

**Finance and Economics Discussion Series  
Divisions of Research & Statistics and Monetary Affairs  
Federal Reserve Board, Washington, D.C.**

**Identification of Monetary Policy Shocks with External  
Instrument SVAR**

**Kyungmin Kim**

**2017-113**

Please cite this paper as:

Kim, Kyungmin (2017). "Identification of Monetary Policy Shocks with External Instrument SVAR," Finance and Economics Discussion Series 2017-113. Washington: Board of Governors of the Federal Reserve System, <https://doi.org/10.17016/FEDS.2017.113>.

NOTE: Staff working papers in the Finance and Economics Discussion Series (FEDS) are preliminary materials circulated to stimulate discussion and critical comment. The analysis and conclusions set forth are those of the authors and do not indicate concurrence by other members of the research staff or the Board of Governors. References in publications to the Finance and Economics Discussion Series (other than acknowledgement) should be cleared with the author(s) to protect the tentative character of these papers.

# Identification of Monetary Policy Shocks with External Instrument SVAR

Kyungmin Kim\*

Federal Reserve Board

November 2, 2017

## **Abstract**

We explore the use of external instrument SVAR to identify monetary policy shocks. We identify a forward guidance shock as the monetary shock component having zero instant impact on the policy rate. A contractionary forward guidance shock raises both future output and price level, stressing the relative importance of revealing policymakers' view on future output and price level over committing to a policy stance. We also decompose non-monetary structural shocks, and find that positive shocks to output and price level lead to monetary contraction. Since information on output and price level is revealed through both monetary and non-monetary channels, some monetary and non-monetary shocks can look alike, leading to linear dependence and violating usual instrument SVAR assumptions. We show that some of the main findings are robust to such dependence.

---

\*E-mail: [kyungmin.kim@frb.gov](mailto:kyungmin.kim@frb.gov). I thank Arsenios Skaperdas for helpful comments. The paper expresses solely my own views, not those of the Federal Reserve Board or the Federal Reserve System.

# 1 Introduction

We explore the use of external instrument structural vector autoregression (instrument VAR) to identify structural shock components correlated with monetary policy announcements. Following Gürkaynak et al. (2005) and Gertler and Karadi (2015), we use price changes of selected interest rate futures contracts around monetary policy announcements as instruments to capture the response of the economy to monetary policy.

We are not the first to identify monetary policy shocks using instrument VAR, but we explore the methodology in depth to fully understand the implications of the model and its potential issues. We start by describing instrument VAR in section 2. Pioneered by Stock and Watson (2012) and Mertens and Ravn (2013), instrument VAR uses the covariance between VAR residuals and external instruments to identify structural shocks correlated with given instruments. This method is useful if instruments are available and other conventional identifying restrictions such as the ‘order’ condition (see Sims (1980) and Watson (1994)) are not applicable.

We present instrument VAR primarily as a matrix operation, rather than as two-stage least squares (2SLS), which is the norm in the literature following Mertens and Ravn (2013). The result is, of course, equivalent, but we believe that our presentation will be more transparent to some readers. Also, it directly leads to a full characterization of the instrument VAR parameter space and our identification strategy. From our presentation, we can clearly see that there is no need to choose variables instrumented by the instruments, which is sometimes misunderstood, and the instruments work for the whole VAR system.

In section 3, we discuss why it is reasonable to treat monetary shocks as two-dimensional shocks in a VAR setup, consistent with Gürkaynak et al. (2005). We discuss two distinct approaches. First, we show that we can reject the hypothesis of weak instruments, using the tests by Stock and Yogo (2005) and Stock et al. (2002). We use weak instrument tests following Mertens and Ravn (2013) and Gertler and Karadi (2015), with the difference that we use a test for two instrumented variables rather than one. Second, we directly test the rank of the covariance matrix between instruments and VAR residuals, as it should theoretically equal the number of structural shocks correlated with the instruments. We use the test developed by Kleibergen and Paap (2006), and normalize the statistic so that it is invariant to linear transformations of VAR variables or instruments. We find that the test

does not strongly reject the hypothesis of one-dimensional monetary structural shock, and thus conclude that both one- and two-dimensional shocks are reasonable choices.

We uniquely identify the forward guidance shock as the component of the two-dimensional shocks that does not affect the current policy rate, which is the federal funds rate. The residual component is defined as the policy rate shock. For our VAR model, we use the simplest possible one with four variables. This is a simpler version of the six-variable baseline VAR model used in Gertler and Karadi (2015). Two of the four variables are industrial production and consumer price index (CPI), following the standard practice of including measures of output and price level in a macroeconomic VAR model. The third variable is the federal funds rate, which acts as the policy rate. We need to use specifically the federal funds rate, not a rate of longer tenor such as the one-year Treasury rate, so that we can define the forward guidance shock as the monetary shock component that has zero instant impact on the policy rate. The fourth variable is the ‘GZ’ excess bond premium (average excess premium on corporate bonds, due to Gilchrist and Zakrajšek (2012)), which we include to capture the financial market’s response to monetary policy announcements. We follow Gertler and Karadi (2015) in including this variable in our VAR, and find that it works better than other commonly used financial spreads such as mortgage and commercial paper (CP) spreads.<sup>1</sup>

The impulse response of the federal funds rate to forward guidance shocks shows that the peak response occurs with about two years of delay. Also, we find that a contractionary forward guidance shock results in an increase, not a decrease, of future output and price level, and has little instant impact on either of them. This outcome is consistent with the idea that forward guidance reveals information about expected future paths of the economy (for example, see Campbell et al. (2012), Campbell et al. (2016), and Nakamura and Steinsson (2013)), rather than changes in future policy stance. For forward guidance, the impact of revealed information seems to dominate the impact of revealed future policy stance. In contrast, a contractionary policy rate shock results in a downward path for future output and price level, as predicted by standard macroeconomic theory.

The excess bond premium increases in response to a contractionary policy rate shock, which is the expected response to tightening credit supply. However, it decreases in response to a contractionary forward guidance. This is again consistent with the idea that a con-

---

<sup>1</sup>Better in the sense of narrower confidence bands.

tractionary forward guidance reveals a positive view about the future of the economy, with increasing output and price level.

In section 4, we discuss novel extensions of the baseline VAR model. First, we decompose non-monetary structural shocks, uncorrelated with monetary policy announcements. We find that non-monetary shocks representing increases in output or price level have a contractionary impact on the policy rate, consistent with how monetary policy is generally expected to respond to such shocks.

Second, we discuss the possible existence of linear dependence between monetary and non-monetary structural shocks. Since information channel can exist for both monetary and non-monetary shocks, certain monetary policy shocks can look like non-monetary shocks, implying a linear dependence between them. Alternatively, one may believe that given  $n$  VAR variables with just one rate variable, there should be  $n - 1$  non-monetary structural shocks, which make the total number of shocks be  $n + 1$ , with two monetary shocks. These violate usual assumptions of instrument VAR. However, we show that some of our results are robust to such problems. Moreover, we show exactly what we can infer about impulse responses to monetary structural shocks even when these problems exist.

Third, we relate our results to those of two recent studies identifying monetary shocks using instrument VAR, Gertler and Karadi (2015) and Lakdawala (2017). Gertler and Karadi (2015) uses a single structural shock to capture monetary policy shocks, and Lakdawala (2017) uses two shocks to identify forward guidance shocks, using identification restrictions different from ours. We discuss how to relate studies using different identifying restrictions and different numbers of instruments, and find a very close agreement between Gertler and Karadi (2015) and our paper, but no such agreement between Lakdawala (2017) and ours. Most interestingly, we describe how we can interpret the result of instrument VAR with just one instrument, when there are indeed two structural shocks correlated with the instruments.

The remainder of this paper is divided into four sections: Section 2 describes instrument VAR methodology and our identification strategy. Section 3 describes the baseline VAR model and the impulse responses to monetary policy shocks. Section 4 discusses extensions of the baseline VAR model. Section 5 concludes.

## 2 Identification with External Instruments

We first explain how instrument VAR works. Omitted technical details are in the appendix. We explicitly construct parameter space consistent with data and restrictions from instruments, and it directly leads to our identification strategy used later in the paper.

A simple reduced-form SVAR has the following well-known form:<sup>2</sup>

$$y_t = \sum_{j=1}^l B_j y_{t-j} + u_t. \quad (1)$$

$y_t$  is an  $n$ -dimensional column vector of VAR variables and  $B_j$  is an  $n$ -by- $n$  matrix of coefficients.  $u_t$  is the reduced-form residual in the form of  $u_t = B\epsilon_t$ , where  $B$  is an  $n$ -by- $n$  nonsingular matrix.  $\epsilon_t$  is an i. i. d. structural shock, which follows a multinomial normal distribution of mean zero and variance  $I_n$ , where  $I_n$  is the  $n$ -by- $n$  identity matrix.

The matrix  $B$  is key to identification, and satisfies  $BB' = \Sigma$ , where  $\Sigma$  is the covariance matrix of VAR residuals. Any  $n$ -by- $n$  matrices  $B_1$  and  $B_2$  satisfying  $B_1B_1' = B_2B_2' = \Sigma$  are rotations of each other, in the sense that  $B_1 = B_2R$  for an  $n$ -by- $n$  orthonormal matrix  $R$ . Structural shocks are identified by choosing a particular rotation  $R$ .

In the external instrument VAR model, only a subset of structural shocks are correlated with given instruments, and the rest are uncorrelated. This places an extra restriction on the choice of the matrix  $B$ . Formally, let  $Z_u$  be the covariance between VAR residuals and instruments,  $Z_u \equiv E[z_t u_t]$ , where  $z_t$  is an  $m$ -dimensional column vector of instruments. The instruments should satisfy the following conditions:

$$\text{rank}(E[z_t \epsilon_{1,t}]) = k. \quad (2)$$

$$E[z_t \epsilon_{2,t}] = 0. \quad (3)$$

$k$  is the number of structural shock components correlated with the instruments. Without loss of generality, the first  $k$  elements of  $\epsilon_t$  are those shock components correlated with the instruments.  $\epsilon_{1,t}$  denotes the first  $k$  elements of  $\epsilon_t$ , and  $\epsilon_{2,t}$  denotes its remaining  $n - k$  elements.

---

<sup>2</sup>Constants are omitted from the expression, even though they are included in actual regressions. This notational convention is typical in the literature.

Note that these can also be regarded as rules of constructing structural shocks from the data. The reason is that it is always possible to define  $n$  structural shocks so that  $\text{rank}(E[z_t\epsilon'_{1,t}]) = k$  and  $E[z_t\epsilon'_{2,t}] = 0$  if  $\text{rank}(Z_u) = k$ . Equivalently, the last  $n - k$  columns of the covariance matrix between the instruments and structural shocks,

$$E[z_t\epsilon'_t] = E[z_tu'_t(B^{-1})'] = Z_u(B^{-1})', \quad (4)$$

are zero.

We can easily characterize the entire set of matrices  $B$  that are consistent with the restrictions coming from the instruments. Any  $B_1$  and  $B_2$  such that (i)  $B_1B'_1 = B_2B'_2 = \Sigma$ , and (ii) the last  $n - k$  columns of both  $Z_u(B_1^{-1})'$  and  $Z_u(B_2^{-1})'$  are zero, are related as follows:

$$B_1 = B_2 \begin{bmatrix} R_k & 0 \\ 0 & R_{n-k} \end{bmatrix}, \quad (5)$$

where  $R_l$  is an  $l$ -by- $l$  orthogonal matrix. For simplicity, we denote by  $R(k, n - k)$  the large matrix composed of  $R_k$  and  $R_{n-k}$ .

Moreover, given any  $B_2$  such that (i)  $B_2B'_2 = \Sigma$ , and (ii) the last  $n - k$  columns of  $Z_u(B_2^{-1})'$  are zero,  $B_1 = B_2R(k, n - k)$  satisfies both of these conditions as well. Therefore, the relationship  $B_1 = B_2R(k, n - k)$  completely characterizes the space of  $B$  consistent with given  $\Sigma$  and  $Z_u$  for instrument VAR.

A simple procedure for actually performing instrument VAR follows directly from the characterization of  $B$ . First, a matrix  $A$  satisfying  $AA' = \Sigma$  can be found, for example, via Cholesky decomposition. Then, we can always find an orthonormal matrix  $R$  such that the last  $n - k$  columns of  $Z_u((AR)^{-1})' = Z_u(A^{-1})'R$  are zero. We define  $B = AR$ . Note that these steps do not involve choosing  $k$  VAR variables that are supposedly correlated with the instruments. It is sometimes misunderstood that such a choice is necessary in performing instrument VAR.

To follow a common practice in the literature deriving from Mertens and Ravn (2013), we can also write  $B$  in the following block form, even though it is unnecessary, as discussed:

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}. \quad (6)$$

$B_{11}$  is a  $k$ -by- $k$  matrix, and  $B_{12}, B_{21}$  and  $B_{22}$  are  $k$ -by- $(n-k)$ ,  $(n-k)$ -by- $k$ , and  $(n-k)$ -by- $(n-k)$ , respectively.  $B_{11}$  and  $B_{22}$  are nonsingular without loss of generality.<sup>3</sup>

Our characterization of the parameter space implies that for any  $B$  consistent with data and instruments,  $B_{11}B'_{11}$  and  $B_{22}B'_{22}$  are constant. This is because for any  $BR(k, n-k)$ , the diagonal blocks are  $B_{11}R_k$  and  $B_{22}R_{n-k}$ . Once we choose  $B_{11}$  and  $B_{22}$  for  $B_{11}B'_{11}$  and  $B_{22}B'_{22}$  given by the data,  $B_{21}$  and  $B_{12}$  are uniquely determined.

This motivates the frequently used closed-form formulas to compute  $B$  in Mertens and Ravn (2013):  $B_{11}B'_{11}$ ,  $B_{22}B'_{22}$ ,  $B_{21}B_{11}^{-1}$  and  $B_{12}B_{22}^{-1}$  can be computed as functions of  $\Sigma$  and  $Z_u$ , and the exact formulas are reproduced in the appendix for quick reference.

In particular, we can make diagonal blocks of  $B$  into lower triangular matrices. Let  $L_1$  be the Cholesky decomposition of  $B_{11}B'_{11}$  and  $L_2$  be that of  $B_{22}B'_{22}$ . Then,

$$B = \begin{bmatrix} L_1 & (B_{12}B_{22}^{-1})L_2 \\ (B_{21}B_{11}^{-1})L_1 & L_2 \end{bmatrix}. \quad (7)$$

We will use this decomposition in identifying the forward guidance shock and in identifying non-monetary structural shocks. Any other  $B$  consistent with data and instruments has the following form:  $BR(k, n-k)$ . For example, note that we can impose order restriction on any  $k$  VAR variables among shocks correlated with instruments, not necessarily among the first  $k$ . Similarly, we can do so on any  $(n-k)$  VAR variables among shocks uncorrelated with instruments, not necessarily among the last  $(n-k)$ .

So far, we have allowed  $k = \text{rank}(Z_u)$  to be smaller than  $m$ , the number of instruments. In practice,  $Z_u$  generated from sample covariance will almost always have full row rank, as long as  $m \leq n$ . However, we might want to assume  $\text{rank}(Z_u) < m$  due to a theoretical reason or to the empirical observation that some linear combination of the rows in  $Z_u$  is close to zero. In such a case, a linear combination of instruments given by  $Mz_t$ , where  $M$  is  $k$ -by- $m$ , needs to be chosen, so that the resulting covariance matrix  $E[Mz_t u'_t] = MZ_u$  is a full-row-rank matrix. Both Mertens and Ravn (2013) and Gertler and Karadi (2015) use the projection of the first  $k$  estimated residuals  $\hat{u}_{1,t}$  onto the instruments  $z_t$  as  $Mz_t$ . This choice of  $M$  gives  $i$ -th row of  $B_{21}B_{11}^{-1}$  as a 2SLS estimator of regressing the  $i$ -th element of  $\hat{u}_{2,t}$  on  $\hat{u}_{1,t}$  projected onto  $z_t$ . We follow the same routine in constructing  $MZ_u$  in the next

---

<sup>3</sup>If  $B_{11}$  or  $B_{22}$  is singular, we can apply a nonsingular linear transformation to  $u_t$  in the form of  $Lu_t$  to make both  $B_{11}$  and  $B_{22}$  nonsingular.

section.<sup>4</sup>

## 3 Identifying Forward Guidance Shocks

### 3.1 Data

The data provided in the appendix of Gertler and Karadi (2015) are used in this paper, both for VAR variables and instruments. Our VAR model includes four variables, and two of them are the logs of industrial production and consumer price index, which represent output and price level, respectively. It is a standard practice to include measures of output and price level in macro VAR models.

We have the federal funds rate as the monetary policy indicator. Including the federal funds rate is essential to identify a forward guidance shock independent from a policy rate shock, because the federal funds rate is the policy rate. We also have the GZ excess bond premium as one of the VAR variables. Financial spreads reflect views on the future path of the economy, and their response to a forward guidance shock tells us how financial markets interpret signals about the future path of monetary policy. Compared to other financial spreads, GZ excess bond premium responds more strongly to monetary policy shocks, and thus we choose to include it as a representative financial spread.

For instruments, we use price changes on current-month federal funds futures and three-quarter-ahead eurodollar futures over a 30-minute window around monetary policy announcements (Federal Open Market Committee statements), which are two of the five intraday price change measures used in Gürkaynak et al. (2005). Price changes of current-month federal funds futures are scaled to reflect the change in the expected federal funds rate due to each monetary policy announcement.

We use monthly VAR variables, starting from 07/1979 and ending in 12/2008. We use data only up to 12/2008 because the federal funds rate had stayed near zero between 2009 and 2015, and it is likely that the relationship between the rate and the other VAR variables changed during the zero-lower-bound period. The instruments are available between 01/1990 and 12/2008, and the covariance matrix between the instruments and the VAR residuals

---

<sup>4</sup>In the baseline specification, we end up using only two instruments for two structural shocks, so there is no need to perform projection. Still, in evaluating potential choices of estimators, we consider their goodness in 2SLS.

over this period is used in our instrument VAR. Gertler and Karadi (2015) has converted the instruments into monthly variables by assigning price changes to both the current and the next month, weighted by the proportion of the 30-day period following announcements belonging to the current and the next calendar month, respectively.

We allow two-dimensional structural shocks, assuming  $\text{rank}(E[z_t u_t']) = 2$ . We noted that it is not necessary to pick two VAR variables that are supposedly correlated with the two instruments to perform instrument VAR. However, we still pick two VAR variables most strongly correlated with the instruments as our ‘instrumented’ variables to perform a test of weak instrument. Also, they are defined as the first two VAR variables for notational convenience. The federal funds rate and the GZ excess bond premium are chosen because they are correlated with the instruments most significantly.

### 3.2 Test of Weak Instruments

We can reject the hypothesis of weak instrument with our chosen instruments. Following Mertens and Ravn (2013) and Gertler and Karadi (2015), we reduce the dimension of instruments to the number of structural shocks,  $k = 2$ , by projecting the first  $k$  VAR residuals onto the instruments and using the resulting  $k$  variables as the instruments. As a result, the  $i$ -th row of  $B_{21}B_{11}^{-1}$  becomes a 2SLS estimator of regressing  $u_{k+i}$  on  $u_1, \dots, u_k$  projected onto the instruments.

We use the minimal eigenvalue of the concentration matrix,  $g_m$ , to quantify the strength of the instruments, following Stock and Yogo (2005) and Stock et al. (2002). This can be interpreted as the analogue of the well-known rule of thumb  $F > 10$  for the case of two instrumented variables.<sup>5</sup> There is no such widely used rule of thumb in the two-instrument case, but the appendix in Stock and Yogo (2005) suggests  $g_m > 8$  as a comparable threshold to  $F > 10$  for the case of two endogenous regressors.

Table 1 reports the minimal eigenvalue, as well as regression coefficients. In the first column, the coefficient on the change in current-month federal funds futures price (CM FF) is somewhat close to 1, as expected. In the second column, no coefficient is significant, but we see that the change in three-quarter-ahead eurodollar futures price (3Q ED) is relatively more important as an explanatory variable than CM FF, compared to the first column. This

---

<sup>5</sup>Gertler and Karadi (2015) uses this rule of thumb, for example.

may reflect the fact that the excess bond premium captured by GZ spread takes into account future paths of the economy.

Table 1 also reports the minimal eigenvalue and regression coefficients from using the full set of price change measures in Gürkaynak et al. (2005), for comparison. The three extra instruments are price changes in three-month-ahead federal funds futures (3M FF), two-quarter- (2Q ED) and four-quarter-ahead (4Q ED) eurodollar futures. The minimal eigenvalue shows that the full set of instruments is substantially weaker as 2SLS instruments.

### 3.3 Determining the Dimension of Structural Shocks

The assumption that monetary structural shocks are two-dimensional is motivated by well-known empirical studies on forward guidance, such as Gürkaynak et al. (2005) and Campbell et al. (2012). As we have discussed, the dimension of structural shocks correlated with the instruments is determined by the rank of  $E[z_t u_t']$ . In this section, we directly test the rank of this matrix with estimated reduced-form residuals, using ‘rk-statistic’ proposed by Kleibergen and Paap (2006). Especially, we are interested in testing whether the rank of this matrix is one or two. We cannot statistically reject the hypothesis that its rank is one, and thus conclude that there is a somewhat equal support for both ranks one and two.

The test is based on the idea that if the rank of an  $m$ -by- $n$  matrix is  $k$ , its smallest ( $\min(m, n) - k$ ) singular values are zero. Under the null hypothesis that the rank is  $k$ , rk-statistic asymptotically follows chi-squared distribution with  $(m - k)(n - k)$  degrees of freedom. If the value of this statistic is too large, the null is rejected in favor of the alternative hypothesis that the rank is greater than  $k$ . Given the form of the matrix  $E[z_t u_t']$ , rk-statistic can be normalized to become invariant to invertible linear transformations such as changing the order of variables or arbitrarily scaling them. We describe the exact normalization in the appendix.

We test whether the rank of the matrix  $E[z_t u_t']$  is zero, one or two, versus the alternative that the rank is greater. This can be done with either the full vectors  $z_t$  and  $u_t$ , or with subvectors of  $z_t$  and  $u_t$  and the corresponding submatrix of  $E[z_t u_t']$ . Mathematically, a submatrix has a rank that is weakly smaller than that of the full matrix, but this needs not to be the case for statistical tests of rank. Indeed, with rk-statistic, removing weakly correlated elements of  $z_t$  and  $u_t$  seems to make the statistic favor hypotheses of higher rank.

Dependent variable:	$u_1$ (Fed funds)	$u_2$ (GZ premium)	$u_1$	$u_2$
CM FF	1.21** (3.70)	0.05 (0.16)	1.02** (2.45)	-0.66 (-1.15)
3Q ED	-0.33 (-1.14)	0.50 (1.10)	-1.39 (-1.02)	-1.59 (-0.79)
3M FF	.	.	0.51 (1.07)	0.93* (1.65)
2Q ED	.	.	0.20 (0.18)	1.85 (1.21)
4Q ED	.	.	0.67 (0.85)	0.24 (0.25)
$R^2$	0.13	0.02	0.14	0.04
Number of obs.		228		228
Number of inst.		2		5
$g_m$		17.7		1.5

Numbers in ( ) are  $t$  statistics.

\*\* Significant at 5 percent.

\* Significant at 10 percent.

Table 1: Regression of Reduced-Form Shocks on Instruments

Included residuals:	All	Fed funds and GZ premium	Fed funds and GZ premium
Included instruments:	All	All	CM FF and 3Q ED
$H_0 : rank = 0$	0.32	0.04	0.05
$H_0 : rank = 1$	0.88	0.47	0.22
$H_0 : rank = 2$	0.92	.	.

Table 2: Tests of Matrix Rank

Table 2 lists the p-values of rk-statistics for a few different combinations of VAR residuals and instruments. With the full matrix  $E[z_t u_t']$ , it is not possible to reject even the hypothesis that its rank is zero. If we include only the first two VAR residuals, then we can reject the zero rank hypothesis at five percent with either all the five instruments or with the two chosen instruments, CM FF and 3Q ED. The hypothesis that the rank of the matrix is one cannot be rejected at conventional significance levels even if we use just the two chosen instruments.

### 3.4 Identifying Restrictions

We estimate the impulse response to the two-dimensional monetary policy shocks. Recall that the following matrix completely characterizes the parameter space of the instrument VAR:

$$B = \begin{bmatrix} L_1 R_k & (B_{12} B_{22}^{-1}) L_2 R_{n-k} \\ (B_{21} B_{11}^{-1}) L_1 R_k & L_2 R_{n-k} \end{bmatrix}. \quad (8)$$

In identifying the monetary policy shocks, we only care about the first  $k = 2$  columns of  $B$ , so we only need to determine  $R_k$ . We decompose the monetary policy shocks into a current policy rate shock ( $\epsilon_1$ ) and a forward guidance shock ( $\epsilon_2$ ). We define the forward guidance shock as the component of the monetary policy shock that has zero instant effect on the federal funds rate,  $y_1$ . In other words,  $B(1, 2) = 0$ , where  $B(i, j)$  denotes the  $i$ -th row and the  $j$ -th column of  $B$ .

This restriction determines the first and the second columns of  $B$  up to sign changes,

and we choose the signs so that a positive shock corresponds to an increase in interest rates. This means that we assume that  $B(1, 1)$  is positive. For  $B(2, 2)$ , we assume that it has a constant sign, and determine its sign to make the impulse response of the short rate,  $y_1$ , to the forward guidance shock generally positive. It turns out that assuming  $B(2, 2) < 0$  unambiguously achieves this purpose.

Alternatively, we can choose the sign of  $B(2, 2)$  so that the impulse response to the forward guidance shock is positive overall. This can be stated as follows:

$$\sum_{t=0, \dots, T} IRF_t > 0, \tag{9}$$

where  $IRF_t$  denotes the  $t$ -month impulse response of the federal funds rate,  $y_1$ , to the forward guidance shock. For reasonable choices of  $T$ , such as 24 or 48 months, this restriction turns out to be identical to  $B(2, 2) < 0$ , and produces similar confidence bands.<sup>6</sup>

### 3.5 Impulse Response to Monetary Policy Shocks

Figure 1 shows the impulse response of the four VAR variables to the two-dimensional monetary policy shocks. Plots on the left are responses to the policy rate shock, while those on the right are responses to the forward guidance shock. The dotted lines show 95-percent confidence intervals.<sup>7</sup> The impulse responses are responses to one-standard-deviation shocks.

The policy rate shock immediately affects the federal funds rate, raising it by about 0.3 percent, and converges to zero in about seven months. This is comparable to what we can typically get from order conditions, as in Christiano et al. (1999). The shock has a weak positive effect on the GZ excess bond premium. An unexpected increase in the interest rate can tighten credit supply, increasing the premium.

A contractionary policy shock also has generally negative effects on output and price level, showing no ‘price puzzle’. This contrasts with the positive responses of output and

---

<sup>6</sup>For alternative choices of instruments, the first restriction based on constant sign produces considerably narrower confidence bands. This gives another reason to prefer the chosen instruments over other possible choices.

<sup>7</sup>Following Mertens and Ravn (2013) and Gertler and Karadi (2015), we use so-called other-percentiles (see Sims and Zha (1999)) of simulated impulse responses with wild bootstrap resampling with multiplier of +1 or -1 with equal probability (see Wu (1986)). We use this method because it is well-known, but there are criticisms of this method, such as Jentsch and Lunsford (2016).

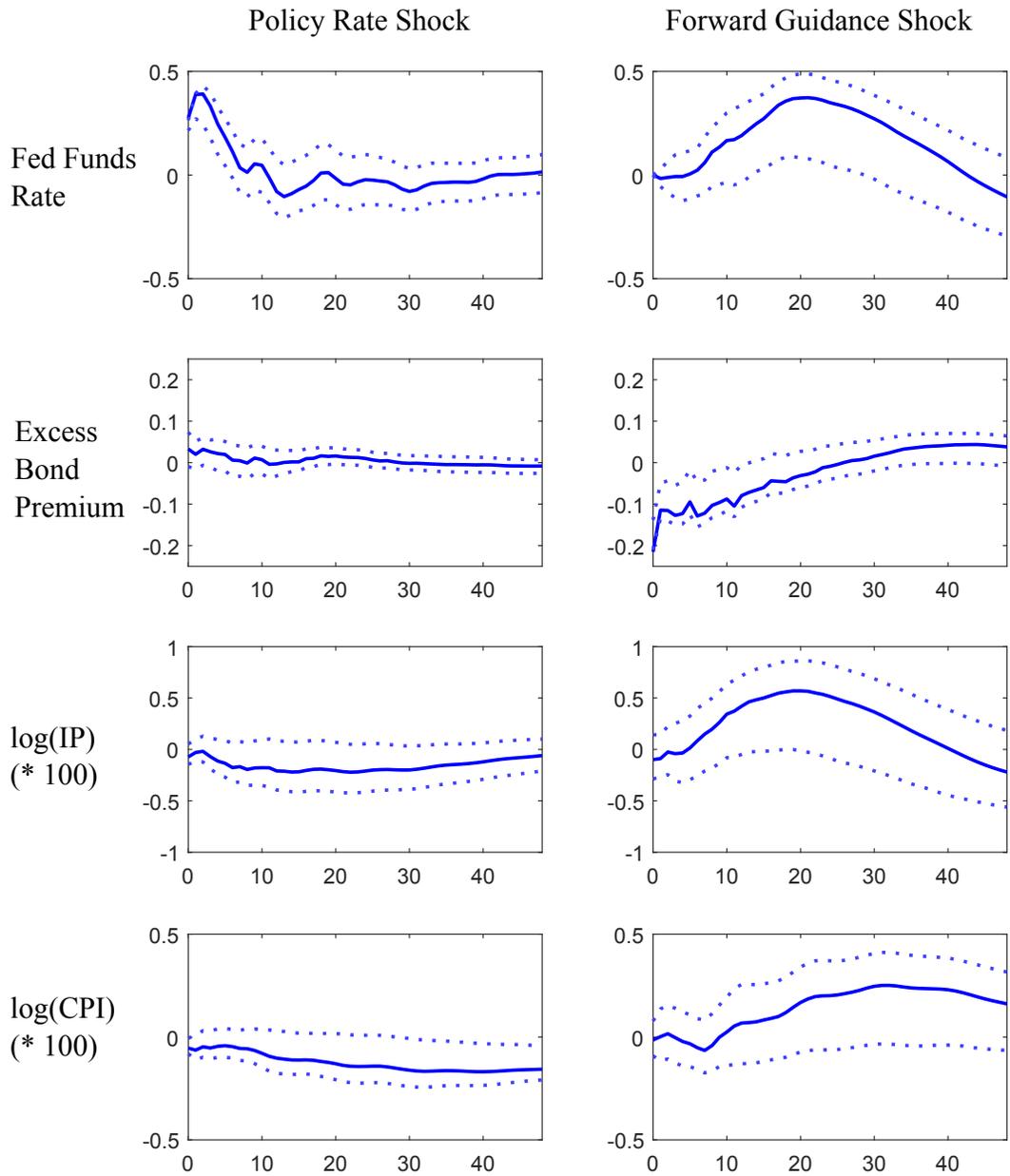


Figure 1: IRF to Monetary Policy Shocks

price level to a contractionary forward guidance shock. Therefore, the forward guidance shock can be interpreted as conveying an optimistic or pessimistic view on the future path of the economy, consistent with the findings of Campbell et al. (2012). It may partly reflect changes in the policy stance, but the net impact looks close to that of revealing information on the future path of the economy.

The forward guidance shock has zero effect on the federal funds rate initially ( $t = 0$ ) by construction. Its hump shape reaches its peak around 2 years after the initial shock, and its maximum magnitude is similar to that of the policy rate shock. A contractionary forward guidance shock is identified by its negative immediate impact on GZ excess bond premium. This is consistent with the interpretation that forward guidance signals optimism about the future of the economy, which can increase credit supply, and thus lower excess bond premium.

## 4 Discussion

### 4.1 Non-Monetary Shocks

In the baseline VAR model, two of the four shocks are correlated with the two instruments. The remaining two shocks are non-monetary shocks, as they are uncorrelated with monetary policy announcements. In section 2, we have shown how to identify the  $(n - k)$  shocks uncorrelated with the instruments. In particular, we are interested in the right-side blocks of the following matrix:

$$B = \begin{bmatrix} L_1 R_k & (B_{12} B_{22}^{-1}) L_2 R_{n-k} \\ (B_{21} B_{11}^{-1}) L_1 R_k & L_2 R_{n-k} \end{bmatrix}. \quad (10)$$

For identification, we place an order restriction between the two remaining shocks,  $\epsilon_3$  and  $\epsilon_4$ . We treat them as shocks to industrial production ( $\epsilon_3$ ) and price level ( $\epsilon_4$ ), and assume that price level moves more slowly than the industrial production, following standard convention.<sup>8</sup> Note that this places order between these two non-monetary shocks only, not between non-monetary and monetary shocks. In other words, we assume the shock to price level,  $\epsilon_4$ , does not affect production,  $y_3$ , at  $t = 0$ . Mathematically, we are choosing  $R_{n-k}$  to be an identity

---

<sup>8</sup>For example, Sims (1980) uses this ordering between price level and production.

matrix, so that the bottom-right block is lower-triangular.

Figure 2 shows the impulse response of the VAR model to non-monetary shocks. By construction, the initial impact of the price shock to production is zero. Also, it turns out that the initial impact of the production shock to price is also close to zero. Therefore, we would end up with similar impulse responses if we ordered price level before production.

Both shocks have positive initial impact on the federal funds rate. Since non-monetary shocks are orthogonal to monetary policy actions, their impact on the federal funds rate represents the expected response of policymakers to a surprise increase in production and price level. Conventional wisdom states that the expected response would be contractionary, which is what we find.

A positive shock to production has a weak negative impact on the GZ excess bond premium. This is consistent with our interpretation of the forward guidance shock: A contractionary forward guidance makes the market expect higher production and price level in the future, thus decreasing the excess bond premium.

The production shock has a stronger positive impact on the federal funds rate and a stronger negative impact on the GZ excess bond premium than the price shock. One reason may be that the production shock has a positive delayed impact on price level, while the price shock has a negative delayed impact on production. With both production and price level responding positively, it is not surprising that the production shock induces more contractionary monetary policy response and more relaxed credit supply than the price shock.

## 4.2 Dependent Monetary and Non-Monetary Shocks

Following the literature, we have assumed that the number of monetary structural shocks and the number of non-monetary structural shocks add up to  $n$ , the number of VAR variables. However, we can reasonably suspect that there may be some linear dependence between these shocks: Their numbers add up to a number greater than  $n$ . The shape of IRFs to the forward guidance shock suggests that forward guidance may work through revealing information about the future path of the economy. If that were the case, we can imagine a non-monetary shock that looks very similar to the forward guidance shock, because similar types of information can be generated outside monetary policy announcements. Also, many VAR models in the literature assume only a single monetary policy structural shock, and to

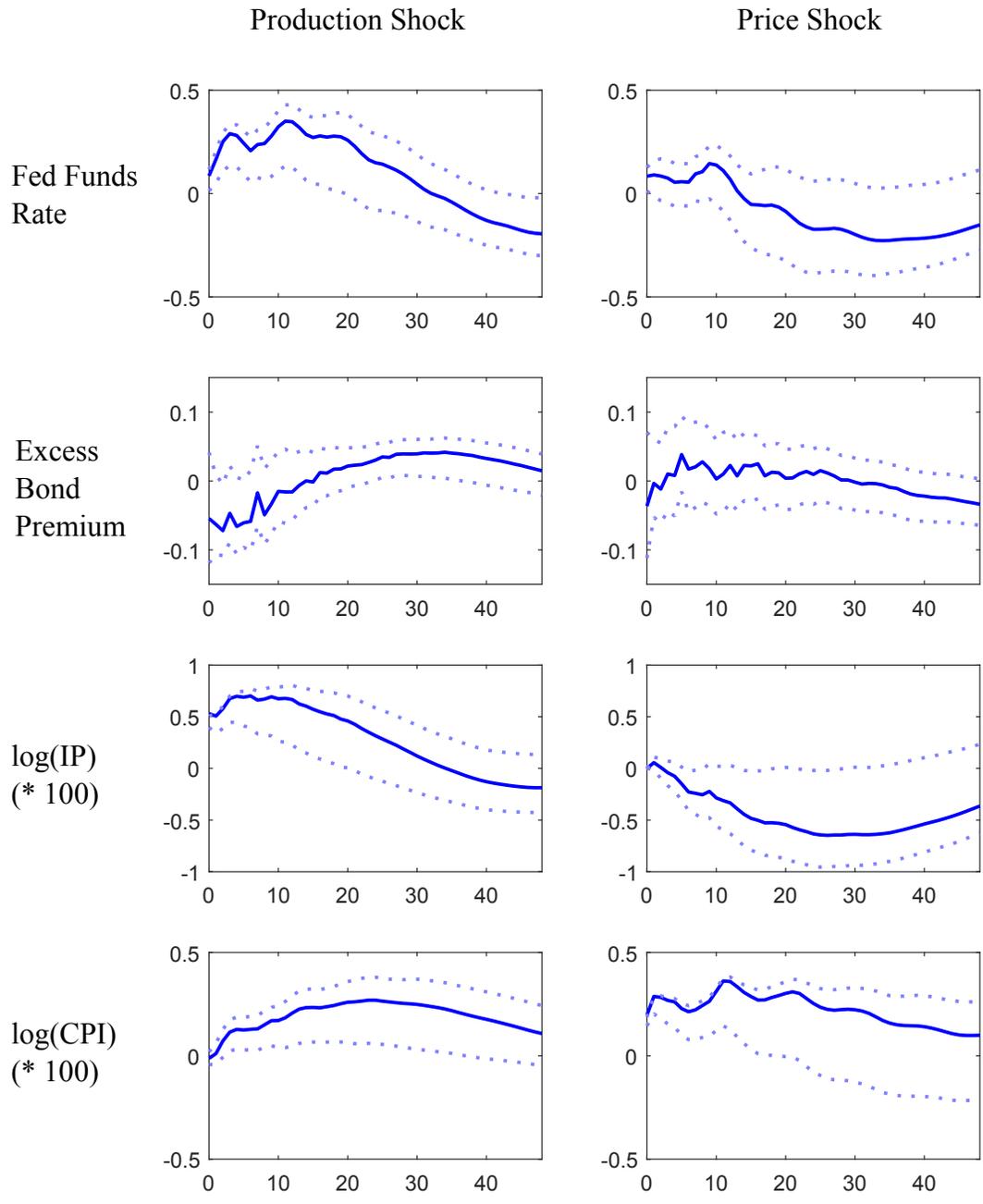


Figure 2: IRF to Non-Monetary Shocks

be compatible with those models, we need to allow  $(n - 1)$ -dimensional non-monetary shocks in addition to 2-dimensional monetary shocks, with a total of  $(n + 1)$  structural shocks.

We show that even with such problems, our baseline identification strategy produces correct shapes for IRFs to forward guidance shocks, and gives good information on the correct shapes of IRFs to policy rate shocks. However, we no longer can determine the magnitude of the shocks.

Formally, the model can be stated as follows:

$$u_t = B\epsilon_t. \tag{11}$$

$B$  is an  $n$ -by- $(k + l)$  matrix of rank  $n$ , and  $k + l > n$ .  $k$  is the number of structural shocks correlated with the  $k$  instruments,  $z_t$ , and  $l$  is the number of structural shocks uncorrelated with the instruments. Note that in our baseline model, we assume  $k + l = n$ .

$B$  has to satisfy the following condition:

$$BB' = E[u_t u_t'] = \Sigma. \tag{12}$$

In addition, there exists a matrix  $Z_\epsilon \equiv E[z_t \epsilon_t']$  that satisfies three conditions. First, with given  $E[z_t z_t']$  and  $E[\epsilon_t \epsilon_t'] = I_{k+l}$ ,  $Z_\epsilon$  forms a valid covariance matrix.<sup>9</sup> Second, the last  $l$  columns of  $Z_\epsilon$  are zero. Finally, it satisfies the following equation:

$$Z_\epsilon B' = E[z_t u_t'] = Z_u. \tag{13}$$

We can characterize  $B$  using the characterization for the baseline case of  $k + l = n$ . There exists an  $n$ -by- $n$  nonsingular matrix  $B_0$  that satisfies all the conditions on  $B$ , with  $l = n - k$  (See section 2). Then,  $B$  can be written as

$$B = B_0 N, \tag{14}$$

for an  $n$ -by- $(k + l)$  matrix  $N$ . In the appendix, we show that any  $N$  that is consistent with

---

<sup>9</sup>This was not a concern under the baseline model, because  $\epsilon_t = B^{-1}u_t$  and the validity of the covariance matrix  $Z_u$  imply that  $Z_\epsilon$  is valid. In the present case,  $B$  is no longer invertible.

the data ( $\Sigma$  and  $Z_u$ ) has the following form:

$$N = \begin{bmatrix} N_{11} & N_{12} \\ 0 & N_{22} \end{bmatrix}. \quad (15)$$

The  $n$  rows of  $N$  are orthonormal vectors in  $\mathbb{R}^{k+l}$ :  $NN' = I_n$ . In addition,  $k$ -by- $k$  matrix  $N_{11}$  is nonsingular.  $N_{12}$  is  $k$ -by- $l$  and  $N_{22}$  is  $(n - k)$ -by- $l$ .

It is easy to see that the rows of  $N$  has to be orthonormal vectors, because  $BB' = B_0B_0' = \Sigma$ . The zeros in the bottom-left block of  $N$  are necessary for the last  $l$  columns of  $Z_\epsilon$  to be zero. This is only necessary, not sufficient, because we still need to make sure that  $Z_\epsilon$  is a valid covariance matrix.

The instant response to monetary policy shocks is characterized by the first two ( $k = 2$ ) columns of  $B$ ,  $B_0[N_{11}' \ 0]' = [b_1 \ b_2]N_{11}$ , where  $[b_1 \ b_2]$  is the first two columns of  $B_0$ . With  $2 + l > n$ , we do not try to identify  $B$  fully, because we need extra identifying restrictions to do so.

To partially identify monetary shocks, we only assume that the instant effect of forward guidance shock on the federal funds rate is zero, and positive shocks are contractionary, as in the baseline model. The forward guidance shock is the linear combination of the rate and the forward guidance shocks in the baseline model, and to make the impact on the federal funds rate zero, the weight on the rate shock should be zero. This is true if and only if  $N_{11}$  is lower triangular:

$$N_{11} = \begin{bmatrix} n_{11} & 0 \\ n_{12} & n_{22} \end{bmatrix}. \quad (16)$$

$n_{ij}$  are entries of the 2-by2 matrix  $N_{11}$ . Furthermore, for positive shocks to be contractionary, we assume  $n_{11}, n_{22} > 0$ .

The instant impact of the forward guidance shock is simply  $n_{22}b_2$ . Therefore, IRFs to the forward guidance shock are identical to those in the baseline model in shape. However, their magnitude is smaller, because  $n_{22}$  is (weakly) smaller than 1.

The instant impact of the policy rate shock is given by  $n_{11}b_1 + n_{12}b_2$ . The instant impact of the policy rate shock on the federal funds rate is smaller than in the baseline, as  $n_{11}$  is (weakly) smaller than 1. Also, the IRFs are given as a linear combination of IRFs to the two monetary structural shocks in the baseline model. This means that IRFs can have any of the

forms given in figure 4 in section 4.3. Therefore, once we allow for the possibility of linear dependence between monetary and non-monetary shocks, IRFs to the policy rate shock can have different shapes depending on identifying restrictions. The only exception is the shape of the short-run response by the federal funds rate, which is quite consistent across different potential linear combinations of  $b_1$  and  $b_2$ , as can be seen in figure 4

### 4.3 Relationship to Alternative Instrument-Based Identification Schemes

We first discuss the consequence of assuming that monetary policy shocks are one-dimensional (1D), as in Gertler and Karadi (2015), when the shocks were, hypothetically, two-dimensional (2D).<sup>10</sup> The discussion can be easily extended to the general case of assuming that there are  $k'$ -dimensional structural shocks correlated with instruments, while the true dimension is  $k > k'$ . Under the assumption of 1D monetary policy shock, the single monetary shock would be a linear combination of the two ‘true’ shocks:

$$b_1 = [\tilde{b}_1 \ \tilde{b}_2]w. \quad (17)$$

In the equation,  $b_1$  is the first column of the matrix  $B$  under 1D identification.  $\tilde{b}_1$  and  $\tilde{b}_2$  are the first two columns of the matrix  $B$  with correct 2D identification, and  $w$  is a vector of unit length. Thus, if we do not have enough instruments, we only get to identify certain rotated components of full structural shocks correlated with instruments.

The particular linear combination,  $w$ , is determined by the choice of an instrument for 1D identification. To see this, let  $z_{1D,t}$  be the instrument used for 1D identification, which is a nonzero linear combination of elements in  $z_t$ , the vector of all potential instruments. There exists another linear combination  $z_{2,t}$  of elements in  $z_t$  such that the rank of  $E[[z_{1D,t} \ z_{2,t}]'u_t]$  is 2. Based on our discussion in section 2, there exists a 2-by- $n$  matrix  $\tilde{Z}_\epsilon$  and an  $n$ -by- $n$  matrix  $\tilde{B}$  such that

$$E[[z_{1D,t} \ z_{2,t}]'u_t] = \tilde{Z}_\epsilon \tilde{B}'. \quad (18)$$

In addition, the last  $n - 2$  columns of  $\tilde{Z}_\epsilon$  are zero, and  $\tilde{B}\tilde{B}' = \Sigma$ . The first two columns of  $\tilde{B}$ , denoted  $\tilde{b}_1$  and  $\tilde{b}_2$ , characterize the structural shocks correlated with the instruments.

---

<sup>10</sup>As we have said earlier, the data seem to support both dimensions somewhat equally.

With only  $z_{1D,t}$  as the instrument, we need to find  $B$  satisfying the following equation for a 1-by- $n$  matrix  $Z_0$ , whose elements are all zero except for the first column:

$$E[z_{1D,t}u_t'] = Z_0B', \quad (19)$$

with  $BB' = \tilde{B}\tilde{B}' = \Sigma$ . At the same time, taking the first row of the matrix equation (18), we have

$$E[z_{1D,t}u_t'] = Z_\epsilon\tilde{B}' = Z_\epsilon RR'\tilde{B}', \quad (20)$$

with  $Z_\epsilon$  being the first row of  $\tilde{Z}_\epsilon$ , and for any orthonormal matrix  $R$ . 1D identification is done simply by solving for  $R$  that makes all the columns of  $Z_\epsilon R$  zero except for the first one. Since only the first two columns of  $Z_\epsilon$  can be nonzero, it is always possible to find a solution  $R$  in the following form:

$$R = \begin{bmatrix} w & w_\perp & 0 \\ 0 & 0 & I_{n-2} \end{bmatrix}, \quad (21)$$

where  $[w \ w_\perp]$  is a 2-by-2 orthonormal matrix. The first column of  $B = \tilde{B}R$  represents the 1D monetary policy shock, which is  $[\tilde{b}_1 \ \tilde{b}_2]w$ .

With our VAR model, we try 1D identification with an instrument defined as the projection of the estimated reduced-form residual of the federal funds rate onto the two baseline instruments (current-month fed funds futures and three-quarter-ahead eurodollar futures). The identified monetary policy shock is  $0.99 \times$  (policy rate shock)  $+ (-0.15) \times$  (forward guidance shock). The resulting impulse response functions, shown by figure 3, is remarkably similar to those in Gertler and Karadi (2015), despite differences in the choice of VAR variables used to identify monetary policy shocks (fed funds rate in our paper vs. 1-year Treasury rate in Gertler and Karadi (2015)) and the choice of instruments (CM FF and 3Q ED in our paper vs. 3M FF in Gertler and Karadi (2015)).<sup>11</sup>

Impulse responses to 1D policy shock under the 2SLS instrument are close to those to the policy rate shock under 2D identification, as figure 4 shows. Under 1D identification, the impulse responses are linear combinations of those to the policy rate shock and the forward guidance shock, and the composition is determined by the choice of the instrument for 1D identification.

---

<sup>11</sup>In addition, the baseline VAR model in Gertler and Karadi (2015) has additional VAR variables.

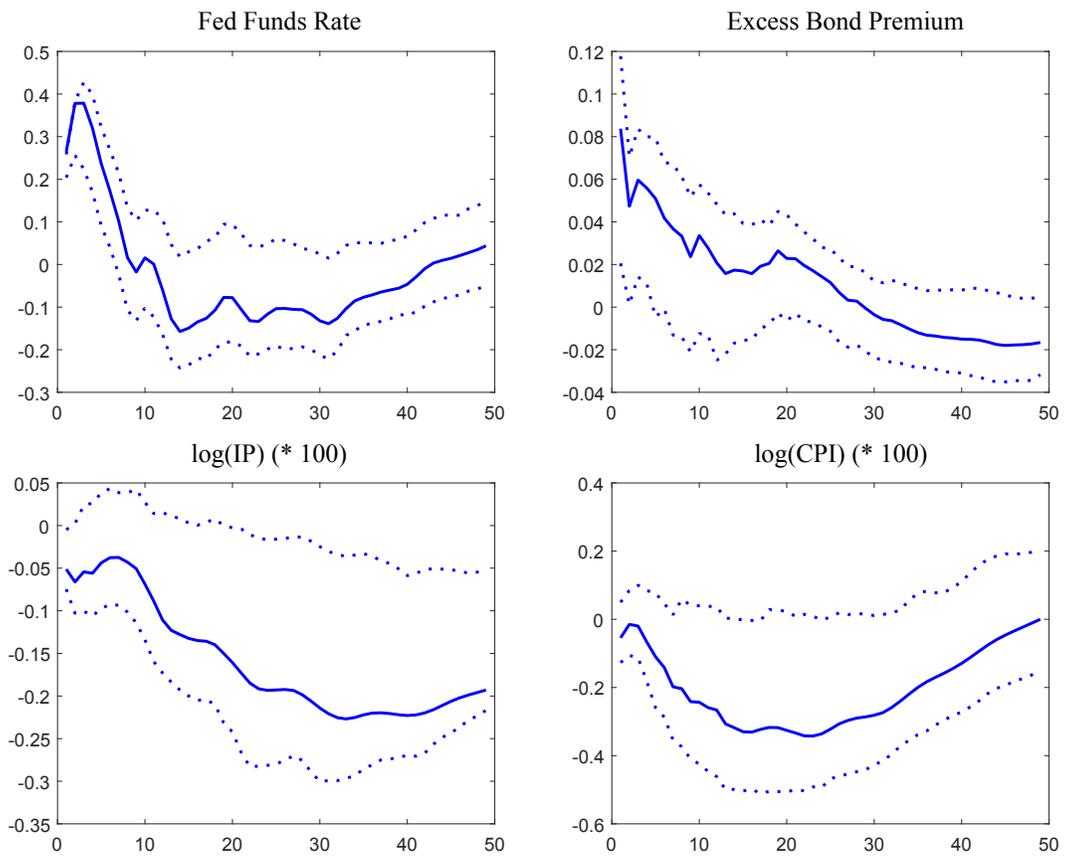


Figure 3: IRF to 1D Monetary Shocks

To see how important the choice of instrument is, we show how much the shape of the IRFs can vary with 1D identification under different linear combinations of the two instruments, CM FF and 3Q ED. The exercise is done by first creating a grid on all possible values of vector  $w$ , which is easy because  $|w| = 1$ . Then, for each linear combination  $w$ , we determine which combination of the two instruments will produce 1D structural shock in the form of  $[\tilde{b}_1 \ \tilde{b}_2]w$ . Then, we check if the instrument is not too weak by making sure that  $F > 10$ , for the regression of the federal funds rate residual on the instrument. Finally, we pick two instruments that are not weak and give most extreme combinations, the largest positive (instrument N) and negative (instrument P) weights on the forward guidance shock.

A simpler interpretation of the exercise is that instrument P is constructed as (CM FF)  $+x \times$  (3Q ED), where  $x$  is roughly the positive number with the largest absolute value such that instrument P is not a weak instrument under the rule of thumb  $F > 10$  (Stock et al. (2002) and Stock and Yogo (2005)). This interpretation is possible because CM FF is a much stronger instrument for the federal funds rate than 3Q ED, and thus, larger  $|x|$  tends to make the instrument weaker. Similarly, instrument N is constructed for negative  $x$ .

More explicitly, the instruments are constructed as follows:

Instrument P: (CM FF)  $+5.09 \times$  (3Q ED),  $F = 11.6$ .

Instrument N: (CM FF)  $-1.18 \times$  (3Q ED),  $F = 11.1$ .

The impulse responses for different instruments show differing shapes, because each instrument choice leads to impulse responses that are distinct composites of those to policy rate and forward guidance shocks. In particular, the monetary policy shock identified with instrument P is  $0.79 \times$  (policy rate shock)  $+(-0.61) \times$  (forward guidance shock), while that identified with instrument N is  $0.90 \times$  (policy rate shock)  $+0.44 \times$  (forward guidance shock). The only highly consistent response across different instruments is the magnitude and duration of the fed funds rate's short-run response. Any reasonable instrument for the federal funds rate gives the policy rate shock a weight close to 1, leading to similar short-run responses given that the forward guidance shock generates only a weak initial response in the federal funds rate.

Another paper that identifies 2D monetary policy shocks is Lakdawala (2017). We note that the impulse response functions in the paper look somewhat different from ours. Moreover, this difference does not seem to be coming just from the difference in identification schemes, as we cannot find rotations of our shocks that look alike the shocks in the paper;

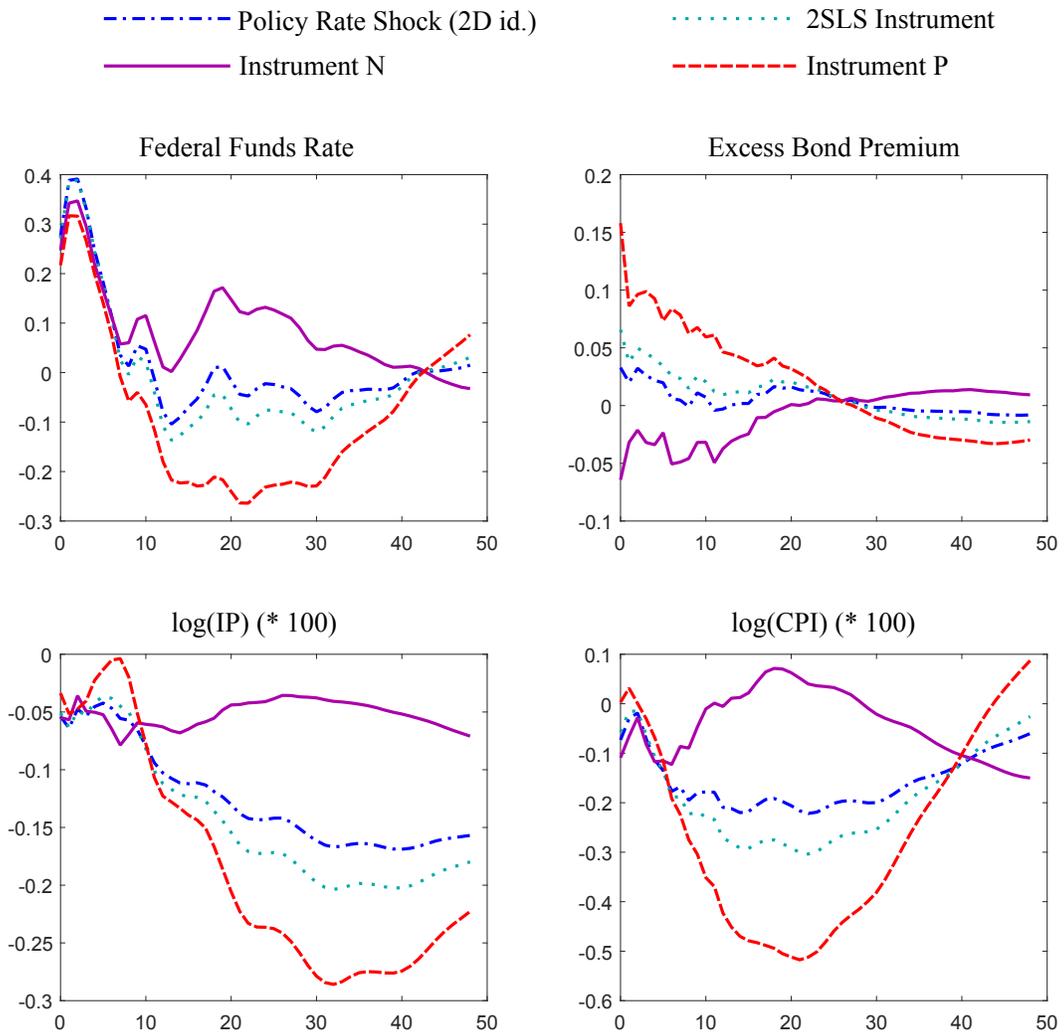


Figure 4: 1D IRF with Different Instruments

note that from the characterization of the matrix  $B$  in section 2, we expect to find such a rotation if our model and the other paper's model were essentially different representations of the same model.

## 5 Conclusion

Monetary policy announcements occur at fixed times, and instrument VAR can be used to measure the impact of monetary policy announcements on macroeconomic variables, using instruments that capture the response of different short-term rates to those announcements. We identify forward guidance shocks as the component of the structural shocks that are correlated with such instruments and have zero instant impact on the federal funds rate. This is different from identifying restrictions used in previous studies that identified monetary policy shocks using instrument VAR, such as Gertler and Karadi (2015) and Lakdawala (2017).

Given the shapes of IRFs, contractionary forward guidance seems to accompany expectations of future increase in output and price level. This suggests that contractionary forward guidance identified through instrument VAR mainly represents a positive view on the future increase of output and price level, rather than unexpected changes in future monetary policy stance. Expectation of future interest rate seemingly adjusts to reflect new information on the future path of the economy after monetary policy announcements, not to take into account unexpected changes in monetary policy stance.

We extend the baseline model in a few novel ways: First, we decompose non-monetary structural shocks, and find that the response of the federal funds rate to shocks to output and price level are what we expect under standard macroeconomic theory. Second, we consider the possibility of linear dependence between monetary and non-monetary structural shocks. This is a reasonable concern given that forward guidance seems to reflect mainly the release of information about the future of the economy, and a similar form of information can be revealed outside monetary policy announcements. We show that our conclusions about the shape of the IRFs (but not magnitude) are still largely valid under this violation of usual instrument VAR assumptions. Finally, we discuss the relationship between our model and those in Gertler and Karadi (2015) and Lakdawala (2017). In doing so, we describe exactly what happens to identification if we use fewer VAR instruments than the number of structural

shocks.

We hope that our presentation of the instrument VAR methodology, combined with these extensions, will be a useful reference for many economists interested in applying it.

## References

- Campbell, Jeffrey R., Charles L. Evans, Jonas D. M. Fisher, and Alejandro Justiniano (2012) “Macroeconomic Effects of Federal Reserve Forward Guidance,” *Brookings Papers on Economic Activity*, pp. 1–54.
- Campbell, Jeffrey R., Jonas D. M. Fisher, Alejandro Justiniano, and Leonardo Melosi (2016) “Forward Guidance and Macroeconomic Outcomes Since the Financial Crisis,” Vol. 31 of NBER Macroeconomics Annual.
- Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans (1999) “Monetary Policy Shocks: What Have We Learned and to What End?”, Vol. 1 of Handbook of Macroeconomics, Amsterdam, The Netherlands: Elsevier B. V. pp. 65–148.
- Gertler, Mark and Peter Karadi (2015) “Monetary Policy Surprises, Credit Costs, and Economic Activity,” *American Economic Journal: Macroeconomics*, Vol. 7, pp. 44–76.
- Gilchrist, Simon and Egon Zakrajšek (2012) “Credit Spreads and Business Cycle Fluctuations,” *American Economic Review*, Vol. 102, pp. 1692–1720.
- Gürkaynak, Refet S., Brian Sack, and Eric 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*, pp. 55–93.
- Jentsch, Carsten and Kurt G. Lunsford (2016) “Proxy SVARs: Asymptotic Theory, Bootstrap Inference, and the Effects of Income Tax Changes in the United States,” Federal Reserve Bank of Cleveland Working Paper 16-19.
- Kleibergen, Frank and Richard Paap (2006) “Generalized Reduced Rank Tests Using the Singular Value Decomposition,” *Journal of Econometrics*, Vol. 133, pp. 97–126.
- Lakdawala, Aeimit (2017) “Decomposing the Effects of Monetary Policy Using an External Instruments SVAR,” Working Paper.
- Mertens, Karel and Morten O. Ravn (2013) “The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States,” *American Economic Review*, Vol. 103, pp. 1212–1247.

- Nakamura, Emi and Jón Steinsson (2013) “High Frequency Identification of Monetary Non-Neutrality: The Information Effect,” Working Paper.
- Sims, Christopher A. (1980) “Macroeconomics and Reality,” *Econometrica*, Vol. 48, pp. 1–48.
- Sims, Christopher A. and Tao Zha (1999) “Error Bands for Impulse Responses,” *Econometrica*, Vol. 67, pp. 1113–1155.
- Stock, James H. and Mark W. Watson (2012) “Disentangling the Channels of the 2007-09 Recession,” *Brookings Papers on Economic Activity*, pp. 81–135.
- Stock, James H., Jonathan H. Wright, and Motohiro Yogo (2002) “A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments,” *Journal of Business & Economic Statistics*, Vol. 20, pp. 518–529.
- Stock, James H. and Motohiro Yogo (2005) “Testing for Weak Instruments in Linear IV Regression,” New York, NY: Cambridge University Press, pp. 80–108.
- Watson, Mark W. (1994) “Vector Autoregressions and Cointegration,” Vol. 4 of Handbook of Econometrics, Amsterdam, The Netherlands: Elsevier B. V. pp. 2843–2915.
- Wu, C. F. J. (1986) “Jackknife, Bootstrap and Other Resampling Methods in Regression Analysis,” *Annals of Statistics*, Vol. 14, pp. 1261–1295.

## 6 Appendix

### 6.1 Supplementary Discussion for Section 2

We provide some technical details for section 2.

First, we show that given  $\text{rank}(Z_u) = k$ , it is possible to define structural shocks so that  $\text{rank}(E[z_t \epsilon'_{1,t}]) = k$  and  $E[z_t \epsilon'_{2,t}] = 0$ . Let  $B_1$  be an  $n$ -by- $n$  matrix such that  $B_1 B_1' = \Sigma$ . For  $u_t = B_1 \epsilon_t$ ,  $E[z_t \epsilon'_t] = Z_u(B_1^{-1})'$ . Since any  $B = B_1 R$  with an orthonormal  $n$ -by- $n$  matrix  $R$  represents an equivalent VAR model with  $u_t = B \epsilon_t$ , we only need to show that there exists  $R$  such that  $E[z_t \epsilon'_t] = Z_u(B^{-1})' = Z_u(B_1^{-1})' R$  satisfies the required properties. It is enough to show that there exists  $R$  such that the last  $n - k$  columns of  $E[z_t \epsilon'_t]$  is zero, because the nonsingularity of  $B_1$  and  $R$  guarantees that once the last  $n - k$  columns are zero, the first  $k$  columns are independent. It follows from standard results in linear algebra that we can find  $R$  that makes the last  $n - k$  columns zero. Also, there is no  $R$  such that only the first  $l$  columns of  $Z_u(B_1^{-1})' R$  are nonzero, for  $l < k$ . There exists  $R$  such that the statement holds for  $l > k$ , but the first  $l$  columns would not be independent in that case.

Second, we show that  $B_1 = B_2 R(k, n - k)$  completely characterizes observationally equivalent models, those with identical  $\Sigma$  and  $Z_u$ . Given any  $B_1$  and  $B_2$  such that  $B_1 B_1' = B_2 B_2' = \Sigma$ ,  $B_1 = B_2 R$  for an  $n$ -by- $n$  orthonormal matrix  $R$ . Then,  $Z_u(B_1^{-1})' = Z_u(B_2^{-1})' R$ . In addition, if the last  $n - k$  columns of both  $Z_u(B_1^{-1})'$  and  $Z_u(B_2^{-1})'$  are zero, their first  $k$  columns must be nonzero and independent. Therefore, the top-right  $k$ -by- $(n - k)$  submatrix of  $R$  must be zero. Similarly,  $Z_u(B_2^{-1})' = Z_u(B_1^{-1})' R'$  holds, and this implies that the bottom-left  $(n - k)$ -by- $k$  submatrix of  $R$  is zero. Therefore,  $R$  has to take the form  $R(k, n - k)$  as a necessary condition for equivalence between  $B_1$  and  $B_2$ .

The converse that  $B_1 = B_2 R(k, n - k)$  is consistent with the restrictions from the data, given that  $B_2$  is so, is even more obvious.  $R(k, n - k)$  is an  $n$ -by- $n$  orthonormal matrix by construction, and the last  $n - k$  columns of  $Z_u(B_1^{-1})' = Z_u(B_2^{-1})' R(k, n - k)$  are zero because they are linear combinations of the last  $n - k$  columns of  $Z_u(B_2^{-1})'$ , which are zero.

Given the block representation of  $B$  with  $B_{ij}$ ,  $i, j = 1, 2$  as in equation (6), the four matrix products  $B_{11} B_{11}'$ ,  $B_{22} B_{22}'$ ,  $B_{21} B_{11}^{-1}$  and  $B_{12} B_{22}^{-1}$  are invariant to multiplying  $B$  by  $R(k, n - k)$  from the right. The reason is that this multiplication makes  $B_{i1}$  into  $B_{i1} R_k$  and  $B_{i2}$  into  $B_{i2} R_{n-k}$ .

Also, these four matrix products uniquely determine  $B$  up to multiplication by  $R(k, n - k)$

from the right.  $B_{11}$  can be determined from  $B_{11}B'_{11}$  uniquely up to multiplication by  $R_k$  from the right, and  $B_{21}$  can be determined from  $B_{21}B_{11}^{-1}$  and  $B_{11}$ .  $B_{22}$  and  $B_{12}$  can be similarly determined uniquely up to multiplication by  $R_{n-k}$  from the right.

## 6.2 Computing $B$

In section 2, we have seen that to determine  $B$ , it is enough to determine  $B_{11}B'_{11}, B_{22}B'_{22}, B_{21}B_{11}^{-1}$  and  $B_{12}B_{22}^{-1}$ , where  $B_{ij}$  is a submatrix of  $B$  as in equation (6). Here, we reproduce formulas in the appendix of Mertens and Ravn (2013) computing these matrix products in terms of  $\Sigma$  and  $Z_u$ .

First, if  $m > k$ , choose any  $k$  independent rows (or  $k$  independent linear combinations of rows) of  $Z_u$  and eliminate the rest, so that  $Z_u$  has the full row rank. Then, let  $Z_{u,1}$  denote the first  $k$  columns of  $Z_u$  and  $Z_{u,2}$  denote its last  $n - k$  columns. Without any loss of generality, we assume that  $Z_{u,1}$  is nonsingular, because if it is not, we can change the order of VAR variables to make  $Z_{u,1}$  nonsingular. Then,

$$B_{21}B_{11}^{-1} = Z'_{u,2}(Z'_{u,1})^{-1}. \quad (22)$$

From  $BB' = \Sigma$ , we obtain the following three equations:

$$B_{11}B'_{11} + bB_{22}B'_{22}b' = \Sigma_{11}. \quad (23)$$

$$B_{11}B'_{11}a' + bB_{22}B'_{22} = \Sigma_{12}. \quad (24)$$

$$aB_{11}B'_{11}a' + B_{22}B'_{22} = \Sigma_{22}. \quad (25)$$

The subscripts on  $\Sigma$  denote its submatrices in the same way as they denote the submatrices of  $B$ , and  $a \equiv B_{21}B_{11}^{-1}$  and  $b \equiv B_{12}B_{22}^{-1}$ . By solving these equations, we get:

$$bB_{22}B'_{22}b' = (\Sigma_{12} - \Sigma_{11}a')(\Sigma_{22} + a\Sigma_{11}a' - a\Sigma_{12} - \Sigma_{21}a')^{-1}(\Sigma_{21} - a\Sigma_{11}). \quad (26)$$

$$B_{11}B'_{11} = \Sigma_{11} - bB_{22}B'_{22}b'. \quad (27)$$

$$B_{22}B'_{22} = \Sigma_{22} - aB_{11}B'_{11}a'. \quad (28)$$

$$b = (\Sigma_{12} - B_{11}B'_{11}a')(B_{22}B'_{22})^{-1}. \quad (29)$$

### 6.3 Normalization of rk-Statistic

Normalization is done by properly defining  $F$  and  $G$  in equation (16) in Kleibergen and Paap (2006):

$$\hat{\Theta} = G\hat{\Pi}F'. \quad (30)$$

We use the following  $G$  and  $F$ , where  $G$  is  $m$ -by- $m$  and  $F$  is  $n$ -by- $n$ :

$$G'G \equiv \frac{1}{T} \sum_{t=1}^T z_t z_t', \quad (31)$$

$$F'F \equiv \frac{1}{T} \sum_{t=1}^T u_t u_t'. \quad (32)$$

These formulas do not uniquely determine  $G$  and  $F$  because for any orthonormal matrices  $R_m$  and  $R_n$ ,  $R_m G$  and  $R_n F$  also satisfy these two equations. However, any  $G$  and  $F$  satisfying these formulas can be used, as all possible choices lead to the same value for rk-statistic. In cases where we use only subvectors of  $z_t$  and  $u_t$ , we use corresponding subvectors in the definition  $G$  and  $F$  as well. Following the construction of rk-statistic outlined in Kleibergen and Paap (2006), this choice of  $G$  and  $F$  can be shown to lead to the same value of rk-statistic under any invertible linear transformations of  $z_t$  and  $u_t$ . In other words, for any nonsingular  $m$ -by- $m$  matrix  $A_m$  and nonsingular  $n$ -by- $n$  matrix  $A_n$ , replacing  $z_t$  and  $u_t$  by  $A_m z_t$  and  $A_n u_t$  does not change the value of rk-statistic. We do not reproduce the computational steps because they follow from standard results in linear algebra.

### 6.4 Characterization of a Model with Linearly Dependent Monetary and Non-Monetary Shocks

As stated in section 4.2, the parameter  $B$  of the VAR model is characterized by several conditions. As discussed in section 2,  $B$  has to satisfy the following condition:

$$BB' = E[u_t u_t'] = \Sigma. \quad (33)$$

In addition, there exists a matrix  $Z_\epsilon \equiv E[z_t \epsilon_t']$  that satisfies three conditions. First, with given  $E[z_t z_t']$  and  $E[\epsilon_t \epsilon_t'] = I_{k+l}$ ,  $Z_\epsilon$  forms a valid covariance matrix. Second, the last  $l$

columns of  $Z_\epsilon$  are zero. Finally, it satisfies the following equation:

$$Z_\epsilon B' = E[z_t u_t'] \equiv Z_u. \quad (34)$$

There exists an  $n$ -by- $n$  nonsingular matrix  $B_0$  that satisfies all the conditions on  $B$ , with  $l = n - k$  (See the discussion in section 2). Then,  $B$  can be written as

$$B = B_0 N, \quad (35)$$

for an  $n$ -by- $(k + l)$  matrix  $N$ .

The equation  $BB' = \Sigma$  is equivalent to  $NN' = (B_0)^{-1}\Sigma(B_0')^{-1} = I_n$ , because  $\Sigma = B_0 B_0'$ . Therefore, the  $n$  rows of  $N$  are orthonormal vectors in  $\mathbb{R}^{k+l}$ .

Let  $Z_0 \equiv Z_u(B_0')^{-1}$ . By definition, the last  $n - k$  columns of  $Z_0$  are zero. Also, equation (34) can be written as follows:

$$Z_\epsilon B' = Z_\epsilon N' B_0' = Z_0 B_0'. \quad (36)$$

Therefore, the third condition on  $Z_\epsilon$  can be stated in terms of  $N$ :  $Z_\epsilon N' = Z_0$ . Along with the second condition on  $Z_\epsilon$  that the last  $l$  columns of  $Z_\epsilon$  be zero, this implies that  $N$  has the following form:

$$N = \begin{bmatrix} N_{11} & N_{12} \\ 0 & N_{22} \end{bmatrix}. \quad (37)$$

$N_{11}$  is  $k$ -by- $k$ ,  $N_{12}$  is  $k$ -by- $l$ , and  $N_{22}$  is  $(n-k)$ -by- $l$ . The first  $k$  columns of  $Z_\epsilon$  are independent because  $\text{rank}(Z_\epsilon N') = \text{rank}(Z_0) = k$ . Since the last  $n - k$  columns of  $Z_0 = Z_\epsilon N'$  are zero, the bottom-left  $(n - k)$ -by- $k$  submatrix of  $N$  must be zero. Since the first  $k$  columns of  $Z_0 = Z_\epsilon N'$  are the first  $k$  columns of  $Z_\epsilon$  multiplied by  $N_{11}$  from the right, and the first  $k$  columns of  $Z_0$  are independent,  $N_{11}$  is nonsingular.

Conversely, for any  $N$  in the form given by equation (37),  $Z_\epsilon$  that satisfies the second and third conditions can be found. This can be simply done by defining the first  $k$  columns of  $Z_\epsilon$  as the first  $k$  columns of  $Z_0$  multiplied by  $(N_{11}')^{-1}$  from the right, and by defining the remaining  $l$  columns of  $Z_\epsilon$  as zero.

Therefore, as long as  $Z_\epsilon$  forms parts of a valid covariance matrix of  $[z_t' \ \epsilon_t']'$ , the following conditions on  $N$  characterizes  $B = B_0 N$ :  $NN' = I_n$  and  $N_{11}$  is nonsingular.  $NN' = I_n$

implies that the rows of  $N_{22}$  are orthonormal and each row of  $N_{12}$  is in the null space of  $N_{22}$ .