Same, but different: testing monetary policy shock measures*

Stephanie Ettmeier† Alexander Kriwoluzky‡

January 31, 2017

Abstract

In this paper we test whether three popular measures for monetary policy, i.e. Romer and Romer (2004), Barakchian and Crowe (2013), and Gertler and Karadi (2015), constitute good proxy variables of monetary policy shocks. To this end, we employ a test statistic which has been developed especially for weak proxy variables (Lunsford, 2015). We find that the measure derived by Gertler and Karadi (2015) constitutes the best measure of monetary policy shocks.

JEL classification: C12, C32, E32, E52
Keywords: Monetary policy shock measures, Proxy-SVAR, Weak proxies, F-test

*We would like to thank Martin Altemeyer-Bartscher and Chi Hyun Kim.

†Martin-Luther-Universität Halle-Wittenberg

‡Martin-Luther-Universität Halle-Wittenberg and Halle Institute for Economic Research (IWH) Halle, Germany, email: alexander.kriwoluzky@wiwi.uni-halle.de, tel: +49 345 55 23358

1 Introduction

What is the effect of a monetary policy shock on the economy? Recently, this question has been addressed by Romer and Romer (2004), Barakchian and Crowe (2013), and Gertler and Karadi (2015) using a two-step approach. First, the authors obtain a measure of the monetary policy shock. Romer and Romer (2004) study the archives of the Federal Reserve system. Barakchian and Crowe (2013) as well as Gertler and Karadi (2015) measure the changes in interest rates and interest-rate futures on days the Federal Open Market Committee (FOMC) meets. The former study employs a factor analysis to construct the monetary policy shock measure from interest-rate futures with different maturities, the latter measures the monetary policy shock using one particular interest rate. Interestingly, as Figure 1 shows, while the three studies measure the same structural shock, these measures are very different from each other.

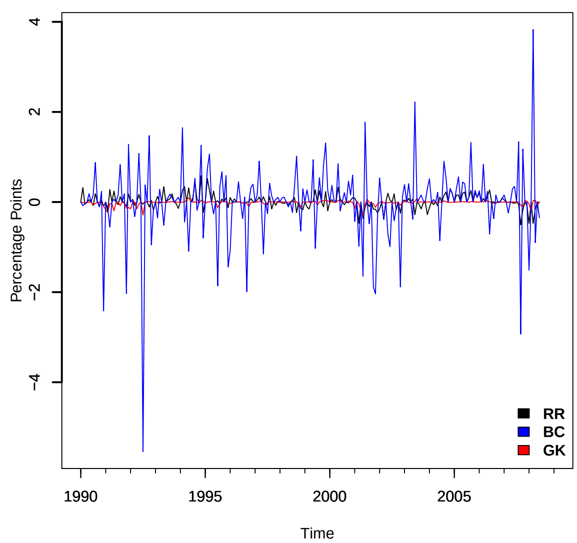

Figure 1: Comparison of different monetary policy shocks measures: the black line depicts Romer and Romer (2004) (RR), the blue line Barakchian and Crowe (2013) (BC), and the red line Gertler and Karadi (2015) (GK).

In a second step, the monetary policy shock measures are used directly (Romer and Romer, 2004; Barakchian and Crowe, 2013) or as a proxy variable of monetary policy shocks (Gertler and Karadi, 2015) to analyze the effects on the variable of interest. This approach has become popular and many studies have recently been applying it.[1] Yet, the quality and

[1] Examples include: Stock and Watson (2012), Midrigan (2011), Miranda-Agrippino and Rey (2015), Piffer (2016), Sinclair, Tien, and Gamber (2016), Zeev, Gunn, and Khan (2015), Georgiadis (2015), Balakrishnan,

accurateness of the estimation results depends on the quality of the measure of the monetary policy shocks. In this paper, we determine which of the three measures should be employed when studying the effects of monetary policy on the economy.

Our analysis is based on a F-test for proxy variables developed by Lunsford (2015) and Stock and Watson (2012), which tests whether a proxy variable is a weak proxy. In contrast to other test statistics this test statistic does not depend on the estimated parameters characterising the relationship between structural errors and the reduced form errors. This allows for precise determination of the critical values of the test statistic. Applying this test, we find that the monetary policy shock measure derived by Gertler and Karadi (2015) is the only measure for which we can reject the hypothesis of being a weak proxy variable on a 1% significance level.

The paper is structured in the following way. Section 2 describes the different measures of monetary policy shocks and shows how much they differ from each other. In Section 3, we first set up our empirical model and the test statistic, before reporting the results of our analysis. The final section concludes.

2 Three different monetary policy shock measures

In this section we briefly describe the three measures of monetary policy shocks which we consider in this paper. We start by illustrating the narrative approach employed by Romer and Romer (2004) and outline the approach by Barakchian and Crowe (2013) and Gertler and Karadi (2015) afterwards.

Romer and Romer (2004) construct their measure of a monetary policy shock by carefully studying the archives of the Federal Reserve. First, the authors use the *Record of Policy Actions of the FOMC*, the *Minutes of the FOMC* and the *Monetary Policy Alternatives* (*Bluebook*) to determine the intended Federal Funds Rate ($\tilde{i}$). Furthermore, Romer and Romer (2004) specify the information set of the FOMC by using the *Greenbook*, i.e. the report on *Current Economic and Financial conditions*. From this source they derive the forecasts of the growth rate of real GDP (x^f), the inflation (p^f), and the unemployment rate (u^f). In a final step, the authors determine the measure of the monetary policy shock (s^{RR})

Laseen, and Pescatori (2016), and Passari and Rey (2015).

as the residual in the following regression:

$$\begin{aligned}\Delta\tilde{i}_t = \beta_0 + \beta_1\tilde{i}_t + \sum_{h=-1}^{2}\beta_{2h}\Delta x^f_{th} + \sum_{h=-1}^{2}\beta_{3h}(\Delta x^f_{th} - \Delta x^f_{t-1,h}) \\ + \sum_{h=-1}^{2}\beta_{4h}p^f_{th} + \sum_{h=-1}^{2}\beta_{5h}(p^f_{th} - p^f_{t-1,h}) + \beta_6 u^f_{t0} + s^{RR}_t\end{aligned} \tag{1}$$

Barakchian and Crowe (2013) and Gertler and Karadi (2015) follow in their approach Söderström (2001), Kuttner (2001), and Faust, Swanson, and Wright (2004). They employ high-frequency data on interest-rate futures to identify a monetary policy innovation. Denote the interest-rate futures traded at date d in month m for the Federal Funds rate in month $m+h$ by f^h_d. Under the assumptions that first, the errors made by targeting the Federal Funds rate are small and typically mean zero, and second, that the risk premium stays constant on the day of the policy announcement, the change in the expected Federal Funds target rate ($E_d\bar{i}_{m+h}$) during the subsequent calendar months ($h \geq 1$) following a policy announcement on day d of month m is given by:

$$\Delta E_d\bar{i}_{m+h} = f^h_d - f^h_{d-1} \tag{2}$$

Barakchian and Crowe (2013) compute the change in the expected Federal Funds target rate in the current month and up to 5 months ahead. In a next step, the authors follow Gürkaynak, Sack, and Swanson (2005) and employ a factor model in order to construct the monetary policy shock measure. The factor model is given by:

$$\omega_t = \phi_t\Lambda + e_t \tag{3}$$

where the vector ω_t collects the changes in the expected Federal Funds target rate for different maturities, e_t comprises unique factors for each entry in ω_t, $\phi_t = \begin{bmatrix} \phi_{1,t} & \phi_{2,t} \end{bmatrix}$ contains two factors and the vector Λ the corresponding loadings. Barakchian and Crowe (2013) argue that the first factor $\phi_{1,t}$ measures the monetary policy shocks and thus set $s^{BC}_t = \phi_{1,t}$.

In contrast to Barakchian and Crowe (2013), Gertler and Karadi (2015) do not estimate a factor model but construct the measure directly from the difference between the expected Federal Funds target rates. Furthermore, they do not consider a time-window of 24 hours around FOMC meetings, but measure the monetary policy shocks within a 30 minute window of the policy announcement. As their baseline measure for monetary policy shocks Gertler and Karadi (2015) consider the change in the three-month ahead futures contracts, i.e. $s^{GK}_t = \Delta E_{30min}\bar{i}_{m+3}$. Here, ΔE_{30min} indicates the 30 minute window around the policy announcement.

Figure 1 plots the different measures for monetary policy shocks.[2] Even though the three measures are supposed to measure the same economic concept, they all look strikingly different. Table 1 further underlines the discrepancies. The measures derived using the high-frequency data approach are only weakly correlated with the narrative time series. In fact, even the correlation between the measures which both have been constructed from futures on interest rate data is only around 0.4. Table 1 further shows that the measure by Barakchian and Crowe (2013) exhibits the highest amplitude and correspondingly the highest standard deviation. The amplitude of the shock measure constructed by Romer and Romer (2004) lies between the two high-frequency measures. The monetary shocks of Gertler and Karadi (2015) exhibit the smallest variance.

While these shock measures are different, they are all subject to criticism. For example, Leeper (1997) and Coibion (2012) suggest that the narrative time series could still contain endogenous components. Furthermore, Ramey (2016) points out that the measure constructed by Gertler and Karadi (2015) can be predicted by Greenbook forecasts. Finally, Ramey (2016) worries that the VAR model is misspecified in the sense that the information set of the VAR model is incomplete. This would imply that the underlying moving average process is not invertible and therefore cannot be approximated by a VAR model. Despite these criticisms, we will pursue in our analysis with the original measures of monetary policy shocks and determine which of these constitutes a good proxy variable for monetary policy shocks for two reasons. First, as shown by the extensive list of references in the introduction the monetary policy shock measures are nevertheless employed in applied research. Second, given that the approaches yield such different measures, it is of interest which approach should be further worked on and improved.[3]

	RR	**BC**	**GK**
RR	1.000		
BC	0.193	1.000	
GK	0.190	0.405	1.000
T	222	222	222
Standard deviation	0.150	0.805	0.052
Maximum value	0.584 (Nov 1994)	3.831 (Mar 2008)	0.092 (May 1995)
Minimum value	-0.505 (Sep 2007)	-5.540 (Jul 1992)	-0.290 (Jul 1992)

Table 1: Correlation and summary statistics

[2]The measures have all been obtained from the supplementary archives of the journals. Appendix A provides further information on the dataset.

[3]First improvement approaches are pursued for instance by Miranda-Agrippino (2016) and Paul (2015).

3 Testing monetary policy shock measures

We start this section by describing the data. In a next step, we set up the empirical model and the test statistic, then we report the results.

3.1 The empirical model

Throughout the paper we consider the time span for which all three series are available, i.e. 1990:M1-2008:M6. In our analysis, we employ the Proxy-SVAR model, which has been developed by Stock and Watson (2012) and Mertens and Ravn (2013). The VAR model contains the endogenous variables which Romer and Romer (2004), Barakchian and Crowe (2013), and Gertler and Karadi (2015) have in common: one measure for GDP, one for inflation, and one for monetary policy. The variables are collected in a vector (y_t):

$$y_t = \begin{bmatrix} i_t & x_t & p_t \end{bmatrix}' , \tag{4}$$

where i_t denotes the Federal Funds rate, x_t the logarithm of industrial production, and p_t the logarithm of the consumer price index. As an robustness exercise, we also consider the producer price index instead of the consumer price index and the interest rate on one-year government bonds as an interest rate measure.[4]

In general, the VAR model with n endogenous variables is given by:

$$y_t = B_0 + B_1 t + B(L) y_{t-1} + u_t, \quad u_t \sim \mathcal{N}(0, \Sigma_u) , \tag{5}$$

where $B(L)$ denotes the reduced form VAR model coefficients, B_1 a time trend, and B_0 the intercept term. u_t denotes the $n \times 1$ vector of reduced form errors with the corresponding variance-covariance matrix Σ_u. The reduced form errors u_t are related to the structural errors ϵ_t as follows:

$$u_t = A\epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0, I) . \tag{6}$$

The identification issue in VAR models arises because it is not possible to determine A uniquely from $\Sigma_u = AA'$. The Proxy-SVAR model estimates the effects of a structural shock by using an additional measure for monetary policy as proxy for the underlying structural shock. Denote the column in A related to the monetary policy shock by a_m. Furthermore,

[4] Appendix B provides the results of the robustness exercise.

partition the structural innovations as follows:

$$\underset{n\times 1}{\epsilon_t} = \begin{bmatrix} \underset{1\times 1}{\epsilon_{m,t}} & \underset{n-1\times 1}{\epsilon_{2,t}'} \end{bmatrix}' \tag{7}$$

The first block ($\epsilon_{m,t}$) is associated with the monetary policy shock, the second block comprises the additional shocks. The three different monetary policy shock measures are subsequently assumed to be a proxy variable which is correlated with the monetary policy shock $\epsilon_{m,t}$:

$$E[s_t^* \epsilon_{m,t}] = \Phi, \quad \Phi \neq 0, \tag{8}$$

and uncorrelated with the remaining structural shocks:

$$E[s_t^* \epsilon_{2,t}] = 0. \tag{9}$$

Here, s_t^* stands for either s^{RR}, s^{GK}, or s^{BC}. In order to test whether the proxy variable constitutes a weak proxy variable we employ the test statistic suggested by Lunsford (2015) and Stock and Watson (2012). The test statistic is based on linearly projecting s_t^* on the reduced form errors and a constant:

$$s_t^* = c_0 + u_t' \mu + \nu_t, \tag{10}$$

where μ is a vector of dimension $n \times 1$ and ν_t an *i.i.d.* error term with variance σ_ν^2. Furthermore, ν_t is assumed to be independent of lags of y_t and independent of ϵ_t. To test whether s^* constitutes a weak proxy variable for ϵ_m Lunsford (2015) suggests a standard F-test statistic.[5] This approach of Lunsford (2015) has the advantage that the critical values of the test statistic do not depend on an estimate of A. In case the proxy variables is a weak proxy, the estimate of A will be imprecise as well and consequently dilute the analysis.

The application of the test demands that the researcher determines the level of statistical significance as well as the level of asymptotic bias. The asymptotic bias arises because the estimator of a_m is not consistent. The corresponding studies in the literature, i.e. Lunsford (2015), Stock and Yogo (2005), and Stock, Wright, and Yogo (2002), do not provide an optimal level of bias tolerance. Instead these studies employ a 10% bias in their applications. Therefore, we also employ an asymptotic bias level of 10%.

[5] Appendix C provides further details on the test for weak proxies.

3.2 Results

In order to obtain a precise estimate of the reduced form coefficients we follow Gertler and Karadi (2015) and estimate the VAR model on the sample 1979:M7-2012:M6. Once we have obtained an estimate for the VAR model parameters, i.e. B_0, B_1, $B(L)$ and Σ_u, and the corresponding reduced form errors u_t, we compute the F-statistic related to the test $\mu = 0$ in Equation (10). To determine whether we can reject the null hypothesis that $\mu = 0$ we use the critical values provided by Lunsford (2015). Table 2 displays the estimation results. We find that the shock measure of Barakchian and Crowe (2013) is a weak proxy variable. Furthermore, regarding the monetary policy shock measures constructed by Romer and Romer (2004), we can reject the null hypothesis with a significance level of 10%. This implies that these shocks are still considered relatively weak proxy variables. The best proxy variable for monetary policy shocks is the monetary policy shock measure provided by Gertler and Karadi (2015). We can reject the null hypothesis with a significance level of 1%.

Table 2: Results

	RR	**BC**	**GK**
F_{WI}	8.134*	3.879	11.901***

Note: * significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level

4 Conclusion

In this paper, we have analysed which monetary policy shock measure constitutes the best proxy variable to analyse the effects of monetary policy. Our results are based on a recently developed test statistic which allows to test whether the shock measures are a weak proxy independent of the linear relationship between the reduced form errors of the VAR model and the underlying structural shocks. We find that the measure by Gertler and Karadi (2015) constitutes the best proxy variable for monetary policy shocks and should thus preferably be employed studying the effect of monetary policy.

References

Balakrishnan, R., S. Laseen, and A. Pescatori (2016): "U.S. Dollar Dynamics; How Important Are Policy Divergence and FX Risk Premiums?" IMF Working Papers 16/125, International Monetary Fund.

Barakchian, S. M. and C. Crowe (2013): "Monetary policy matters: Evidence from new shocks data," *Journal of Monetary Economics*, 60, 950–966.

Coibion, O. (2012): "Are the effects of monetary policy shocks big or small?" *American Economic Journal: Macroeconomics*, 4, 1–32.

Faust, J., E. T. Swanson, and J. H. Wright (2004): "Identifying VARS based on high frequency futures data," *Journal of Monetary Economics*, 51, 1107–1131.

Georgiadis, G. (2015): "To bi, or not to bi? differences in spillover estimates from bilateral and multilateral multi-country models," Globalization and Monetary Policy Institute Working Paper 256, Federal Reserve Bank of Dallas.

Gertler, M. and P. Karadi (2015): "Monetary policy surprises, credit costs, and economic activity," *American Economic Journal: Macroeconomics*, 7, 44–76.

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.

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

Leeper, E. M. (1997): "Narrative and VAR approaches to monetary policy: Common identification problems," *Journal of Monetary Economics*, 40, 641–657.

Lunsford, K. G. (2015): "Identifying Structural VARs with a Proxy Variable and a Test for a Weak Proxy," Working Paper 1528, Federal Reserve Bank of Cleveland.

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.

Midrigan, V. (2011): "Menu Costs, Multiproduct Firms, and Aggregate Fluctuations," *Econometrica*, 79, 1139–1180.

Miranda-Agrippino, S. (2016): "Unsurprising shocks: information, premia, and the monetary transmission," Bank of England working papers 626, Bank of England.

Miranda-Agrippino, S. and H. Rey (2015): "World Asset Markets and the Global Financial Cycle," Tech. Rep. 21722, National Bureau of Economic Research.

Passari, E. and H. Rey (2015): "Financial Flows and the International Monetary System," *The Economic Journal*, 125, 675–698.

Paul, P. (2015): "The Time Varying Transmission of Monetary Policy Surprises," Unpublished, University of Oxford.

Piffer, M. (2016): "Monetary Policy and Defaults in the US," Discussion Papers of DIW Berlin 1559, DIW Berlin, German Institute for Economic Research.

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.

Sinclair, T., P.-L. Tien, and E. Gamber (2016): "Do Fed Forecast Errors Matter?" Working Papers 2016-14, The George Washington University, Institute for International Economic Policy.

Söderström, U. (2001): "Predicting monetary policy with federal funds futures prices," *Journal of Futures Markets*, 21, 377–391.

Stock, J. H. and M. W. Watson (2012): "Disentangling the Channels of the 2007-09 Recession," *Brookings Papers on Economic Activity*, 44, 81–156.

Stock, J. H., J. H. Wright, and M. Yogo (2002): "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments," *Journal of Business & Economic Statistics*, 20, 518–529.

Stock, J. H. and M. Yogo (2005): "Testing for Weak Instruments in Linear IV Regression," in *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, ed. by D. W. K. Andrews and J. H. Stock, Cambridge University Press, chap. 5, 80–108.

Zeev, N. B., C. M. Gunn, and H. Khan (2015): "Monetary News Shocks," Carleton Economic Papers 15-02, Carleton University, Department of Economics.

A Data description

The frequency of all data used is monthly.

Romer and Romer (2004) Shock: The original narrative monetary policy shock measure from 1969-1996 is provided by Romer and Romer (2004). Monthly updates till 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).

Barakchian and Crowe (2013) Shock: 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).

Gertler and Karadi (2015) Shock: 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).

Federal Funds Rate: The effective federal funds rate is retrieved as *FEDFUNDS* from the Federal Reserve Bank of St. Louis. The series is measured in percent and not seasonally adjusted.
https://fred.stlouisfed.org/series/FEDFUNDS (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 2012. It is used in its log transformation.
https://fred.stlouisfed.org/series/INDPRO (01/30/2017).

Consumer Price Index: The consumer price index for all urban consumers and all items is seasonally adjusted, chained 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).

Producer Price Index: The producer price index by commodity for finished goods is seasonally adjusted, chained 1982, and log transformed. This series can be retrieved from the Federal Reserve Bank of St. Louis as *PPIFGS*.
https://fred.stlouisfed.org/series/PPIFGS (01/30/2017).

1-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).

B Robustness results

Table 3: Results with producer price index

	RR	**BC**	**GK**
F_{WI}	10.102**	4.872	13.491***

Note: * significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level

Table 4: Results with one-year government bond rate

	RR	**BC**	**GK**
F_{WI}	1.786	0.149	6.993

Note: * significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level

C Test for weak proxies

The standard F-statistic for the null hypothesis that $\mu = 0$ in equation (10) is given by:

$$F_{WI} = \left(\frac{T-n}{n}\right) \frac{(s_t^* - \bar{s_t^*})^2 - [(s_t^* - \bar{s_t^*}) - \hat{u}_t{}'\hat{\mu}]^2}{[(s_t^* - \bar{s_t^*}) - \hat{u}_t{}'\hat{\mu}]^2} \tag{C.1}$$

$\bar{s_t^*}$ denotes the mean of the proxy variable.
Critical values are provided by Lunsford (2015). They depend on the tolerated asymptotic bias, the statistical significance level and the number of endogenous variables in the VAR.

Table 5: Critical values for trivariate VAR

	statistical significance		
asymptotic bias	10%	5%	1%
10%	7.38	8.53	10.92