).

Gertler and Karadi (2015) Shock Measure: The high-frequency monetary policy shock measure is provided by the authors as *ff4*.
https://www.aeaweb.org/aej/mac/data/0701/2013-0329_data.zip (01/30/2017).

Industrial Production Index: The industrial production index is taken from the Federal Reserve Bank of St. Louis. The series *INDPRO* is seasonally adjusted and chained (reference year 2012). It is used in its log transformation.
https://fred.stlouisfed.org/series/INDPRO (01/30/2017).

Miranda-Agrippino and Ricco (2018b) Shock Measure: The hybrid monetary policy shock measure is provided by the authors as *MM_IV1*.
http://silviamirandaagrippino.com/s/Instruments_web.xlsx (03/11/2019).

One-Year Government Bond Rate: This interest rate measure is the 1-Year Treasury Constant Maturity Rate from the Federal Reserve Bank of St. Louis. The series is not seasonally adjusted, measured in percent and available as *GS1*.
https://fred.stlouisfed.org/series/GS1 (01/30/2017).

Romer and Romer (2004) Shock Measure: The original narrative monetary policy shock measure from 1969 to 1996 is provided by Romer and Romer (2004). Monthly updates through June 2008 are made available by Barakchian and Crowe (2013) and can be retrieved as *resid08*.
http://dx.doi.org/10.1016/j.jmoneco.2013.09.006 (01/30/2017).

Smets and Wouters (2007) Monetary Policy Shock: The monetary policy shock is the estimate of the structural shock derived from the Kalman smoother at the model's posterior mean. Code and data for estimating the model can be found at
https://www.aeaweb.org/aer/data/june07/20041254_data.zip (03/11/2019).

Two-Year Government Bond Rate: This interest rate measure is the 2-Year Treasury Constant Maturity Rate from the Federal Reserve Bank of St. Louis. The series is not seasonally adjusted, measured in percent and available as *GS2*.
https://fred.stlouisfed.org/series/GS2 (01/30/2017).

Appendix B Invertibility test

To examine if the VAR model with four variables (FF^*, log IP, log CPI, EBP) and 12 lags is able to recover the structural monetary policy shock, we follow Stock and Watson (2018) and test the invertibility of the system. We do this by comparing the identified structural estimates of the Proxy-SVAR with the directly obtained estimates of a local projection estimation with proxy variables (LP-IV) as applied in Jorda et al. (2015). In LP-IV, the policy indicator (FF^*) is instrumented with the proxy variable and regressed on log IP, log CPI, and EBP to estimate the dynamic causal effects of monetary policy.

If the spanning condition holds, the Proxy-SVAR and LP-IV deliver consistent results, but the Proxy-SVAR is more efficient. If, however, the spanning condition does not hold, only the results from LP-IV are consistent. Hence, comparing the Proxy-SVAR and LP-IV estimators provides a Hausman-type test of the null hypothesis of invertibility. In Table B1, we show the p-values for the test of the null hypothesis that the Proxy-SVAR's and IV-LP's causal effects are the same for different horizons ($h = 0, 6, 12, 24$). The p-values are obtained by a parametric bootstrap as in Stock and Watson (2018). For all shock proxies, we cannot reject the null hypothesis of invertibility.

Table B1
VAR(4) model invertibility test

	FF^*	**IP**	**CPI**	**EBP**
RR	0.49	0.66	0.77	0.64
BC	0.96	0.47	0.34	0.76
GK	0.93	0.96	0.59	0.56
MAR	0.47	0.44	0.43	0.23

Note: The table shows p-values for the test of invertibility (H_0 : invertibiliy).

Appendix C Choice of policy indicator

Besides the federal funds rate, the one-year nominal government bond rate (1Y) and the two-year nominal government bond rate (2Y) are used as alternative policy indicators in the literature. To guide our choice of the policy indicator, we checked the invertibility of the VAR system when using one of the alternatives. Table C1 and Table C2 contain the results.

Table C1
VAR(4) model invertibility test

	1Y	**IP**	**CPI**	**EBP**
RR	0.03	0.01	0.18	0
BC	0	0	0	0
GK	0.67	0.87	0.88	0.75
MAR	0.43	0.01	0.12	0.01

Note: The table shows p-values for the test of invertibility (H_0 : invertibiliy).

Table C2
VAR(4) model invertibility test

	2Y	**IP**	**CPI**	**EBP**
RR	0.59	0	0.05	0
BC	0	0	0	0
GK	0.73	0.97	0.95	0.64
MAR	0.6	0.23	0.74	0.08

Note: The table shows p-values for the test of invertibility (H_0 : invertibiliy).