A Study of the Romer and Romer Monetary Policy Shocks Using Revised Data

Andersson, Fredrik N G; Kilman, Josefin

2021

Working Paper 2021:19

Department of Economics
School of Economics and Management

A Study of the Romer and Romer Monetary Policy Shocks Using Revised Data

Fredrik N. G. Andersson
Josefin Kilman

December 2021

LUND UNIVERSITY

A Study of the Romer and Romer Monetary Policy Shocks Using Revised Data*

Fredrik N G Andersson[†], Josefin Kilman[‡]

Abstract

Romer and Romer (2004) propose a simple method to estimate monetary policy shocks using forecasts and real-time data. However, such data is not always (publicly) available, especially in a historical context. We explore the consequences of using revised data instead of the original forecast and real-time data when estimating policy shocks using the Romer and Romer framework. To this end, we estimate policy shocks for the same period as Romer and Romer. We find that using revised data has little impact on actual shock estimates, and the estimated effects of monetary policy shocks are similar.

JEL Classification: E2, E3, E4, E5, E6.
Keywords: Monetary policy shocks, prices, GDP.

*We thank David Romer, James Cloyne, and Klas Fregert for their valuable input.
[†]Department of Economics, Lund University. E-mail: ngf.andersson@nek.lu.se.
[‡]Department of Economics, Lund University. E-mail: josefin.kilman@nek.lu.se.

1 Introduction

Estimating the impact of monetary policy on, for example, output (Christiano et al., 1999; Bernanke et al., 2005; Barakchian and Crowe, 2013), inflation (Cloyne and Hürtgen, 2016; Coibion, 2012), exchange rates (Kim and Lim, 2018), and financial markets (Gertler and Karadi, 2015) is an important research field. However, this line of research faces the key problem of identifying the casual effect of monetary policy. Changes in central bank policy rates are caused by both discrete policy changes, so-called policy shocks, aimed at affecting economic outcomes, and endogenous changes in response to economic events. The latter causes an endogeneity problem that may result in biased estimates of the impact of monetary policy on the economy (Ramey, 2016).

Several methods have been proposed to separate endogenous movements in the policy rate from exogenous monetary policy shocks (see Ramey, 2016, for a review). These include structural vector autoregressive (SVAR) models (Christiano et al., 1999), high-frequency identification methods (Kuttner, 2001; Gertler and Karadi, 2015), the narrative approach (Cloyne, 2013; Romer and Romer, 1989), and dynamic stochastic general equilibrium (DSGE) models (Smets and Wouters, 2007). An additional method related both to the SVAR and the narrative literature is the two-step method proposed by Romer and Romer (2004). The method has gained in popularity over time due to its simplicity (see e.g., Cloyne and Hürtgen, 2016; Coibion, 2012; Coibion et al., 2017; Leahy and Thapar, 2019; Miranda-Agrippino and Rey, 2020; Tenreyro and Thwaites, 2016; Doniger, 2019; Holm et al., 2021; Aikman et al., 2018). In the first estimation step, policy shocks are identified by regressing changes in the central bank policy rate on forecasts of future changes in the economic environment. Similar to a SVAR model, any change in the policy rate that is explained by a change in the variables included in the model is defined as an endogenous movement in the policy rate. The remaining variation, the residuals, are defined as the policy shocks. In the second step, the effect of monetary policy on the economy is explored by regressing the economic variable on the policy shocks.

Unlike a SVAR model, both the first and the second step regressions are carried out as single equation regressions rather than a system of equations. The modeling is thus more flexible compared to a SVAR model, and the variables included in the first and the second stage regressions can differ. The estimation technique relates to the narrative approach in which Romer and Romer (R&R) collect data on intended policy changes from Federal Open Market Committee (FOMC) minutes, and forecasts from the Federal Reserve's Greenbook, to ensure that the variables used to estimate the policy shocks are the actual variables available to the governors at the time of the interest rate decision.[1] They argue that "[t]he resulting series for monetary shocks should be relatively free of both endogenous and anticipatory

[1]The research staff at the Board of Governors prepare the Greenbook before each meeting of the Federal Open Market Committee, and it contains macroeconomic variable forecasts (Hoesch et al., 2020).

actions" (Romer and Romer, 2004, p. 1056).

The purpose of this paper is to explore to what extent the policy shock estimates and the estimated effects of monetary policy on output and inflation are affected by using revised outcome data instead of using real-time and forecast data, within the (R&R) framework. Arguments against using revised data are obvious, as it is not the data that was available to the central bank at the time of the interest rate decision. Significant and persistent differences between the real-time and revised data may cause biased parameter estimates in the first-stage regression, which also biases the shock estimates (Orphanides, 2001, 2003; Molodtsova et al., 2008). The argument for using revised data is more practical. The availability of real-time data is in many cases restricted, not least in a historical context. Central bank forecast data from before the 1990s can be even harder to obtain.[2] Researchers are simply forced to use revised data as it is the only data available.

We estimate two sets of policy shocks using revised data for the United States between 1969 and 2008, and compare these to the original R&R shocks.[3] For the first set of shocks, we closely follow R&R and regress the policy rate change on GDP growth, inflation, and unemployment, using revised outcome data for the independent variables instead of real-time forecast data. In our second set of shocks, we try to proxy a larger information set that R&R captured using forecast data, by regressing the policy rate change on four principal components based on 18 macroeconomic variables (also revised and outcome variables). We find that these two sets of shocks are similar to the original R&R shocks. The correlation between our shocks and the R&R shocks is 0.82 for the full sample period. The correlation increases to 0.88 if we exclude the 1970s. More importantly, the estimated effects of monetary policy on GDP growth and consumer price inflation are of a similar magnitude and contain the same turning points. These results indicate that the use of real-time and forecast data is not crucial when modeling the impact of monetary policy, at least not for the United States and the period we consider. We contribute to the literature by providing a simple method to estimate monetary shocks that can extend the number of applications in the two-step approach in Romer and Romer (2004).

The rest of the paper is set out as follows. We discuss the identification of monetary policy shocks in Section 2. Section 3 contains the empirical analysis, and section 4 concludes the paper.

[2]Real-time data is available for the United States (Archival Federal Reserve Economic Data (ALFRED) database), and a number of the OECD countries (Federal Reserve Bank of Dallas and ECB). The ECB provides a real-time database for the euro area (but not individual euro area countries), Japan, and the United states. The IMF provides forecast data for some advanced countries from 1990. The variables in these databases are few though. There are several private sector forecasts, but these are often not publicly available.

[3]The original R&R shocks are available between 1969 and 1996. We use the shock series provided by Coibion et al. (2017) which are updated through 2008. Between 1969 and 1996, these shocks are close to being identical (correlation is 0.99).

2 Monetary policy shocks

Monetary policy shocks are defined as unexpected discretionary changes in the central bank policy rate. Identifying these shocks is crucial when studying the impact of monetary policy on the economy. The observed central bank policy rate contains both unexpected policy shocks and endogenous changes in the policy rate due to shifts in the economic climate. This creates an endogeneity problem that may bias the estimated effects of monetary policy on the economy (Ramey, 2016). To identify the policy shocks, it is common to linearly decompose the central bank policy rate into the two components,

$$\triangle i_t = f(\Omega_t) + p_t \tag{1}$$

where i is the policy rate at time t, Ω_t is the central bank's information set at time t, f is the central bank's reaction function to the information set, and p_t is the policy shock at time t (see e.g., Cloyne and Hürtgen, 2016; Croushore and Evans, 2006). In practice, neither the central bank's reaction function nor the central bank information set are fully observable and must be estimated. The estimation creates an error in variables problem. For simplicity assume that

$$\hat{p}_t = p_t + \theta_t \tag{2}$$

where $\hat{p}_t$ is the estimated policy shock, and θ_t is the error. When the shocks are applied to estimate the effect of monetary policy, the error may bias the estimated effect. Assume that the impact of monetary policy is estimated by the following model

$$y_t = \alpha_0 + \alpha_1 p_t + \epsilon_t \tag{3}$$

Because only the estimated shocks, $\hat{p}_t$, are observed and not the true shocks, the model that is actually estimated is given by

$$y_t = \alpha_0 + \alpha_1(\hat{p}_t - \theta_t) + \epsilon_t = \alpha_0 + \alpha_1\hat{p}_t + (\epsilon_t - \alpha_1\theta_t) \tag{4}$$

There is an endogeneity problem between the estimated policy shocks and the error term in (4). The impact of the error on the estimate of the parameter of interest, α_1, depends on the nature of the error θ. In the simple case when θ is an independent random variable with a mean zero and variance σ_θ^2, the OLS parameter estimate becomes

$$plim\ \hat{\alpha}_1 = \alpha_1 \frac{\sigma_p^2}{\sigma_p^2 + \sigma_\theta^2} \tag{5}$$

where σ_p^2 is the variance of the policy shocks. As can be seen in (5), the estimated impact is biased towards zero. The size of the bias depends on the size of the estimation error in comparison to the size of the variance of the shocks. The larger the variance of the error, the more biased the estimated parameter ceteris paribus is. For more complicated forms of measurement errors, such as when the estimated policy shocks and the error are correlated, the character of the bias becomes more complex. There is no easy method to deal with measurement error, especially when the error is of an unknown form. The focus of the literature on policy shocks is to find estimates that are as close to the true shocks as possible.

2.1 Methods to estimate monetary shocks

Previous papers that estimate monetary policy shocks contain a wide set of different methods (Ramey, 2016, provides a review of the most common methods). Some methods are purely econometric, some are primarily theoretical, while others combine econometrics with economic theory. Vector autoregressive (VAR) models, or more specifically SVAR models (see e.g., Bernanke and Mihov, 1998; Christiano et al., 1999), and DSGE models (see e.g., Smets and Wouters, 2007; Schmitt-Groh and Uribe, 2012) are two examples where economic theory is used together with an empirical model to identify exogenous policy shocks. These methods first try to obtain the central bank's reaction function, f. They then define the policy shocks as deviations from the expected response to changes in the variables included in the central bank's information set, where the variables included can be both real-time and revised data. This approach suffers from several shortcomings. First, the central bank's reaction functions are estimated either using a theoretical or econometric model. Further, the reaction function is assumed to be stable over longer periods. Changes in the composition of the monetary policy committee may change how the central bank reacts to changes in the economy. Second, it assumes that the public knows the central bank's reaction function and that all changes in the policy rate not captured by the reaction function are unexpected shocks. It could be that the central bank's policy decision was anticipated due to e.g., the central bank's communication before the meeting, even if it breaks with the bank's normal response pattern. Third, the policy shocks are residuals and thus capture all changes not explained by the model, including policy errors and measurement errors in the central bank's information set. Thus, not all the shocks are unexpected discrete policy changes from the central bank aimed at affecting the economic outcome.

An alternative to a model-based approach is to use high-frequency shocks. In the high-frequency approach, monetary shocks are estimated by comparing movements in real-time financial market data, as proxies for monetary policy expectations, in a short time window around monetary policy announcement. For the United States, these shocks are commonly identified using federal funds futures and a 30-minute window around policy announcements from the Fed. The identifying assumption is that when market rates move in this short period, it reflects unexpected changes in monetary policy, or shocks, since the announcement is the only new information to market participants (see e.g., Kuttner, 2001; Gertler and Karadi, 2015; Nakamura and Steinsson, 2018). By studying the reaction of financial markets, the monetary surprises are identified without having to estimate the central bank's reaction function.[4] The method's weakness is the assumption that all changes in the interest rates during the window are caused by the central bank. The method also assumes that movements in the interest rate during the short window translates into changes in the real economy. It could be that the movements during the window are only temporary and not sufficiently persistent to affect the real economy.

A third approach is the so-called narrative approach. Then researchers study policy documents in detail to gauge the motivation behind each policy decision. Those policy changes that were not motivated by the state (or expected state) of the economy, or where the policy was in the opposite direction compared to conventional wisdom given the state of the economy, are defined as exogenous policy events (see e.g., Friedman and Schwartz, 1963; Romer and Romer, 1989). This method is sensitive to the personal interpretation of the researcher and their subjective interpretation of the documents.

2.2 Romer and Romer monetary policy shocks

Romer and Romer (2004) proposes a fourth method to estimate monetary policy shocks. The method relates to both the econometric and the narrative approaches. Similar to a SVAR model, they estimate the Federal Reserve's reaction function using an econometric model. Like the narrative approach, they go through the FOMC's minutes and the Greenbook to collect data that was available to the Federal Reserve at the time of their interest rate decision. This gives them a data set that contains real-time data and the Federal Reserve's forecast of future economic developments. The inclusion of forecast data has two advantages. First, monetary policy operates with a time lag whereby the central bank is likely to respond to information about future economic developments, and not just on observed changes. Forecasts are a good proxy for these anticipatory movements of the economy. Second, forecasts are commonly generated using a large set of variables, such that

$$\tilde{x}_{t+h} = g(\Omega_t) \tag{6}$$

where $\tilde{x}_{t+h}$ is the forecast of the variable x for $h > 0$ future periods. Because the forecasts are likely based on a wider set of variables, it is possible to reduce the dimension of the variable set included in the estimation of the reaction function

[4] More recently, it is common to use proxy SVARs with external instruments for monetary shocks, where the instrument is usually high-frequency monetary shocks (Stock and Watson, 2012; Mertens and Ravn, 2013; Gertler and Karadi, 2015; Cesa-Bianchi et al., 2020).

and thereby avoid problems of over-parameterization.[5,6]

R&R is a two-step method. First, the central bank's reaction function is estimated using the forecast data, and the shocks are defined as the residuals from that regression. Second, the shocks are applied to estimate the impact on the economy. The two-step approach allows for a large amount of flexibility in the modeling. The variables included in the first and second stages can be different, and the functional form can vary between the two modeling stages. It is also possible to apply the policy shock estimates in a local projection estimation to generate impulse-response figures, rather than relying on a VAR model to generate them.[7] It is unsurprising that the method is widely used (see e.g., Coibion, 2012; Leahy and Thapar, 2019; Miranda-Agrippino and Rey, 2020; Doniger, 2019; Holm et al., 2021), especially when the focus is on the effect of monetary policy on one specific variable and where the possible mechanisms through which the monetary policy operates is of less interest (see e.g., Blinder and Watson, 2016; Tenreyro and Thwaites, 2016; Coibion et al., 2017; Lennard, 2018). Our method, to replace real-time forecast data with revised outcome data in the R&R approach is similar to a SVAR model including revised data. However, the R&R approach provides flexibility that a SVAR model does not. To identify potential policy channels, a SVAR model or a DSGE model are more appropriate.[8]

2.3 Romer and Romer monetary policy shocks using revised data

Real-time data and central bank forecasts are not always available, especially in a historical context. This raises the question if it is possible to apply the same two-step modeling approach as in R&R but with revised data instead. Excluding forecast data is potentially less of a problem compared to using revised instead of real-time data. Provided that forecasts are a mapping of the central bank's information set, we can exclude the forecasts and use a larger set of variables instead. The main potential problem lies with the replacement of the actual data available to the central bank at the time of the monetary policy decision with revised data. Approximating the central bank's information set Ω_t with $\hat{\Omega}_t$ introduces an additional layer of measurement error in (2), i.e., $\theta_{RR,t} < \theta_{RD,t}$ where $\theta_{RR,t}$ is the measurement error from the R&R policy shock estimation and $\theta_{RD,t}$ is the measurement error from the estimation with revised data. Under the assumption that there is a

[5] Alternatives to avoid problems with over-parameterization is to use a factor augmented VAR that incorporates more information into VAR models (Bernanke et al., 2005).

[6] The central bank's forecast model can be i) an explicit econometric forecast model, ii) the forecasters own personal judgments, or iii) a hybrid of the first two (Cloyne and Hürtgen, 2016).

[7] If a VAR model is misspecified, the bias in the estimated coefficients worsens over horizons. Local projections are less likely to be misspecified, which makes it a more flexible method to estimate monetary shock impacts. The local projection method is less successful with small samples though (Miranda-Agrippino and Ricco, 2021).

[8] One potential limitation of the two-step approach is the loss of observations in both steps (equal to the number of dynamic responses estimated and the number of lags included in the information set). Christiano et al. (1999) argues that it is convenient to map the two-step procedure into a VAR-based procedure since only the number of lags included in the information set is dropped.

simple form of measurement error, the larger error will cause an even larger bias of the estimated impact of monetary policy in (5). For more complex forms of the measurement error, the impact of the error depends on the nature of the error, and it can both increase and decrease the bias, resulting in a negative or a positive bias. A larger measurement error implies that using revised data should never be the first option. However, it could be a second-best choice provided that real-time data is not available. There is no method to judge the size of the error theoretically. It is a question that needs to be empirically explored.

3 Empirical analysis

We evaluate how large an impact the use of revised data rather than real-time and forecast data has on the shock estimates, by comparing the original R&R shocks with shock estimates using the same two-step approach including revised data. The original R&R shocks cover the period from 1969 to 1996. Coibion et al. (2017) extends the shocks through 2008 using the same methodology and data sources as R&R, extending the sample from 1969 to 2008.[9]

3.1 Revised data

R&R model the central bank's information set as consisting of forecast data on inflation, GDP growth, and unemployment, as well as real-time data on inflation and GDP growth. Replacing the original data with revised data may bias the policy shock estimates and the estimated economic responses to a policy shock. However, the correlation between the real-time data on growth and inflation (collected from R&R's dataset), and the revised data is high, as seen in Table 1. For GDP growth, the correlation between the real-time and the revised data is 0.80, and the correlation between real-time inflation and revised inflation is 0.85. This suggests that using revised data, while not optimal, is potentially not a great concern.

Table 1: Correlation between R&R's real-time data and our revised data

R&R variable/Our data	GDP growth	Inflation
GDP growth	0.80	
Inflation		0.85

Notes: The data is at the quarterly frequency and runs between 1969q1-2008q4. R&R measure inflation with the GDP/GNP deflator, while we measure inflation with CPI. I do not include unemployment since R&R do not include real-time data on unemployment in their specification.

In our estimation, we also exclude the original forecast data. As shown in (6), most forecasts are a mapping of the available information set. We can thus replace the forecasts with the actual information set when we estimate the policy shocks.

[9]The correlation between the original shocks and the updated shocks is 0.992 for the overlapping period.

However, because the central bank's information set is larger than inflation, GDP growth, and unemployment, our model is potentially too small to reproduce the Federal Reserve's forecasts. Thus, we consider two sets of shocks: our *baseline* shocks include data on inflation, GDP growth, and unemployment only, while our *PCA* shocks include 18 macroeconomic variables to extend the information set. We follow Bernanke et al. (2005) and Stock and Watson (2002) and reduce the dimension of the data set by applying a principal component analysis (PCA) before we estimate the second set of shocks. More information on the shocks is found in the next section.

Table 2 presents the R-squared value from regressing the Federal Reserve's forecasts of GDP growth, inflation, and unemployment on our revised data set. The explanatory power of those regressions is relatively high: between 77 and 89 percent of the variation in the forecast data for GDP-growth and inflation. The explanatory power is higher when using the principal components compared to the baseline regressions with only three explanatory variables. For unemployment, the explained variation is lower between 14 and 41 percent. The high explanatory power suggests that we can replace the forecast data with our revised data set, with the possible exception of unemployment. In summation, we find that our revised data set correlates highly with the R&R data set, suggesting that our shock estimates should be similar to the R&R estimates. In the next section, we will explore whether this is the case.

Table 2: Proportion of variance in forecasts explained by our approaches

Dependent variable	**Independet variables**	**R2-value**
Forecast of GDP growth	GDP growth and inflation (levels and first differences), unemployment (level)	0.77
Forecast of GDP growth	4 principal components (levels and first differences)	0.80
Forecast of inflation	GDP growth and inflation (levels and first differences), unemployment (level)	0.78
Forecast of inflation	4 principal components (levels and first differences)	0.89
Forecast of unemployment	GDP growth and inflation (levels and first differences), unemployment (level)	0.14
Forecast of unemployment	4 principal components (levels and first differences)	0.41

Notes: The data is at the quarterly frequency and runs between 1969q1-2008q4. R&R measure inflation with the GDP/GNP deflator, while we measure inflation with CPI. The independent variables are either the ones included in our *baseline* shocks (GDP growth, inflation and unemployment in levels and first differences) or *PCA* shocks (four principle components in levels and first differences). The four principal components are based on 18 macroeconomic variables found in Appendix Table A1. See section 3.2.1 for further details on the shocks and variables included.

3.2 Evaluation of our monetary policy shocks

3.2.1 Estimation of the shocks

We estimate a similar model as in R&R, but replace the forecast data with revised outcome data[10]

$$\triangle i_t = \beta_0 + \beta_1 i_{t-1} + \beta_{2i} x_{it} + \beta_{it} \triangle x_{it} + \epsilon_t \tag{7}$$

where x_{it} is a vector including GDP growth, inflation, and unemployment, and $\triangle x_{it}$ includes the first difference of inflation and GDP growth. ϵ_t is the monetary policy shocks that we refer to as the *baseline* shocks.[11] i_t is the intended federal funds rate collected from Romer and Romer (2004).[12] However, as highlighted in Cloyne and Hürtgen (2016), the intended funds rate is the actual policy rate in most countries. Lastly, i_{t-1} is the previous period's intended federal funds rate in line with the specification in Romer and Romer (2004).[13] We include x_{it} and not x_{it-1} in our specification, since we try to capture the information set ($\hat{\Omega}_t$) that R&R include in their specification (Ω_t), not mimic the reaction function assumed in R&R. R&R also control for the previous period's forecast, which is most often real-time data on GDP growth and inflation rather than a forecast, as well as forecasts of up to two quarters into the future. Shocks are close to identical if we include lagged values of our variables as well. We do not include any leads, since we do not use forecast data. The forecasts are instead captured by the information set (see (6)).[14]

In our second set of shocks, we include additional variables to extend the central bank's information set. We include 18 variables found in the Survey of Professional Forecasters, also included in the Federal Reserve's Greenbook.[15] There is a risk of

[10] Romer and Romer (2004) estimate the following model at a meeting frequency m: $\triangle i_m = \alpha + \beta i_{m-1} + \sum_{i=-1}^{2} \gamma_i \tilde{x}_{mi} + \sum_{i=-1}^{2} \delta_i \triangle \tilde{x}_{mi} + \rho u_{m0} + \epsilon_m$, where $\tilde{x}_{mi}$ is a vector including forecasts on GDP growth and inflation at the FOMC meeting date m for horizon i, $\triangle \tilde{x}_{mi}$ is the corresponding change in the forecast since the previous meeting, u_{m0} is the current forecast for unemployment, and ϵ_m is the monetary policy shock at the meeting m. Horizon refers to quarters, and they include forecasts (often real-time data) from the previous quarter to capture the present state of the economy, and forecasts of up to two quarters into the future to capture anticipatory changes in economic conditions. The policy rate is modeled using intended federal fund rate changes.

[11] We measure inflation with the consumer price index (CPI), because the Fed followed the CPI closely up until 2000 (Hanson, 2004). However, shocks do not differ if we use personal consumption expenditures (PCE) inflation.

[12] The intended federal funds rate is the implied policy target rate from the FOMC's minutes. Intended fund rate changes capture what the Federal Reserve intends to happen to the fund rate because of the actions agreed upon at each meeting. We use the series on the intended fund rate from Coibion et al. (2017), which is identical to the R&R data but extended through 2008.

[13] Data on both the intended change in the fed fund rate and the previous period's intended fed fund rate are at the monthly frequency. To estimate monetary shocks at the quarterly frequency, we sum up these series over each quarter in line with the method in Coibion et al. (2017).

[14] The R&R approach handles the reverse causality issue when modeling the policy rule neatly, by using forecast data that is available just before the policy decision. Since the first step is made on a meeting-by-meeting frequency, they only include information that is available close to the policy decision. One limitation with our approach is that we cannot account for this issue, since we do not include forecast data. However, we do not want to capture their reaction function, only the information set incorporated in their reaction function, so this should not be a large problem.

[15] The Survey of Professional Forecasters (conducted by the Federal Reserve Bank of Philadelphia) is a quarterly survey of macroeconomic forecasts in the United States.

over-fitting the model by including up to 18 variables. Therefore, we reduce the dimension of the data set by applying a principal component analysis before we estimate the second set of shocks. This method relates to the factor-augmented VAR (FAVAR) approach in Bernanke et al. (2005), who also use revised data, but we use the R&R two-step approach by first estimating the shocks. Appendix A provides details on the variables included in the PCA and the components. We include four common components in the analysis representing 72 percent of the variance in the data.[16] We run model (7) where the vector x_{it} now includes the four principal components, $\triangle x_{it}$ includes the first difference of the principal components, and ϵ_t is the monetary policy shocks that we refer to as the *PC* shocks.[17,18] The final sample consists of two sets of monetary shocks, baseline shocks and PC shocks, at the quarterly frequency between 1969q1 and 2008q4.

Table 3 reports the coefficient estimates from estimating model (7) for the two types of shocks. Column (1) presents the coefficients from the first-stage regression when estimating the baseline shocks, controlling for GDP growth, inflation, and unemployment. Column (2) presents the coefficients from the first-stage regression when estimating the PC shocks, controlling for four principal components based on the macroeconomic variables presented in Appendix Table A1. The coefficient estimates are not important in themselves, since we are only interested in the residuals of model (7).[19] However, the signs of most variables are in line with theory. We find that a higher inflation rate significantly increases the intended policy rate and that a higher unemployment rate significantly decreases the intended policy rate. In the PC approach, only the first and second principal components significantly impact the intended policy rate. The Appendix Table A2 show that the first and second principal components mainly capture variables connected to GDP and inflation, which makes the positive coefficient in line with theory. The R-squared values are similar to those in Romer and Romer (2004) and Cloyne and Hürtgen (2016).

[16] Shocks are close to identical if we only include two or three components. See Appendix Figure A1 for a plot of the eigenvalues, and Appendix Table A2 for a specification on which variables the different components include.

[17] We also estimated a model with all 18 variables included. Those shocks are similar to the PC shocks, but slightly less correlated with the R&R shocks.

[18] This approach is somewhat related to the one in Aikman et al. (2018), who estimate monetary shocks for the UK and include both forecast data and principal components estimated with macroeconomic and financial data.

[19] It is important to highlight that this is not the same specification as the one R&R use.

Table 3: Regression output from estimating our shocks

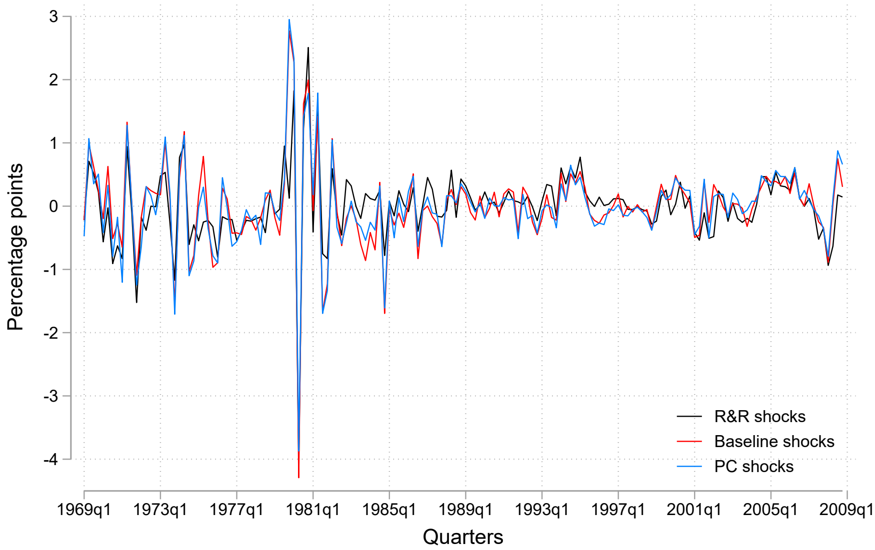
	(1) Baseline approach		(2) PC approach
i_{t-1}	-0.015	i_{t-1}	-0.011
	(0.022)		(0.010)
GDPgrowth$_t$	0.162	PC1$_t$	0.096***
	(0.132)		(0.034)
$\triangle$GDPgrowth$_t$	-0.020	PC2$_t$	0.154***
	(0.082)		(0.042)
Inflation$_t$	0.219**	PC3$_t$	-0.036
	(0.108)		(0.063)
$\triangle$Inflation$_t$	-0.097	PC4$_t$	0.075
	(0.127)		(0.075)
Unemployment rate$_t$	-1.013***	$\triangle$PC1$_t$	0.051
	(0.266)		(0.047)
Constant	-0.003	$\triangle$PC2$_t$	0.011
	(0.002)		(0.058)
		$\triangle$PC3$_t$	-0.000
			(0.056)
		$\triangle$PC4$_t$	-0.017
			(0.060)
		Constant	0.128
			(0.160)
Observations	160		160
R-squared	0.278		0.282

Notes: Coefficients are from estimating model (7). Levels of significance: *** p<0.01, ** p<0.05, * p<0.1.

3.2.2 Comparing the shock estimates

Figure 1 illustrates the original R&R shocks (black line), the baseline shocks (red line) and the PC shocks (blue line). Both of our sets of shocks track the main movements in the original R&R shocks relatively closely, indicating that the forecast error and the data revision error are relatively small. In terms of policy, we can observe large swings between expansionary and contractionary policies during the 1970s, followed by the Volcker disinflation policy in the late 1970s to the early 1980s resulting in large shocks. The volatility in the policy shocks declines thereafter with relatively small periods of policy tightening and loosening. Irrespective of the shock series, we draw similar conclusions as to when policy is expansionary and contractionary, and the size of the shocks is similar. Again, this points towards relatively small forecast model errors and data revision errors.

Figure 1: Graphical comparison of our monetary shocks and the R&R shocks



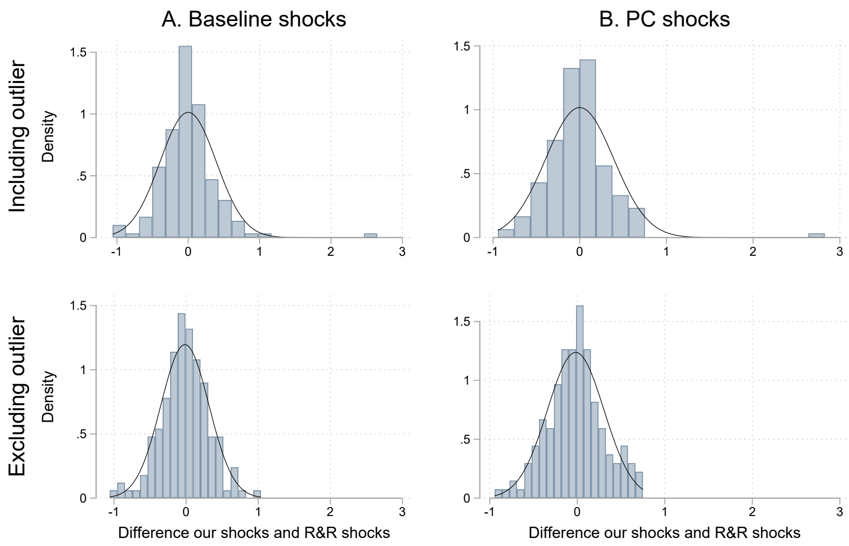
Notes: The original R&R shocks are in black. The baseline shocks are in red and the PC shocks are in blue. The sample runs between 1969q1 and 2008q4.

The correlation between the R&R shocks and the baseline shocks is 0.822, and the correlation between the R&R shocks and the PC shocks is 0.823. If we exclude the 1970s and start the sample period around the time of the Volker disinflation period (sample period 1980 to 2008), the correlation increases to 0.888 and 0.881 respectively. Irrespective of the period, the correlation is high and there is almost no gain in correlation by expanding the number of variables in the shock estimation. Shocks from the simple model with only three variables do not perform significantly worse compared to the shocks from the PC-based model (the correlation between the baseline and the PC shocks is 0.904). This suggests that a larger model is not necessary.

The volatility in the baseline shocks and the PC shocks is, as expected, higher than the volatility in the R&R shocks since they include a larger measurement error. The standard deviation of the R&R shocks is 0.588, the standard deviation of the baseline shocks is 0.691, and the standard deviation for the PC shocks is 0.689. Volatility is the highest for all series during the activist 1970s and 1980s but is reduced after 1985 to 0.294, 0.304, and 0.308. The reduction in volatility is similar for our shocks compared to the R&R shocks.

So far, the analysis indicates that the difference between our shocks and the original R&R shocks are small. Next, we illustrate the difference between the shocks by using histograms plotting the difference between our baseline shocks and the R&R shocks in Figure 2 Panel A, and between our PC shocks and the R&R shocks in Figure 2 Panel B.

Figure 2: Histograms on the difference between our shocks and the R&R shocks



Notes: The figure illustrates histograms on the differences between our shocks and the original R&R shocks (x-axis), using 20 bins. The upper row shows histograms when all observations are included. The lower row shows histograms without an outlier in 1979q4. The black line is a normal density. Panel A is the difference between our baseline shocks and the R&R shocks. Panel B is the difference between our PC shocks and the R&R shocks.

There is one observation when our shocks are clearly different from the R&R shocks in Figure 1, which is in the fourth quarter in 1979. Both of our shocks are more contractionary compared to the original R&R shocks in this quarter. R&R find a positive policy shocks of 0.13 percentage points, while our shocks are 2.63 and 2.95 percentage points.[20] If we exclude this outlier period, see the second row of Figure 2, the difference in shocks becomes bell-shaped. The Jarque-Bera normality test cannot rule out that the difference in shocks is normally distributed.[21] This means that there is no systematic bias in our shocks. In summation, the high correlation between our shocks and the R&R shocks indicates that we should obtain similar estimates of the effect of monetary policy in the second-step regressions.

3.3 Estimating the impact of monetary policy on output and prices

Following Romer and Romer (2004), we estimate the effect of monetary policy shocks on output and prices using a VAR model. A VAR model is not necessary for second-step regressions given that the policy shocks are exogenous. Nevertheless, R&R use a VAR model to track the dynamic effect of the shocks in impulse response figures.[22] We estimate two VAR models. In one, output (y_t) is measured using the

[20]The discrepancy between our shocks and the original R&R shocks should be explained by the lack of data from the FOMC's October meeting in 1979. The R&R shocks are therefore missing for this meeting, see Romer and Romer (2004)'s data file. Hence, our shocks capture movements during this time that R&R could not capture.

[21]We find the same results when running the Shapiro-Wilk W test for normality.

[22]This is sometimes referred to as a hybrid VAR since the actual monetary shocks are included instead of the fed fund rate (this set-up is also used in Cloyne and Hürtgen (2016), Gertler and

log of real GDP and the price level (pl_t) is measured with the log of the CPI. The model is estimated using quarterly data. In the second model, we replace GDP with the industrial production index and CPI with the producer price index (PPI). The second model specification is close to the model used by R&R. However, it captures a much smaller segment of the economy compared to using the GDP and CPI, whereby we consider both. Specifically, we estimate the following VAR model

$$X_t = B(L)X_{t-1} + \epsilon_t \tag{8}$$

where B(L) is the lag polynomial with p lags and the vector of observables is $X_t = [y_t, pl_t, \hat{p}_t]$, where $\hat{p}_t$ is our monetary shocks. Like the set-up in R&R, we assume that monetary policy responds to changes in the economy but does not cause changes contemporaneously, and therefore order the shocks last. To estimate quarterly shocks, we follow Romer and Romer (2004) and Cloyne and Hürtgen (2016) and cumulate the shocks to represent the sum of all previous quarters' shocks, since VARs usually include interest rates in levels. In each model, we include p = 3 years of lags in line with R&R.

Impulse-response figures from the first VAR model with real GDP (Panel B) and CPI (Panel C) are shown in Figure 3. The thick black line represents the response following a shock in the original R&R shocks. The dashed black line is the estimated 95 percent confidence bound for these shocks. The red line illustrates the response to our baseline shocks, and the blue line illustrates the response to our PC shocks. The figure illustrates the responses to a one standard deviation (contractionary) monetary shock. GDP falls in response to the shock and the largest drop in output occurs approximately two years after the shock. Prices initially increase in response to the shock and start to fall two years after the shock.[23] As is evident from the figure, the estimated policy response of both GDP and CPI is similar irrespective of which shock series we use. This is of course expected, considering the high correlation over 0.8 between the various shocks. The response of GDP differs somewhat for the PC shocks after ten quarters. During the initial quarters, the effects are similar. Most importantly, the differences are small, and they all fall within the 95 percent confidence bound from the original R&R shocks.[24] Although this is not strictly a test of the effects being statistically the same, it is a clear indication that we would not draw different policy conclusions from the respective policy shock series. Both

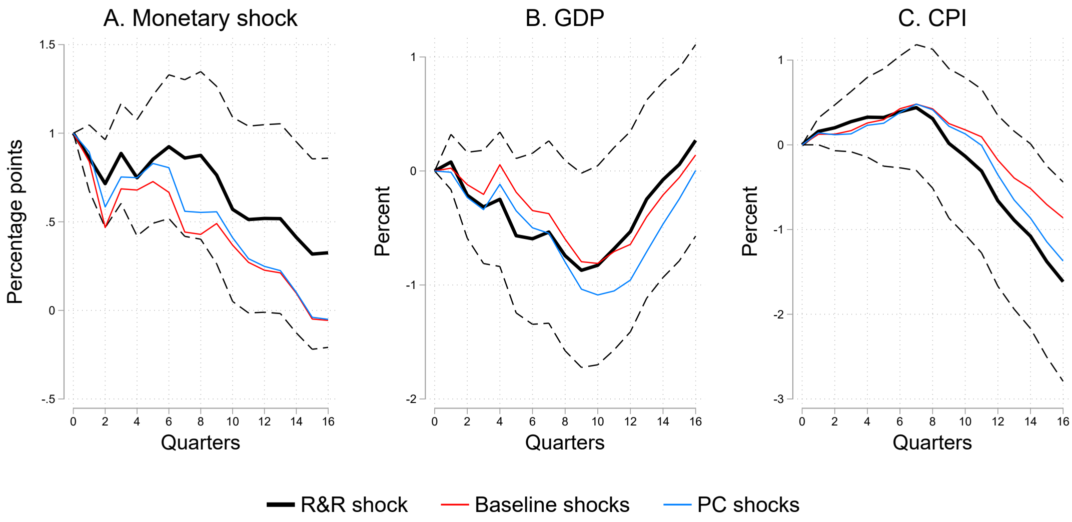
Karadi (2015) and Caldara and Herbst (2019)).

[23] That prices initially increase as a response to the contractionary policy shock goes against the standard theory. However, this result is not unusual in the VAR literature and is sometimes referred to as the "price puzzle" (Sims, 1992). A common explanation for this puzzle is model misspecification, such as missing variables. The purpose of this paper is to compare the effect of the R&R shocks with our shocks. Since we find the same puzzle effect for all series, the puzzle does not arise from how we generate our shock series. A common solution to this problem is to include commodity prices in the VAR model. We have tried this approach, and it does mitigate the price puzzle effect, but it does not eliminate it. Hanson (2004) highlights that both the mitigation of the price puzzle and the responses of output are sensitive to the choice of the commodity price index. We have not found any consistent use of a particular commodity price index in the previous literature.

[24] The response to output and prices fall inside 68% confidence bands as well.

the magnitude of the estimated effects, and the turning points, are similar.[25]

Figure 3: Impulse responses to a monetary policy shock



Notes: The figure compares impulse responses of the monetary shocks (left), output measured with the log of real GDP (middle) and the price level measured with the log of CPI (right) to a one standard deviation impulse in the monetary shock (normalized to start at one). Data are from 1969q1-2008q4. Dashed lines are 95% confidence bands.

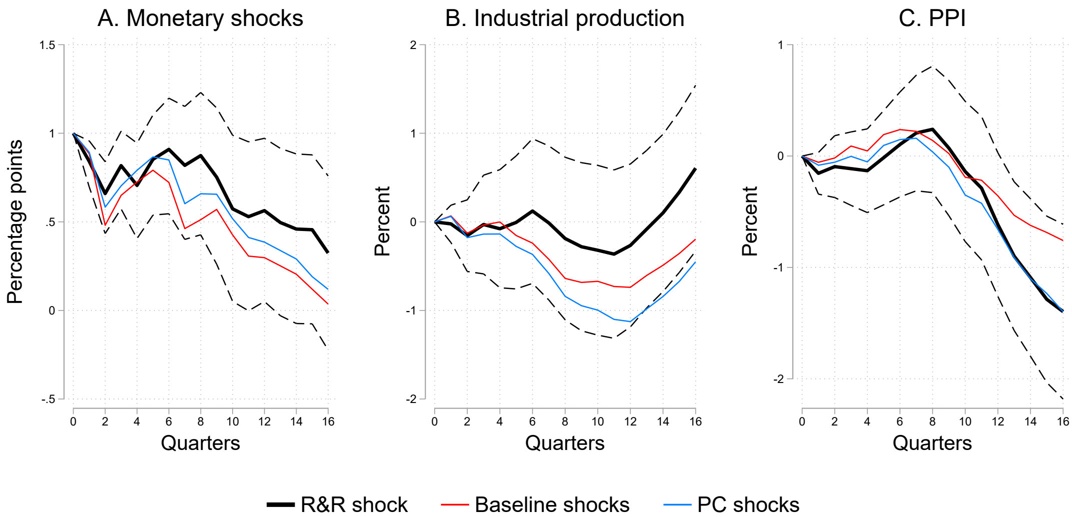
Figure 4 presents impulse responses from an additional VAR model where we measure output with the log of industrial production and the price level with the log of the PPI.[26] In this case, there are larger differences in how the economy responds to our shocks compared to the original R&R shocks. For PPI responses, all differences fall within the 95 percent confidence bounds and the turning points are similar. The response to output is stronger for our shocks and turning points do not always coincide. The PC shocks even fall outside the confidence bands after 13 quarters.[27] When comparing the baseline shocks with the PC shocks, we find that the effects are similar overall.

[25] The results are similar if we measure the price level with the log of the PCE index instead of the CPI index in the VAR model.

[26] The producer price index on finished goods is from the OECD.

[27] It could be the case that the approximation error, from including outcome data instead of forecast data, is exaggerated when we cumulate the monetary shocks. We, therefore, run the VAR model with industrial production and the PPI with monetary shocks in changes instead of levels. Responses are more similar for output when comparing shocks, but less similar for prices. However, they all lie within the confidence bands.

Figure 4: Impulse responses to a monetary policy shock with IP and PPI



Notes: The figure compares impulse responses of the monetary shocks (left), output measured with the log of industrial production (middle) and the price level measured with the log of PPI (right) to a one standard deviation impulse in the monetary shock (normalized to start at one). Data are from 1969q1-2008q4. Dashed lines are 95% confidence bands.

In summation, we find that the estimated effects on output and prices are similar irrespective of shock series in terms of the direction and the magnitude of the effect. These results hold especially for the estimated effects during the first six to eight quarters. For longer periods, the effects become more uncertain and there are some larger differences between the estimated effects from our shock estimates and the estimated effects from the R&R shocks. However, the differences fall within the 95 percent confidence bounds of the R&R shocks, whereby we would draw similar policy conclusions irrespective of the shock series. The main difference between our shocks and the R&R shocks are shown for responses of industrial production. However, this is a less important variable compared to GDP, given the relatively small size of the manufacturing, mining, and utility industries in the US economy (Yuskavage and Pho, 2004).

4 Conclusion

All estimates of monetary policy shocks are uncertain and contain a measurement error that affects the estimation of the impact of monetary policy on the economy. The size and nature of the measurement error is difficult to judge theoretically. The aim should always be to minimize the error by using data that comes as close as possible to the actual data used by the central bank to set the interest rate, and a method that comes close to capturing the reaction function of the central bank. For regression-based methods, this implies using real-time and forecast data as in Romer and Romer (2004). However, such data is sometimes difficult to obtain in a historical setting and it is not uncommon for studies to rely on revised data instead.

We explore the effect of using revised outcome data rather than real-time and forecast data, by estimating monetary policy shocks for the United States from 1969 to 2008 using a similar two-step approach as in R&R. We find that the differences in shock series are generally small. Most importantly, the estimated effects of a monetary policy shock on the economy are similar in terms of the sign of the effect, the magnitude, and turning points when the effect increases or decreases. These results suggest that it is possible to use revised data rather than real-time and forecast data when such data is not available. However, the result should not be exaggerated. Whether the same result holds for other countries is uncertain. Our results only cover the United States. Furthermore, the fact that revised data yield similar estimates of the effects of monetary policy on the economy could also be seen as an indication that the R&R method does a poor job in identifying the true policy shocks. Whether this is the case is impossible to evaluate as it would require access to the "true" policy shocks, which are unobservable.

The possibility of using revised data when real-time and forecast data is not available opens up the possibility to apply the R&R method more widely as complementary to other methods. However, we conclude that the effect of monetary policy should be explored using as many different approaches as possible.

References

Aikman, D., Bush, O., and Taylor, A. M. (2018). Monetary Versus Macroprudential Policies: Causal Impacts of Interest Rates and Credit Controls in the Era of the UK Radcliffe Report. *NBER Working Paper Series No. 22380.*

Barakchian, S. M. and Crowe, C. (2013). Monetary Policy Matters: Evidence From New Shocks Data. *Journal of Monetary Economics*, 60:950–966.

Bernanke, B. S., Boivin, J., and Eliasz, P. (2005). Measuring the Effects of Monetary Policy: A Factor-Augmented Vector Autoregressive (FAVAR) Approach. *The Quarterly Journal of Economics*, (1):387–422.

Bernanke, B. S. and Mihov, I. (1998). Measuring Monetary Policy. *The Quarterly Journal of Economics*, 113(3):869–902.

Blinder, A. S. and Watson, M. W. (2016). Presidents and the US Economy: An Econometric Exploration. *American Economic Review*, 106(4):1015–1045.

Caldara, D. and Herbst, E. (2019). Monetary Policy, Real Activity, and Credit Spreads: Evidence from Bayesian Proxy SVARs. *American Economic Journal: Macroeconomics*, 11(1):157–192.

Cesa-Bianchi, A., Thwaites, G., and Vicondoa, A. (2020). Monetary Policy Transmission in the United Kingdom: A High Frequency Identification Approach. *European Economic Review*, 123.

Christiano, L. J., Eichenbaum, M., and Evans, C. L. (1999). Chapter 2: Monetary Policy Shocks: What Have We Learned and To What End? Part A in Handbook of Macroeconomics, pages 65 – 148.

Cloyne, J. (2013). Discretionary Tax Changes and the Macroeconomy: New Narrative Evidence from the United Kingdom. *The American Economic Review*, 103(4):1507–1528.

Cloyne, J. and Hürtgen, P. (2016). The Macroeconomic Effects of Monetary Policy: A New Measure for the United Kingdom. *American Economic Journal: Macroeconomics*, 8(4):75–102. 75.

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

Coibion, O., Gorodnichenko, Y., Kueng, L., and Silvia, J. (2017). Innocent Bystanders? Monetary Policy and Inequality. *Journal of Monetary Economics*, 88:70–89.

Croushore, D. and Evans, C. L. (2006). Data Revisions and the Identification of Monetary Policy Shocks. *Journal of Monetary Economics*, 53(6):1135–1160.

Doniger, C. L. (2019). Do Greasy Wheels Curb Inequality? *FED Finance and Economics Discussion Series 2019-021.*

Friedman, M. and Schwartz, A. J. (1963). *A Monetary History of the United States, 1867-1960.* National Bureau of Economic Research Publications. Princeton University Press.

Gertler, M. and Karadi, P. (2015). Monetary Policy Surprises, Credit Costs, and Economic Activity. *American Economic Journal: Macroeconomics*, 7(1):44–76.

Hanson, M. S. (2004). The 'Price Puzzle' Reconsidered. *Journal of Monetary Economics*, 51(7):1385–1413.

Hoesch, L., Rossi, B., and Sekhposyan, T. (2020). Has the Information Channel of Monetary Policy Disappeared? Revisiting the Empirical Evidence. *Federal Reserve Bank of San Francisco Working Paper 2020-08.*

Holm, M. B., Paul, P., and Tischbirek, A. (2021). The Transmission of Monetary Policy Under the Microscope. *Journal of Political Economy*, 129(10):2861–2904.

Kim, S. and Lim, K. (2018). Effects of Monetary Policy Shocks on Exchange Rate in Small Open Economies. *Journal of Macroeconomics*, 56:324–339.

Kuttner, K. N. (2001). Monetary Policy Surprises and Interest Rates: Evidence from the Fed Funds Futures Market. *Journal of Monetary Economics*, 47(3):523–544.

Leahy, J. V. and Thapar, A. (2019). Demographic Effects on the Impact of Monetary Policy. *NBER Working Paper Series No. 26324*.

Lennard, J. (2018). Did Monetary Policy Matter? Narrative Evidence From the Classical Gold Standard. *Explorations in Economic History*, 68:16–36. 16.

Mertens, K. and Ravn, M. O. (2013). The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States. *The American Economic Review*, 103(4):1212–1247.

Miranda-Agrippino, S. and Rey, H. (2020). U.S. Monetary Policy and the Global Financial Cycle. *The Review of Economic Studies*, 87(6):2754–2776.

Miranda-Agrippino, S. and Ricco, G. (2021). The Transmission of Monetary Policy Shocks. *American Economic Journal: Macroeconomics*, 13(3):74–107.

Molodtsova, T., Nikolsko-Rzhevskyy, A., and Papell, D. H. (2008). Taylor Rules With Real-Time Data: A Tale of Two Countries and One Exchange Rate. *Journal of Monetary Economics*, 55:S63–S79.

Nakamura, E. and Steinsson, J. (2018). High-Frequency Identification of Monetary Non-neutrality: The Information Effect. *Quarterly Journal of Economics*, 133(3):1283–1330.

Orphanides, A. (2001). Monetary Policy Rules Based on Real-Time Data. *The American Economic Review*, 91(4):964–985.

Orphanides, A. (2003). Historical Monetary Policy Analysis and the Taylor Rule. *Journal of Monetary Economics*, 50(5):983–1022.

Ramey, V. (2016). *Macroeconomic Shocks and Their Propagation*, volume 2 of *Handbook of Macroeconomics*. Elsevier B.V.

Romer, C. D. and Romer, D. H. (1989). Does Monetary Policy Matter? A New Test in the Spirit of Friedman and Schwartz. *NBER Macroeconomics Annual*, 4:121–170.

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

Schmitt-Groh, S. and Uribe, M. (2012). What's News in Business Cycles. *Econometrica*, 80(6):2733–2764.

Sims, C. A. (1992). Interpreting the Macroeconomic Time Series Facts: The Effects of Monetary Policy. *European Economic Review*, 36(5):975–1000.

Smets, F. and Wouters, R. (2007). Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach. *The American Economic Review*, 97(3):586–606.

Stock, J. H. and Watson, M. W. (2002). Macroeconomic Forecasting Using Diffusion Indexes. *Journal of Business & Economic Statistics*, 20(2):147–162.

Stock, J. H. and Watson, M. W. (2012). Disentangling the Channels of the 200709 Recession. *Brookings Papers on Economic Activity*, pages 81–135.

Tenreyro, S. and Thwaites, G. (2016). Pushing on a String: US Monetary Policy Is Less Powerful in Recessions. *American Economic Journal: Macroeconomics*, 8(4):43–74.

Yuskavage, R. E. and Pho, Y. H. (2004). Gross Domestic Product by Industry for 1987-2000 - New Estimates on the North American Industry Classification System. Technical report.

Appendix

Appendix A. Principal component analysis

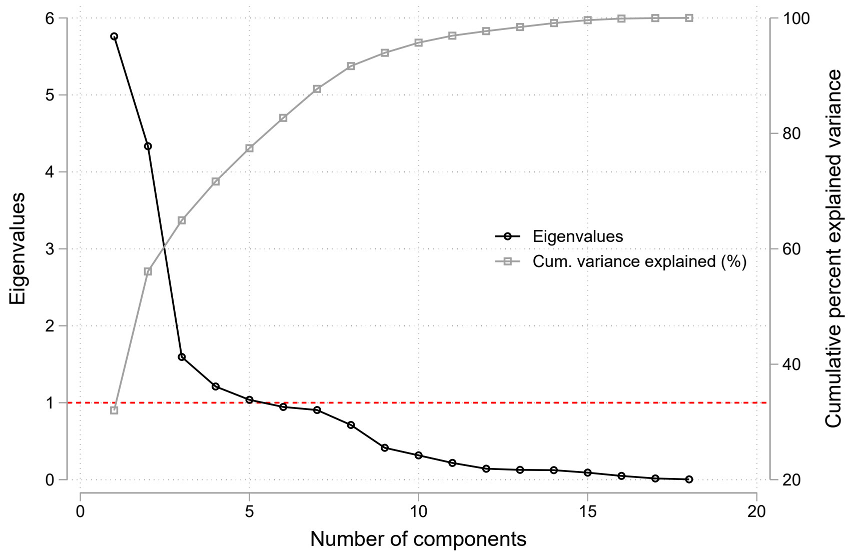
The Appendix Table A1 presents the data used to estimate the PC shocks. The Appendix Figure A1 illustrates the eigenvalues and the cumulative percentage of the variance explained by the common components. The first two common components are the most important. To ensure that we adequately capture the variance in the dataset, we also include the third and fourth components in the shock estimation. These first four components explain 72 percent of the total variance in the data. Results are close to identical if we include two or three components. The Appendix Table A2 presents loadings of each common component. The first two components are most clear where the first component captures variables connected to the real economy such as real GDP, industrial production, and unemployment. The second component captures variables connected to price indices, such as CPI, GDP deflator, and core inflation. The third component mainly captures the variable on new housing units, while the fourth component mainly captures personal consumption expenditures and government consumption and investment.

Table A1: Variables included in the principal components

Variable name	Source	Transformation
GDP (million USD), real	BEA	2
GDP (PI)	BEA	2
Corporate profits after tax (million USD)	BEA	2
Unemployment rate (%)	BLS	1
Employment, nonfarm industries (thousands)	BLS	2
Industrial production (index)	FRED	2
Housing started (thousands)	UCB	2
PCE (million USD), real	BEA	2
PCE (PI)	BEA	2
PCE core (PI)	BEA	2
Net exports of goods & services (% of GDP)	BEA	1
Changes in inventories (% of GDP)	BEA	1
Gov. consumption & gross investment, state and local (million USD), real	BEA	2
Gov. consumption & gross investment, federal (million USD), real	BEA	2
Private domestic investment, residential (million USD), real	BEA	2
Private domestic investment, nonresidential (million USD), real	BEA	2
CPI, all items	BLS	2
CPI, core	BLS	2

Notes: The table includes a list of all the variables included in the principal component analysis, together with information on their sources and transformations. The transformation codes are 1 = first difference and 2 = first difference of logarithms. All relevant variables have been transformed into real values using base date 2010q1. Abbreviations for sources are the following: BEA = US Bureau of Economic Analysis, BLS = US Bureau of Labor Statistics, FRED = Federal Reserve Economic Data, and UCB = US Census Bureau. PI stands for price index.

Figure A1: Plot on eigenvalues and the cumulative percentage explained for each component's variance



Notes: The first eigenvalue explains 32% of the total variance, the second explains 24%, the third 9%, and the fourth 7%. Cumulative, the first four components explain 72% of the total variance.

Table A2: Loadings of the principal components

Variables	**Comp. 1: GDP and employment**	**Comp. 2: Price indices**	**Comp. 3: Housing**	**Comp. 4: Private and gov. cons.**
GDP, real	0,332	0,184	0,055	0,068
GDP (PI)	-0,238	0,341	0,197	0,112
Corporate profits after tax	0,186	0,236	0,081	-0,314
Employment, full- and part-time	0,266	0,279	-0,183	0,249
Industrial production	0,336	0,221	-0,104	0,059
New private housing units started	0,197	0,002	0,573	-0,183
PCE (PI)	-0,179	0,390	0,017	-0,122
PCE, real	0,271	-0,128	0,336	0,265
PCE, core (PI)	-0,229	0,329	0,270	0,110
Gov. cons. & invest., state & local, real	0,074	-0,002	0,214	0,282
Gov. cons. & invest., federal, real	-0,025	-0,013	0,001	-0,441
Gross fixed invest., residential, real	0,295	0,040	0,434	-0,153
Gross fixed invest., non-res., real	0,260	0,219	-0,191	0,260
CPI	-0,222	0,390	0,017	-0,024
CPI, core	-0,248	0,324	0,099	0,131
Unemployment rate	-0,331	-0,185	0,184	-0,051
Net exports	-0,141	-0,150	0,070	0,509
Changes in inventories	0,136	0,169	-0,271	-0,210

Notes: PI stands for price index.