Ettmeier, Stephanie; Kriwuluzky, Alexander

Working Paper

Same, but different: Testing monetary policy shock measures

IWH Discussion Papers, No. 9/2017

Provided in Cooperation with:
Halle Institute for Economic Research (IWH) – Member of the Leibniz Association

Suggested Citation: Ettmeier, Stephanie; Kriwuluzky, Alexander (2017) : Same, but different: Testing monetary policy shock measures, IWH Discussion Papers, No. 9/2017, Leibniz-Institut für Wirtschaftsforschung Halle (IWH), Halle (Saale)

This Version is available at:
http://hdl.handle.net/10419/156656

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
Same, but Different: Testing Monetary Policy Shock Measures

Stephanie Ettmeier, Alexander Kriwoluzky
Authors

Stephanie Ettmeier  
Martin Luther University Halle-Wittenberg  
E-mail: stephanie.ettmeier@wiwi.uni-halle.de  
Tel +49 345 552 3455

Alexander Kriwoluzky  
Martin Luther University Halle-Wittenberg and Halle Institute for Economic Research (IWH) – Member of the Leibniz Association  
E-mail: alexander.kriwoluzky@wiwi.uni-halle.de  
Tel +49 345 552 3358

The responsibility for discussion papers lies solely with the individual authors. The views expressed herein do not necessarily represent those of the IWH. The papers represent preliminary work and are circulated to encourage discussion with the authors. Citation of the discussion papers should account for their provisional character; a revised version may be available directly from the authors.

Comments and suggestions on the methods and results presented are welcome.

IWH Discussion Papers are indexed in RePEc-EconPapers and in ECONIS.

Editor

Halle Institute for Economic Research (IWH) – Member of the Leibniz Association  
Address: Kleine Maerkerstrasse 8  
D-06108 Halle (Saale), Germany  
Postal Address: P.O. Box 11 03 61  
D-06017 Halle (Saale), Germany  
Tel +49 345 7753 60  
Fax +49 345 7753 820  
www.iwh-halle.de  
ISSN 2194-2188
Same, but Different: Testing Monetary Policy Shock Measures*

Abstract

In this study, we test whether three popular measures for monetary policy, that is, Romer and Romer (2004), Barakchian and Crowe (2013), and Gertler and Karadi (2015), constitute suitable proxy variables for monetary policy shocks. To this end, we employ different test statistics used in the literature to detect weak proxy variables. We find that the measure derived by Gertler and Karadi (2015) is the most suitable in this regard.

Keywords: monetary policy shock measures, Proxy-SVAR, weak proxies, F-test

JEL Classification: C12, C32, E32, E52

* We would like to thank Martin Altemeyer-Bartscher, Oliver Holtemöller, Chi Hyun Kim, Martin Kliem, Gernot Müller, and Moritz Schularick.
1 Introduction

What is the effect of a monetary policy shock on the economy? Recently, this question was addressed by Romer and Romer (2004), Barakchian and Crowe (2013), and Gertler and Karadi (2015). In a first step, all of these 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) and Gertler and Karadi (2015) measure the changes in interest rates and interest-rate futures on days when 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, whereas the latter measures this shock using one particular interest rate. 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 for monetary policy shocks (Gertler and Karadi, 2015) to analyze the effects on the variable of interest. These measures have become popular, and many recent studies have applied them. Interestingly, as Figure 1 shows, although the three studies measure the same structural shock, these measures are very different from each other. Since the quality and accuracy of the estimation results depend on the quality of the measure of monetary

\[\text{Figure 1: Comparison of different monetary policy shock measures; the black line depicts Romer and Romer (2004), the blue line Barakchian and Crowe (2013), and the red line Gertler and Karadi (2015).}\]

policy shocks, we determine in this study which of the three measures should be employed when studying the effects of monetary policy on the economy.

Two different tests for weak proxies are used in the literature to evaluate the quality of proxy variables in a VAR model framework. Both are based on an F-statistic that determines whether the correlation between a proxy variable and the structural shock of interest is close to zero, but both start from a different setup. On the one hand, Gertler and Karadi (2015) follow the microeconometric instrumental variable approach and employ the proxy as an independent variable. In particular, they regress the reduced form error related to the monetary policy variable on the proxy and derive the F-statistic from this regression. For the test decision, Gertler and Karadi (2015) rely on the commonly accepted rule of thumb that states that an F-statistic greater than 10 signals proxy strength (Stock and Yogo, 2005). On the other hand, Lunsford (2015) follows Stock and Watson (2012) and applies the proxy as the dependent variable. He computes the F-statistic from the regression of the proxy variable on a constant and the entire vector of reduced form errors. Additionally, he provides exact critical values for this test statistic. In order to determine the quality of the three shock measures, we employ both tests for weak proxies. We find that the F-test that is based on the regression with the proxy as an independent variable rejects the null hypothesis of being a weak proxy for all three shock series. Using the F-test linked to the regression with the proxy as the dependent variable, we find that the monetary policy shock measure derived by Gertler and Karadi (2015) is the only measure for which we can reject the null hypothesis of being a weak proxy.

The remainder of this paper is structured as follows. Section 2 describes the three measures of monetary policy shocks and shows how different they are from each other. In Section 3, we set up our empirical model and the test statistics and then report 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 that we consider in this study. We start by illustrating the narrative approach employed by Romer and Romer (2004) and outline the approaches used 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*, that is, 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 rate ($p^f$), and the unemployment rate ($u^f$). In a final step, the authors determine the measure of the monetary policy shock ($s^{RR}$) as the residual in the following regression:

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

(1)

Barakchian and Crowe (2013) and Gertler and Karadi (2015) follow Söderström (2001), Kuttner (2001), and Faust, Swanson, and Wright (2004) in their approach. They employ high-frequency data on interest-rate futures to identify a monetary policy innovation. We 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 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 ($E_{d\tilde{m}+h}$) during the subsequent calendar months ($h \geq 1$) following a policy announcement on day $d$ of month $m$ is given by

$$
\Delta E_{d\tilde{m}+h} = f^h_d - f^h_{d-1}.
$$

(2)

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

$$
\omega_t = \phi_t \Lambda + e_t,
$$

(3)

where the vector $\omega_t$ collects the changes in the expected federal funds target rate for different maturities, $e_t$ comprises unique factors for each entry in $\omega_t$, $\phi_t = \begin{bmatrix} \phi_{1,t} & \phi_{2,t} \end{bmatrix}$ contains two factors, and the vector $\Lambda$ includes 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 instead 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 rather measure the monetary policy shocks within a 30-minute window around the policy announcement. As their baseline measure for monetary policy shocks, Gertler and Karadi (2015) consider the change in three-month ahead futures contracts, that is, \( s_{GK}^t = \Delta E_{30\text{min}} t_{m+3} \). Here, \( \Delta E_{30\text{min}} \) indicates the 30-minute window around the policy announcement.

Figure 1 plots the different measures for monetary policy shocks in their original and standardized forms.\(^2\) 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 two measures that have been constructed from interest-rate futures data is only around 0.4. Table 1 further shows that the measure by Barakchian and Crowe (2013) exhibits the greatest amplitude and, correspondingly, the largest 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) display the smallest variance.

Although these shock measures are different, they are all subject to criticism. For example, Leeper (1997) and Coibion (2012) suggest that a 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) highlights a concern that the VAR model is misspecified in the sense that the information set of the VAR model is incomplete. This misspecification 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 persist in our analysis of 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, these monetary policy shock measures are nevertheless employed in applied research. Second, given that the approaches yield such different measures, it is of interest which measure should be further developed and improved.\(^3\)

3 Testing monetary policy shock measures

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

\(^2\)The measures have all been obtained from the supplementary archives of the journals. Appendix A provides further information on the dataset.

\(^3\)Initial improvement approaches were pursued by, for instance, Miranda-Agrippino (2016) and Paul (2015).
The empirical model

Throughout this study, we consider the time span for which all three series are jointly available, that is, 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 that 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 \( y_t \):

\[
y_t = \begin{bmatrix} i_t & x_t & p_t \end{bmatrix}',
\]

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. We employ a lag order of 12 as all our data has monthly frequency.

In general, the VAR model with \( n \) endogenous variables is given by

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

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 \( \Sigma_u \). The reduced form errors \( u_t \) are related to the structural errors \( \epsilon_t \) as follows:

\[
u_t = A \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0, I).
\]

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 a proxy for the underlying structural
shock. We denote the column in $A$ related to the monetary policy shock as $a_m$. Furthermore, we partition the structural innovations and the reduced form errors as follows:

\[
\begin{align*}
\epsilon_t &= \begin{bmatrix} \epsilon_{m,t} \\ \epsilon_{2,t} \end{bmatrix}_{n \times 1} \\
u_t &= \begin{bmatrix} u_{m,t} \\ u_{2,t} \end{bmatrix}_{n \times 1}
\end{align*}
\]

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

\[
E[s^*_t \epsilon_{m,t}] = \Phi, \quad \Phi \neq 0,
\]

and uncorrelated with the remaining structural shocks:

\[
E[s^*_t \epsilon_{2,t}] = 0.
\]

Here, $s^*_t$ stands for either $s^{RR}$, $s^{GK}$, or $s^{BC}$. In order to determine whether the proxy variable constitutes a weak proxy, we employ two different statistical tests. Our first test is the F-test applied by Gertler and Karadi (2015). Following the microeconometric instrumental variable approach, Gertler and Karadi (2015) use the proxy as an independent variable to derive the test statistic from the linear projection of the reduced form error associated with monetary policy on $s^*_t$:

\[
u_{m,t} = \gamma (s^*_t - \bar{s}^*_t) + \eta_t.
\]

$\bar{s}^*_t$ denotes the mean of the proxy variable, and $\eta_t$ is an i.i.d. error term with variance $\sigma^2_{\eta_t}$. It is assumed to be independent of $\epsilon_t$ and lags of $y_t$. To determine whether $s^*$ is sufficiently correlated with $\epsilon_m$, Gertler and Karadi (2015) compute the F-statistic of Equation (11) for the null hypothesis that $\gamma = 0$:

\[
F_{IV} = (T - 1) \frac{\hat{u}^2_{m,t} - [\hat{u}_{m,t} - \hat{\gamma} (s^*_t - \bar{s}^*_t)]^2}{[\hat{u}_{m,t} - \hat{\gamma} (s^*_t - \bar{s}^*_t)]^2}.
\]

For the test decision, they rely on the commonly accepted rule of thumb that states that an F-statistic greater than 10 signals proxy strength (Stock and Yogo, 2005).

Our second test is the F-test suggested by Lunsford (2015). In contrast to Gertler and Karadi (2015), Lunsford (2015) employs the proxy as the dependent variable. More precisely, the F-statistic is based on the linear projection of $s^*_t$ on the entire vector of reduced form
errors and a constant:

\[ s_t^* = c_0 + u_t' \mu + \nu_t, \]  

(13)

where \( \mu \) is a vector of dimension \( n \times 1 \) and \( \nu_t \) is an i.i.d. error term with variance \( \sigma_\nu^2 \). Furthermore, \( \nu_t \) is assumed to be independent of \( \epsilon_t \) and lags of \( y_t \). To test whether \( s^* \) constitutes a weak proxy variable for \( \epsilon_m \), Lunsford (2015) suggests computing the F-statistic of Equation (13) for the null hypothesis that \( \mu = 0 \):

\[ F_{WP} = \left( \frac{T - n}{n} \right) \frac{(s_t^* - \bar{s}_t^*)^2 - [(s_t^* - \bar{s}_t^*) - \hat{u}_t' \hat{\mu}]^2}{(s_t^* - \bar{s}_t^*) - \hat{u}_t' \hat{\mu}]^2}. \]

(14)

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 proxy variable is only imperfectly correlated with the structural shock of interest and biases the estimator of \( a_m \) towards zero. The corresponding studies in the literature, that is, 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 a tolerated asymptotic bias level of 10\%.

### 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, that is, \( B_0, B_1, B(L), \) and \( \Sigma_u \), and the corresponding reduced form errors \( u_t \), we compute the two F-statistics. \( F_{IV} \) is the statistic for the test \( \gamma = 0 \) in Equation (11). To decide whether we can reject the null hypothesis of weak proxies, we use a critical value of 10, as in Gertler and Karadi (2015). Table 2 presents the results for this F-statistic. All three monetary policy shock measures constitute good proxy variables for monetary policy shocks. The second test statistic, \( F_{WP} \), is related to

<table>
<thead>
<tr>
<th>RR</th>
<th>BC</th>
<th>GK</th>
</tr>
</thead>
<tbody>
<tr>
<td>( F_{IV} )</td>
<td>23.353*</td>
<td>11.423*</td>
</tr>
</tbody>
</table>

* significant at the 5\% level according to the rule of thumb \( F_{IV} > 10 \)

Table 2: Results \( F_{IV} \)

the test \( \mu = 0 \) in Equation (13). To determine whether we can reject the null hypothesis that \( \mu = 0 \), we use the critical values provided by Lunsford (2015). Table 3 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 measure constructed by Romer and Romer (2004), we cannot reject the null hypothesis at a significance level of 5%. This result implies that also these shocks are considered 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 that this measure is a weak proxy variable at a significance level of 5%.

<table>
<thead>
<tr>
<th></th>
<th>RR</th>
<th>BC</th>
<th>GK</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F_{WP}$</td>
<td>8.134</td>
<td>3.879</td>
<td>11.901*</td>
</tr>
</tbody>
</table>

*significant at the 5% level according to critical value of Lunsford (2015) $F_{WP} > 8.53$

Table 3: Results $F_{WP}$

Both test procedures produce different results. $F_{IV}$ qualifies all three shock measures as relevant proxy variables, whereas $F_{WP}$ determines that the best proxy variable is the shock measure of Gertler and Karadi (2015). In his analysis, Lunsford (2015) points to an important difference between the two tests. More precisely, he shows that $F_{WP}$ converges in distribution to a non-central $\chi^2_n$ distribution that has as its non-centrality parameter the signal-to-noise ratio $\frac{C^2}{\sigma^2}$ of the proxy variable, where $C$ captures the correlation between the proxy and the structural shock. The signal-to-noise ratio measures the signal that the proxy contains for the structural shock relative to the noise of the proxy. The non-centrality parameter of the asymptotic distribution of $F_{WP}$ is therefore independent of any VAR model coefficients. Thus, critical values for Lunsford’s $F$-statistic are determined precisely and hold independently of the VAR model specification. In contrast, $F_{IV}$ converges in distribution to a non-central $\chi^2_1$ distribution that has a non-centrality parameter of

$$\frac{a_{11}^2 C^2}{a_{11}^2 + a_{12} \Sigma_{\epsilon_2} a_{12}'}.$$  \hspace{1cm} (15)

$a_{11}$ and $a_{12}$ denote elements from the structural coefficient matrix in Equation (6), which is partitioned to correspond to Equation (7) and (8), and $\Sigma_{\epsilon_2} = E[\epsilon_{2,t} \epsilon_{2,t}']$. The relation in (15) shows that the signal-to-noise ratio $\frac{C^2}{\sigma^2}$ is scaled by $\frac{a_{11}^2}{a_{11}^2 + a_{12} \Sigma_{\epsilon_2} a_{12}'}$ in the case of $F_{IV}$. Thus, critical values of $F_{IV}$ are not independent of $A$ and will vary with the VAR model specification. This argument, as well as the availability of precise critical values for the test statistic, leads us to base our conclusion on the $F_{WP}$ statistic rather than on the $F_{IV}$ statistic. Consequently, we find that the measure by Gertler and Karadi (2015) constitutes the best proxy variable for monetary policy shocks.
4 Conclusion

In this study, we attempt to identify the monetary policy shock measure that constitutes the best proxy variable to study the effects of monetary policy. In our analysis, we employ two F-statistics that allow us to test whether the shock measures are a weak proxy. On the one hand, we apply a test that treats the proxy variable as an independent variable. This test finds that none of our monetary policy shock measures is a weak proxy variable. On the other hand, we use a test in which the proxy is the dependent variable. This test rejects the null hypothesis that the proxy variable constitutes a weak proxy only for the measure provided by Gertler and Karadi (2015). Since the latter test employs precise critical values and its F-statistic converges to a distribution that is independent of the VAR model, we base our conclusion on this test. Therefore, 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 for studying monetary policy effects.
References


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.


**Barakchian and Crowe (2013) Shock:** The high-frequency monetary policy shock measure is downloadable as shock.


**Gertler and Karadi (2015) Shock:** The high-frequency monetary policy shock measure is provided by the authors as ff4.


**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).

https://fred.stlouisfed.org/series/PPIFGS (01/30/2017).