Universität Trier

Quantitative Finance and Risk Analysis

Working Paper Series Nº 23 - 01:

(Almost) Recursive Shock Identification with Economic Parameter Restrictions

Jan Pablo Burgard

Matthias Neuenkirch

Dennis Umlandt

(Almost) Recursive Shock Identification with Economic Parameter Restrictions*

Jan Pablo Burgard — Trier University and RIFOSS

Matthias Neuenkirch† — Trier University and CESifo

Dennis Umlandt — University of Innsbruck

First Draft: January 9, 2023
This Version: October 4, 2023

Abstract

We propose to estimate the parameters of a vector autoregressive model under the restriction that economic theory is not violated, while the shocks are still recursively identified. We use an augmented Lagrange solution approach, which adjusts the coefficients to meet the theoretical requirements. In a generalization, we additionally allow for a (minimal) rotation of the Cholesky matrix. Based on a Monte Carlo study and an empirical application, we show that the "almost recursive identification with parameter restrictions" leads to a solution that avoids an estimation bias, generates theory-consistent impulse responses, and is as close as possible to the recursive scheme.

JEL Codes: C32, C61, C82.

Keywords: Non-Linear Optimization, Recursive Identification, Rotation, Sign Restrictions.

*We offer our gratitude to Helmut Herwartz, Tobias Kranz, Pascal Langer, as well as participants of the 2022 Annual Conference of the German Economic Association in Basel for their helpful comments on earlier versions of the paper. This paper was previously circulated under the title "(Almost) Recursive Identification of Monetary Policy Shocks with Economic Parameter Restrictions."

†Corresponding author: Trier University, Department of Economics, D-54286 Trier, Germany, Tel.: +49-651-201-2629, E-mail: neuenkirch@uni-trier.de.

1 Introduction

In this paper, we propose a novel strategy to identify shocks in a vector autoregressive (VAR) model. We extend what is arguably the simplest estimation and identification approach, that is, least squares parameter estimation and recursive identification of shocks. Based on the work by Gafarov et al. (2018), we propose to estimate the VAR parameters under the restriction that economic theory is not violated, but stick to a recursive identification of shocks. This leads to a mathematical optimization problem under non-linear constraints, for which we propose an augmented Lagrange solution approach. As an illustration, we estimate a standard monetary policy transmission VAR model under the additional constraint that the impulse response functions (IRFs) for output and prices (the interest rate) after a contractionary monetary policy shock should not be positive (negative) for a period of h months after the shock. Hence, our approach resembles the idea of sign restrictions (Uhlig, 2005), but imposes restrictions right away during the estimation process and not on the IRFs in a "second stage."

However, since the identification scheme is fixed, a situation could arise where the parameters of the data-generating process (DGP) might not fulfill the economic restrictions. As a consequence, the estimated coefficients may no longer resemble the DGP and, therefore, are biased. To address this issue, we employ a penalized regression approach that allows for a rotation of the Cholesky matrix in addition to the parameter restrictions. The resulting optimization procedure favors identification schemes with as little as possible departures from the initial recursive identification scheme to fulfill the economic restrictions. At the same time, this procedure should prevent biased VAR estimates.

To investigate the abilities and limitations of our approach, we first conduct a Monte Carlo simulation where we show that — if the economic parameter constraints are valid for the DGP — the restricted (but unrotated) VAR produces consistent estimates while having a similar small sample performance as the standard VAR. However, if the DGP does not fulfill the economic restrictions, we can only enforce the constraints by allowing for a (substantial) estimation bias. In this case, the "almost recursively identified approach" that slightly rotates the initial propagation of shocks can prevent the bias while generating theory-consistent impulse responses.

Second, we estimate a standard monetary policy transmission VAR for the euro area and the period January 1999–December 2019. Our results indicate that a restriction on the first two months in the IRFs of output, prices, and the interest rate is sufficient to get rid of the price puzzle and the counterintuitive response of output found in a standard Cholesky-identified VAR. This holds for both, the recursively identified VAR with parameter restrictions and the model where we additionally allow for a (small) rotation of the Cholesky matrix. However, the almost recursively identified approach outperforms the Cholesky identification as it features virtually no change in the model parameters and no noteworthy decrease in the model's fit.

The additional flexibility of this approach is underscored by the possibility of having non-zero responses on impact for output and inflation, which is by definition ruled out in the recursive identification scheme. Consequently, in particular the "almost recursively identified approach with parameter restrictions" is a useful complement to the Bayesian sign restriction approach commonly used in empirical macroeconomic studies as it leads to a solution without an estimation bias, generates theory-consistent impulse responses, and is as close as possible to the recursive scheme.

The remainder of this paper is organized as follows. Section 2 briefly introduces VAR models and identification via Cholesky decomposition. Thereafter, a recursive identification scheme with restrictions in the estimation process is proposed, before we additionally allow for a rotation of the Cholesky matrix in the non-linear optimization procedure. Section 3 compares the small sample properties of both restricted VAR specifications in a Monte Carlo study to those of a standard VAR. Section 4 provides an empirical application for the euro area. Section 5 concludes.

2 VAR Models with Different Identification Approaches

2.1 Recursive Identification

In a first step, we introduce the standard estimation and recursive identification approach for VAR models.[1] The VAR model of order p for a K-dimensional time series vector x_t can be written in its reduced form as:

$$y_t = \delta + A_1 y_{t-1} + A_2 y_{t-2} + \ldots + A_p y_{t-p} + u_t, \qquad t = 1, \ldots T \tag{1}$$

where the A_i $(i = 1, \ldots, p)$ are $K \times K$-dimensional coefficient matrices, u_t is a independently and identically distributed K-dimensional vector of random errors with mean $\mathbb{E}(u_t) = 0$ and covariance matrix $\mathbb{E}(u_t u_t^\top) = \Sigma_U$, and δ is the K-dimensional vector of intercepts. The model can be stacked in matrix notation as follows:

$$\mathbf{Y} = \mathbf{Z}\boldsymbol{\beta} + \mathbf{U}, \tag{2}$$

with

$$\mathbf{Y} := \left[x_{p+1}, x_{p+2}, \ldots, x_T\right]^\top \in \mathbb{R}^{(T-p)\times K}, \qquad \mathbf{Z_t} := \left[1, x_t^\top, x_{t-1}^\top, \ldots, x_{t-p+1}^\top\right]^\top \in \mathbb{R}^{(Kp+1)},$$
$$\mathbf{Z} := \left[Z_p, Z_{p+1}, \ldots, Z_{T-1}\right]^\top \in \mathbb{R}^{(T-p)\times(Kp+1)}, \qquad \mathbf{U} := \left[u_{p+1}, u_{p+2}, \ldots, u_T\right]^\top \in \mathbb{R}^{(T-p)\times K},$$

and

$$\boldsymbol{\beta} := \left[\delta, A_1, A_2, \ldots, A_p\right]^\top \in \mathbb{R}^{(Kp+1)\times K}.$$

Eq. (2) can be estimated by equation-wise ordinary least squares to obtain $\hat{\boldsymbol{\beta}}$.

To obtain the IRFs for time horizon ζ, we write the fitted VAR(p) model in its moving average (MA) representation:

$$y_\zeta = \Phi_0 u_\zeta + \Phi_1 u_{\zeta-1} + \Phi_2 u_{\zeta-2} + \ldots, \tag{3}$$

[1]For thorough discussion of VAR(p) models we recommend the textbook by Lütkepohl (2005).

where $\Phi_0 := I_K$ and Φ_t can be computed recursively using the coefficient matrices $A_1, \ldots, A_p$ of the VAR(p) model according to

$$\Phi_t := \sum_{i=1}^{t} \Phi_{t-i} A_i, \tag{4}$$

with $A_i := 0$ for $i > p$. Accordingly, Φ_t is a non-linear function of the model parameters for $t > 1$.

Our interest is in studying the response of a variable to an impulse of another variable. For instance, let the first variable in Y be the one giving the impulse. Then we create a vector $u_0 = (1, 0, \ldots, 0)^\top$ being the impulse at time point 0. The response on the time points $1, \ldots, \zeta$ is given by

$$y_0 = u_0, \quad y_1 = \Phi_1 u_0, \quad y_2 = \Phi_2 u_0, \ \ldots$$

leading to the cumulative impact at time t of

$$\widetilde{\psi}_t = \sum_{i=0}^{t} \Phi_i u_0.$$

The reduced form innovations in u_t tend to be correlated with each other and do not feature a structural interpretation. For the purpose of identifying the uncorrelated structural shocks, the residual covariance matrix is factorized with a matrix P such that $\Sigma_u = PR^\top RP^\top$. R is a (orthogonal) rotation matrix of dimension K and is assumed to be the identity matrix for this and the next subsection. In subsection 2.3, a general rotation matrix will be allowed.

The identification problem therefore boils down to choosing one of the infinitely many orthogonal matrices $\widetilde{P} := PR^\top$ that decompose the covariance matrix. One popular approach is to choose the matrix P with a Cholesky factorization. Given a determined P, the structural shocks are defined as $\widetilde{u_t} = \widetilde{P}^{-1} u_t$ and are orthogonal (i.e., $\mathbb{E}\left(\widetilde{U_t}\right) = I_K$). Given P, the MA representation (3) can be written in terms of structural innovations as

$$y_\zeta = \Psi_0 \widetilde{u}_\zeta + \Psi_1 \widetilde{u}_{\zeta-1} + \Psi_2 \widetilde{u}_{\zeta-2} + \ldots, \tag{5}$$

with $\Psi_i := \Phi_i \widetilde{P}$ leading to the cumulative orthogonal impact of

$$\widetilde{\Psi}_t = \sum_{i=0}^{t} \Psi_i \widetilde{U}_0.$$

2.2 Recursive Identification with Parameter Restrictions

Recursive identification of shocks might lead to theoretically counterintuitive IRFs (see also Figure 1 in Section 4.2.1 below). Accordingly, one might need to assume some predefined short-run behavior of the IRFs based on economic theory. As a first modification to the standard recursively identified VAR, we propose to impose this theory-conform behavior of the IRFs as restrictions in the estimation process.

We start with the matrix notation from Eq. (2). As an illustration, we assume that the response of the second variable given an impulse of the first variable should be non-positive at time points $t+1$, $t+2$, and $t+3$. Hence, the estimated model should reflect this theoretical restriction. Put differently, we use information from economic theory to impose a prior on the reduced-form estimation. The resulting optimization problem is finding the parameter vector $\text{vec}(\boldsymbol{\beta})$ that minimizes the sum of squared residuals (SSR) of the VAR(p) model under the restriction that the IRFs generated by the set of model parameters satisfies the sign restrictions:[2]

$$\begin{aligned} \underset{\text{vec}(\boldsymbol{\beta})}{\text{minimize}} \quad & \text{vec}(\mathbf{U})^\top \text{vec}(\mathbf{U}) \\ \text{subject to} \quad & \mathbf{U} = \mathbf{Y} - \mathbf{Z}\boldsymbol{\beta}, \\ & (\widetilde{\Psi}_{t+1})_{1\,2} \leq 0, \\ & (\widetilde{\Psi}_{t+2})_{1\,2} \leq 0, \\ & (\widetilde{\Psi}_{t+3})_{1\,2} \leq 0. \end{aligned} \tag{6}$$

[2]Note that we consider the impact of an impulse not to be seen at time 0, that is, no instantaneous reaction on the full system is assumed. Mathematically, however, this would not impose major changes.

More generally, one can write for a response size matrix $\mathcal{R}_i \in \mathbb{R}^{K\times K}$ and a sign restriction matrix $\mathcal{S}_i \in \mathbb{R}^{K\times K}$ for each forecast point i

$$(\mathcal{S}_i)_{jk} := \begin{cases} -1, & \text{if response } j \text{ from impulse } k \text{ is to be greater than } (\mathcal{R}_i)_{jk} \\ 1, & \text{if response } j \text{ from impulse } k \text{ is to be lower than } (\mathcal{R}_i)_{jk} \\ 0, & \text{if response } j \text{ from impulse } k \text{ is not to be restricted} \end{cases} \tag{7}$$

the following optimization problem with $\cdot$ being the element-wise multiplication operator and ζ the final prediction period:

$$\begin{aligned} \underset{\operatorname{vec}(\boldsymbol{\beta})}{\text{minimize}} \quad & \operatorname{vec}(\mathbf{U})^\top \operatorname{vec}(\mathbf{U}) \\ \text{subject to} \quad & \mathbf{U} = \mathbf{Y} - \mathbf{Z}\boldsymbol{\beta}, \\ & \mathcal{S}_{t+i} \cdot \widetilde{\Psi}_{t+i} - \mathcal{R}_{t+i} \leq 0, \quad \forall i = 1,\ldots,\zeta. \end{aligned} \tag{8}$$

While the loss function is convex, the restrictions are non-linear. Hence, the overall optimization problem is not convex anymore, and a non-linear optimization procedure has to be applied. We choose to use an augmented Lagrange approach implemented in the function `auglag` of the R package `alabama` (Varadhan 2015).

Algorithm 1: Residual Bootstrap for the Restricted VAR(p) Model

Estimate the model parameters using the restricted approach and save the residuals in the matrix $\mathbf{U} = \mathbf{Y} - \mathbf{Z}\boldsymbol{\beta}$.;

for *r = 1,…,R bootstrap resamples* **do**

1.) Sample with replacement U^{r*} from U.;

2.) Generate the bootstrapped time series Y^{r*} and Z^{r*} recursively, starting with the first observation and the sampled residuals.;

3.) Estimate the restricted VAR(p) model and save the parameter estimates and the IRFs.

end

Compute bootstrapped confidence intervals for the IRFs and standard errors for the model parameters.

The model uncertainty is taken into account using a residual bootstrap in analogy to Lütkepohl (2000). The bootstrapping procedure is described in Algorithm 1.

2.3 Almost Recursive Identification with Parameter Restrictions

In contrast to a standard VAR(p) model, we restrict the parameter space to force the resulting IRFs to satisfy a priori economically meaningful sign constraints. Since the structural shocks are still recursively identified via the unique Cholesky factorization of the innovations' covariance matrix, the economic restrictions can only be satisfied by (substantially) adjusting the VAR coefficients. As long as the recursively identified shocks are met by the true parameters of the DGP, the procedure leads to consistent estimates. However, if this is not the case, the estimates of the restricted optimization approach will be considerably biased and inconsistent. In order to tackle this issue, we propose an extended constrained estimation procedure that relaxes the assumption of recursively identified shocks as a second modification to the standard approach.

Let P be the Cholesky factor of the innovations' covariance matrix and $R(\theta)$ an orthogonal rotation matrix, which is parameterized by the vector θ from an adequate parameter space Θ such that $R(0) = I_K$. For example, the rotations can be parameterized with $\theta \in [-\pi, \pi]$ and

$$R(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \tag{9}$$

for $K = 2$ dimensions.[3] The rotated Cholesky factor $\tilde{P}(\theta) = PR(\theta)$ implies a valid identification of shocks because of $\tilde{P}(\theta)\tilde{P}(\theta)' = PR(\theta)R(\theta)'P' = PP' = \Sigma_U$. The additional parameter vector θ allows to satisfy economic restrictions on the structural IRFs by additionally modifying the identification scheme and therefore does not strictly restrict the space of the VAR parameter vector β.

Since the particular identification of structural shocks does not affect the likelihood objective function, we introduce a L_p-regularization term $\|\theta\|_p$ for the angle θ. As the

[3]Similar tractable parameterizations exist for higher dimensions and can be used in the discussed approach. For example, we use the Tait-Bryan angles with $\theta \in [-\pi, \pi]^3$ for the case of $K = 3$ dimensions in the following empirical application.

rotation matrix coincides with the identity matrix when $\theta = 0$, the procedure favours identification schemes with small deviations from the original recursive identification. Hence, this identification scheme should break the recursive scheme only as much as necessary to fulfill the economic restriction and — at the same time — to prevent biased VAR parameter estimates. The new optimization problem is given by:

$$\begin{aligned} \underset{\text{vec}(\boldsymbol{\beta}),\ \boldsymbol{\theta}}{\text{minimize}} \quad & \frac{1}{T} \text{vec}(\mathbf{U})^\top \text{vec}(\mathbf{U}) + \lambda \|\boldsymbol{\theta}\|_p \\ \text{subject to} \quad & \mathbf{U} = \mathbf{Y} - \mathbf{Z}\boldsymbol{\beta}, \\ & \mathcal{S}_{t+i} \cdot \widetilde{\Psi}_{t+i} - \mathcal{R}_{t+i} \leq 0, \quad \forall i = 1, \ldots, \zeta, \end{aligned} \tag{9}$$

where $\widetilde{\Psi}_t$ is now the IRF computed with the rotated structural shocks and λ is a data-dependent hyperparameter that controls the trade-off between residual minimization and recursive identification. The bootstrap procedure is identical to the one in Algorithm 1.

The described method can also be applied to arbitrary point identification schemes represented by a matrix P with $PP' = \Sigma_U$. It is particularly useful in situations where economically meaningful IRFs are not feasible under strict enforcement of the identification scheme. We continue to use the Cholesky factorization as a traditional target to shrink to.

3 Simulation Study

In the following, we investigate the small sample properties of our restricted VAR specifications (without and with an additional rotation of the Cholesky factorization) using a Monte Carlo study and compare it to the performance of the standard (unrestricted) case. We consider a two-dimensional VAR model with $p = 2$ lags as DGP, that is, we sample from

$$x_t = A_1 x_{t-1} + A_2 x_{t-2} + u_t, \tag{10}$$

with $(A_k)_{i,j} = a_{i,j}^{(k)}$. The DGP is calibrated based on the industrial production and interest rate time series that are also used in the empirical application in Section 4. Accordingly, the parameter matrices in (10) are set as:

$$A_1 = \begin{pmatrix} 0.83 & 0.89 \\ 0.05 & 1.14 \end{pmatrix} \quad \text{and} \quad A_2 = \begin{pmatrix} 0.17 & -1.08 \\ -0.03 & -0.24 \end{pmatrix}.$$

Table 1 shows average estimates of 1000 Monte Carlo replications and the corresponding root mean squared error (RMSE) in parentheses for three different sample sizes $T = 100, 250, 1000$. Each simulated sample has been drawn after discarding 100 burn-in periods. The DGP features the price puzzle that is sometimes observed empirically in monetary policy transmission models. In the given setting, this results because of the MA coefficient $(\Psi_1)_{1,2} = a_{1,2}^{(1)} P_{2,2} = 0.89 * 0.25 > 0$ being positive while $(\Psi_t)_{1,2} < 0$ for $t > 5$.

The benchmark estimator is given by a recursively identified VAR(2) model referred to as "Unrestr." in the following. The most promising scenario for the constrained estimator arises when the IRF puzzle in the standard VAR model occurs due to an incorrectly predicted sign of $(\Psi_t)_{1,2}$. We reproduce this setting by running a restricted VAR estimation with the first impulse response required to be positive (column "Restricted +"). The economic sign restriction is then in line with the DGP. In addition, we compute an estimator constraining the first impulse response to be negative (column "Restricted −"). Such a misspecification could arise if the sign restriction on the IRF is correct, but the puzzle results from a non-recursive identification scheme and not from incorrectly estimated parameters.

Table 1 shows that the performance of the correctly specified restricted estimator (with a positive sign restriction) does not crucially differ from the one of the unrestricted VAR. The results are even identical in the larger samples with $T = 250$ and $T = 1000$. This implies that the restricted estimator successfully solves the potential impulse response puzzle while the estimation performance does not deteriorate in comparison to the standard VAR estimator.

Table 1: Results of Simulation Study

	True	$T = 100$				$T = 250$				$T = 1000$			
		Unrestr.	Restricted			Unrestr.	Restricted			Unrestr.	Restricted		
			+	–	Rot.		+	–	Rot.		+	–	Rot.
$a_{1,1}^{(1)}$	0.830	0.792	0.792	0.834	0.792	0.820	0.820	0.858	0.819	0.828	0.828	0.864	0.828
		(0.109)	(0.109)	(0.107)	(0.109)	(0.065)	(0.065)	(0.073)	(0.065)	(0.031)	(0.031)	(0.047)	(0.031)
$a_{1,2}^{(1)}$	0.890	0.927	0.930	−0.004	0.927	0.892	0.892	0.000	0.892	0.885	0.885	0.000	0.885
		(0.446)	(0.438)	(0.895)	(0.446)	(0.262)	(0.262)	(0.890)	(0.262)	(0.130)	(0.130)	(0.890)	(0.130)
$a_{1,1}^{(2)}$	0.170	0.165	0.165	0.157	0.165	0.168	0.168	0.156	0.168	0.169	0.169	0.155	0.169
		(0.105)	(0.105)	(0.112)	(0.105)	(0.065)	(0.065)	(0.070)	(0.065)	(0.032)	(0.032)	(0.036)	(0.032)
$a_{1,2}^{(2)}$	−1.080	−1.124	−1.127	−0.327	−1.124	−1.093	−1.093	−0.306	−1.093	−1.075	−1.075	−0.285	−1.075
		(0.423)	(0.417)	(0.782)	(0.423)	(0.253)	(0.253)	(0.782)	(0.253)	(0.123)	(0.123)	(0.796)	(0.123)
$a_{2,1}^{(1)}$	0.050	0.052	0.052	0.054	0.052	0.051	0.051	0.052	0.051	0.051	0.051	0.052	0.051
		(0.025)	(0.025)	(0.025)	(0.025)	(0.015)	(0.015)	(0.015)	(0.015)	(0.008)	(0.008)	(0.008)	(0.008)
$a_{2,2}^{(1)}$	1.140	1.099	1.099	1.068	1.100	1.128	1.128	1.098	1.128	1.134	1.134	1.105	1.135
		(0.109)	(0.109)	(0.124)	(0.109)	(0.066)	(0.066)	(0.078)	(0.066)	(0.031)	(0.031)	(0.047)	(0.031)
$a_{2,1}^{(2)}$	−0.030	−0.024	−0.024	−0.024	−0.024	−0.028	−0.028	−0.028	−0.028	−0.030	−0.030	−0.030	−0.030
		(0.026)	(0.026)	(0.026)	(0.026)	(0.016)	(0.016)	(0.016)	(0.016)	(0.008)	(0.008)	(0.008)	(0.008)
$a_{2,2}^{(2)}$	−0.240	−0.244	−0.244	−0.217	−0.244	−0.243	−0.243	−0.217	−0.243	−0.238	−0.238	−0.212	−0.238
		(0.095)	(0.094)	(0.097)	(0.095)	(0.062)	(0.062)	(0.066)	(0.062)	(0.030)	(0.030)	(0.041)	(0.030)

Notes: Table shows average parameter estimates and root mean squared errors (in parentheses) of 1000 Monte Carlo draws for different estimation procedures and sample sizes $T = 100, 250, 1000$. Unrestricted results are derived with the maximum likelihood estimator. Restricted results are derived from a constrained VAR model with positive (+) or negative (–) sign restrictions on the first impulse response. Results for a constrained VAR model with negative sign restriction that additionally allows for a rotated identification are shown in column "Rot."

However, the constrained VAR with a negative sign restriction is misspecified as the true parameter is not included in the restricted parameter space. This causes considerably biased estimates as shown in the columns "Restricted −" of Table 1. Hence, the recursive identification with parameter restrictions clearly has some limitations if the economic restriction on the impulse response is not fulfilled by the DGP.

This potential estimation bias stems from the inflexibility of the recursive identification scheme. In the columns labeled "Restricted Rot.", we further report results of a constrained estimator that allows rotating the Cholesky factorization (which has been fixed so far) in addition to the negative sign restriction. The results reveal that the estimation bias largely vanishes when additionally rotating the covariance matrix factorization. Even the RMSEs show similar values as the unrestricted VAR. This is particularly remarkable as the employed L2-penalization yields a rotation of, on average, only 13 degrees. Hence, the procedure that uses economic restrictions and rotations leads to an identification that is close to the recursive scheme while avoiding an estimation bias and an impulse response puzzle.

To summarize, if the economic parameter constraints are valid for the DGP, the restricted (but unrotated) VAR produces consistent estimates while having a similar small sample performance as the standard VAR. However, if the DGP does not fulfill the economic restrictions, we can only enforce the constraints by allowing for a (substantial) estimation bias. In this case, the "almost recursively identified approach" that slightly rotates the initial propagation of shocks can prevent the bias while generating theory-consistent impulse responses.

4 Empirical Application

In the following, we estimate a standard monetary policy VAR for the euro area and the period January 1999–December 2019.

4.1 Data and Preliminary Steps

Our data set includes the (i) the industrial production index (IP, in logs), (ii) the harmonized index of consumer prices inflation rate, and (iii) a composite indicator of the monetary policy stance as endogenous variables. As monetary policy indicator, we use the European Central Bank's main refinancing operations rate (MRO rate) until October 2008.[4] After that date, we replace the MRO rate with the shadow short rate by Krippner (2015), which provides a quantification of all unconventional monetary policy measures in a single shadow interest rate and also allows for negative interest rates. In our view, this is the most parsimonious description of monetary policy in normal times and crisis times in a single variable.[5] Figure A1 in the Appendix plots the three series over time.

All variables enter the VAR model linearly de-trended, following Burgard et al. (2019).[6] The selection of the lag structure is based on three criteria. First, there should be no autocorrelation left in the residuals of the VAR models. Second, the IRFs should converge to zero, at least asymptotically. Third, either model should be as parsimonious as possible, that is, redundant (i.e., insignificant) lags should be removed. Employing three lags yields no autocorrelation in the standard VAR as well as in the restricted and rotated VAR. However, there would still be autocorrelation left at the 10% significance level in the industrial production equation of the unrotated VAR with estimations restrictions on the first two, three, and six lags. To ensure comparability across models and to avoid any risk of remaining serial correlation in the residuals, we choose a lag length of four for all VARs in the following empirical application. In this case, Ljung-Box (1978) tests indicate no autocorrelation at any reasonable significance level for all models considered. Including additional lags in either model only leads to a less sharp identification of the IRFs due to a loss in the degrees of freedom.

[4]Note that replacing the MRO rate with the EONIA leaves the results virtually unchanged.

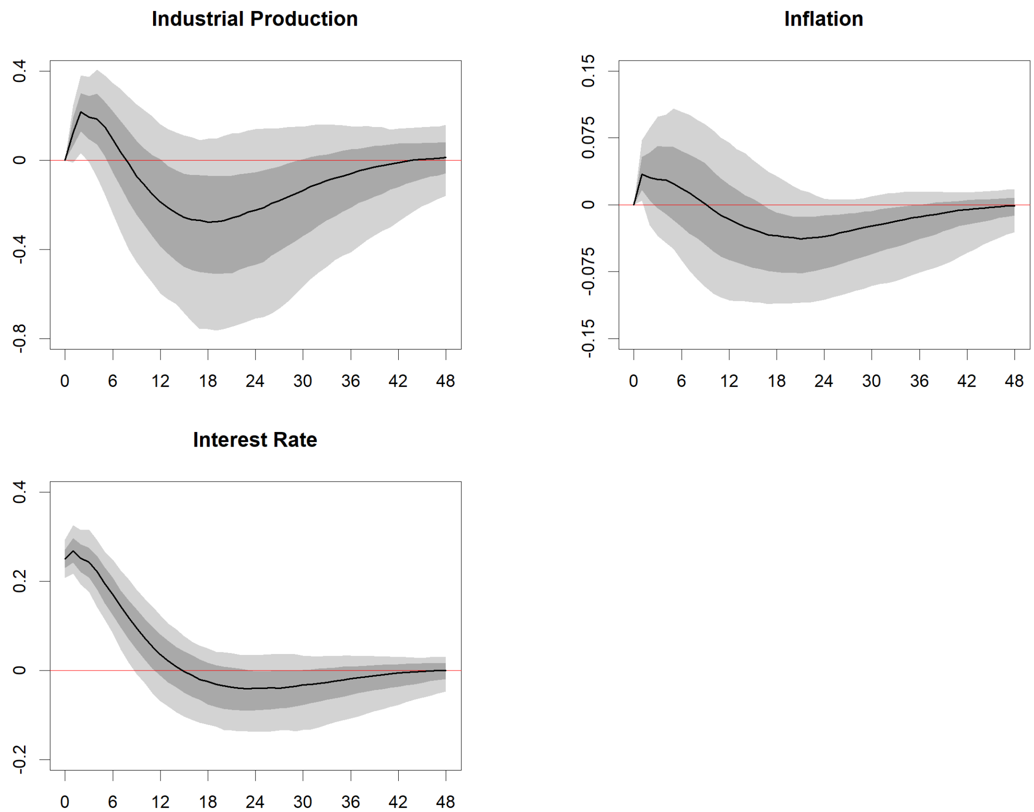
[5]We decided not to use the shadow rate by Wu and Xia (2016) for several reasons. During the recent years, this variables has be subject to criticism (e.g., Krippner, 2020) since its underlying assumptions lead to rather "extreme" values (e.g., almost –8% for the euro area in 2019). In addition, employing this variable in our context leads to highly persistent monetary policy shocks that die out very slowly.

[6]Note that this is equivalent to including a linear trend in the VAR system. The trends are statistically significant in all three cases at the 5% level.

4.2 Empirical Results

4.2.1 Benchmark

Figure 1: Impulse Responses of Recursively Identified VAR



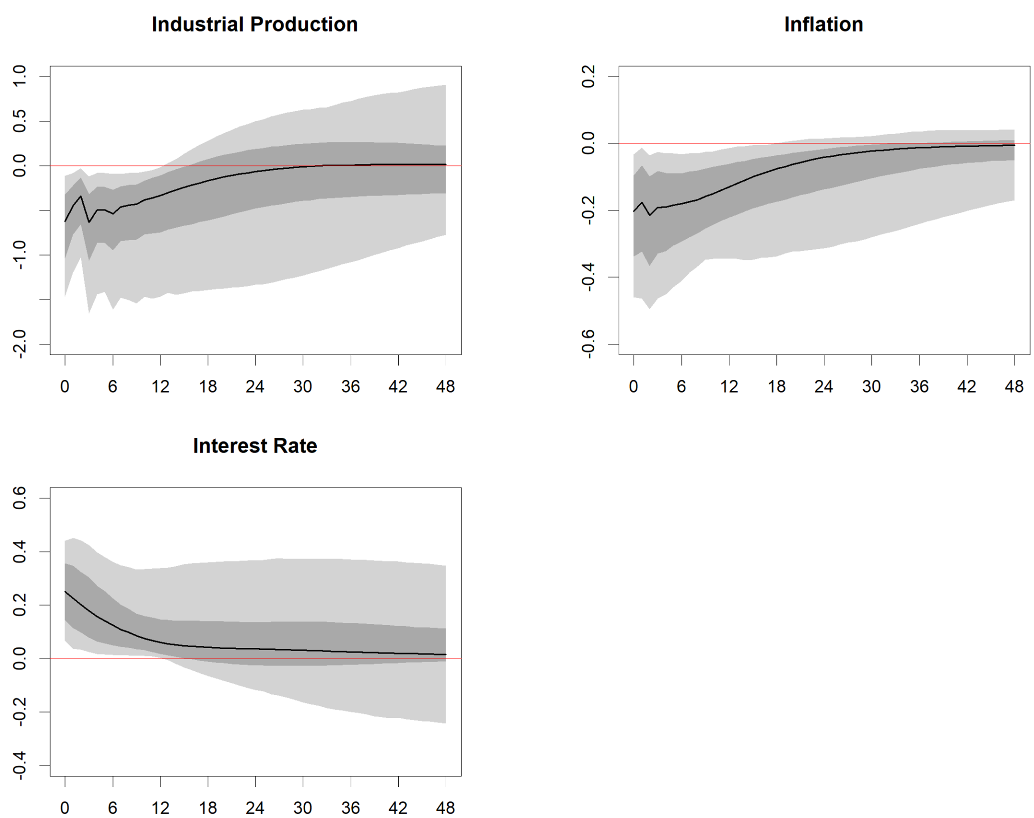
Notes: Solid lines represent median IRFs (in percentage points) to a contractionary monetary policy shock of 25 basis points. Dark gray shaded (light gray shaded) areas indicate 68% (95%) confidence bands derived by bootstrapping and 500 replications. Cholesky decomposition is based on the ordering: (i) industrial production, (ii) inflation, and (iii) interest rate.

Figure 1 shows median IRFs after a contractionary monetary policy shock of 25 bps in a standard recursively identified VAR. Whereas the medium-term responses of output (13–29 months after the shock) and inflation (17–35 months after the shock) to a contractionary monetary policy shock are well in line with a priori expectations, the short-term responses are not. Output and prices increase and we have evidence for the well-known "price puzzle" (Eichenbaum, 1992).[7]

[7] In addition to this empirical issue, it is also debatable from a theoretical point of view that prices do not respond contemporaneously to monetary policy shocks. In a standard New Keynesian model, some firms are able to react contemporaneously to shocks and others are not, which implies that prices are

Various approaches such as, for instance, the inclusion of commodity prices (Sims, 1992) or identification with the help of high frequency data (Faust et al., 2004) have been employed to get rid of these initially counterintuitive responses. Since Uhlig (2005), the most commonly used method is to restrict the responses of some variables in the VAR system to be positive or negative for a certain period of time or to employ a penalty function approach. That is, identification is based on IRFs that satisfy the restrictions imposed by the researcher. All counterintuitive responses are discarded (or penalized) in the process of identifying a monetary policy shock.[8]

Figure 2: Impulse Responses of Sign-Restricted VAR



Notes: Solid lines represent median IRFs (in percentage points) to a contractionary monetary policy shock of 25 basis points. Dark gray shaded (light grey shaded) areas indicate 68% (95%) credible sets based on 5,000 accepted MCMC draws. Responses of output and prices (the interest rate) after a contractionary monetary policy shock are assumed to be negative (positive) for 12 months after the shock.

sticky (but not fixed). Nevertheless, zero restrictions on impact for prices (and output) after a monetary policy shock are also assumed in other recent papers (e.g., Peersman, 2011; Gambacorta et al., 2014; Galí and Gambetti, 2015).

[8]Uhlig (2017) provides an intuitive introduction into identification via sign restrictions.

Figure 2 shows median IRFs after a contractionary monetary policy shock of 25 bps for the same VAR model as in Figure 1, but based on sign restrictions. The shock is assumed to decrease output and prices on impact and for 12 months thereafter and to increase the interest rate for the same period (following Principle 16 on p. 124 in Uhlig, 2017). Even with this less "agnostic" identification scheme as the one in Uhlig (2005), the response of output becomes "insignificant" quickly once the restrictions are lifted. In general, the responses are less precisely measured as in the recursive identification scheme.

As illustrated by the two sets of impulse responses, both, recursive identification and identification via sign restrictions have some drawbacks, which our proposed approaches aim to overcome.

4.2.2 Impulse Responses with Restrictions

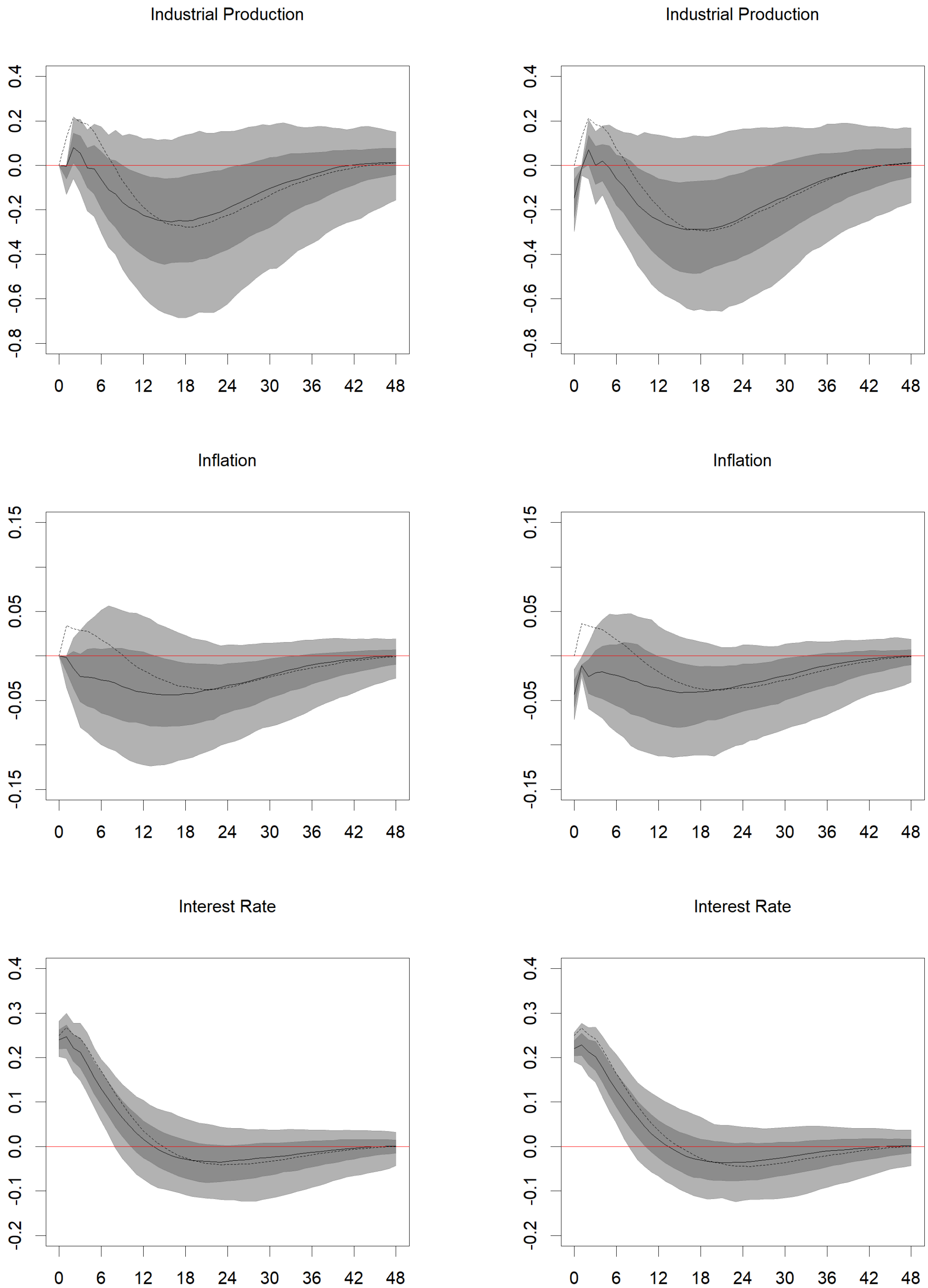
In a first step, we estimate a linear VAR model under the restriction that the IRFs for industrial production and inflation should be non-positive during the *first* month after a contractionary monetary policy shock.[9] In addition, the response of the interest rate indicator is restricted to be non-negative on impact and during the *first* month.[10] Figure 3 shows the results with the dashed lines replicating the median IRFs of the unrestricted model in Figure 1. The left panel shows the results with the sign restrictions only, whereas in the right panel we additionally allow for a rotation of the Cholesky factorization. Figure 3 illustrates that this restriction is already sufficient to get rid of the "price puzzle" in both cases. It is also sufficient to prevent a significant positive response of industrial production one month after the shock. However, two months after the shock we still find a (brief) positive response in both panels.

[9] Note that the instantaneous response of both variables is already restricted to zero due to the recursive identification scheme in the left panel of Figure 3.

[10] Note that this restriction is implemented for comparability reasons to the sign restriction approach in Bayesian VAR but non-binding in our application.

Figure 3: Impulse Responses with 1-Month Restriction

Panel A: No Rotation **Panel B: Rotation of Cholesky Matrix**



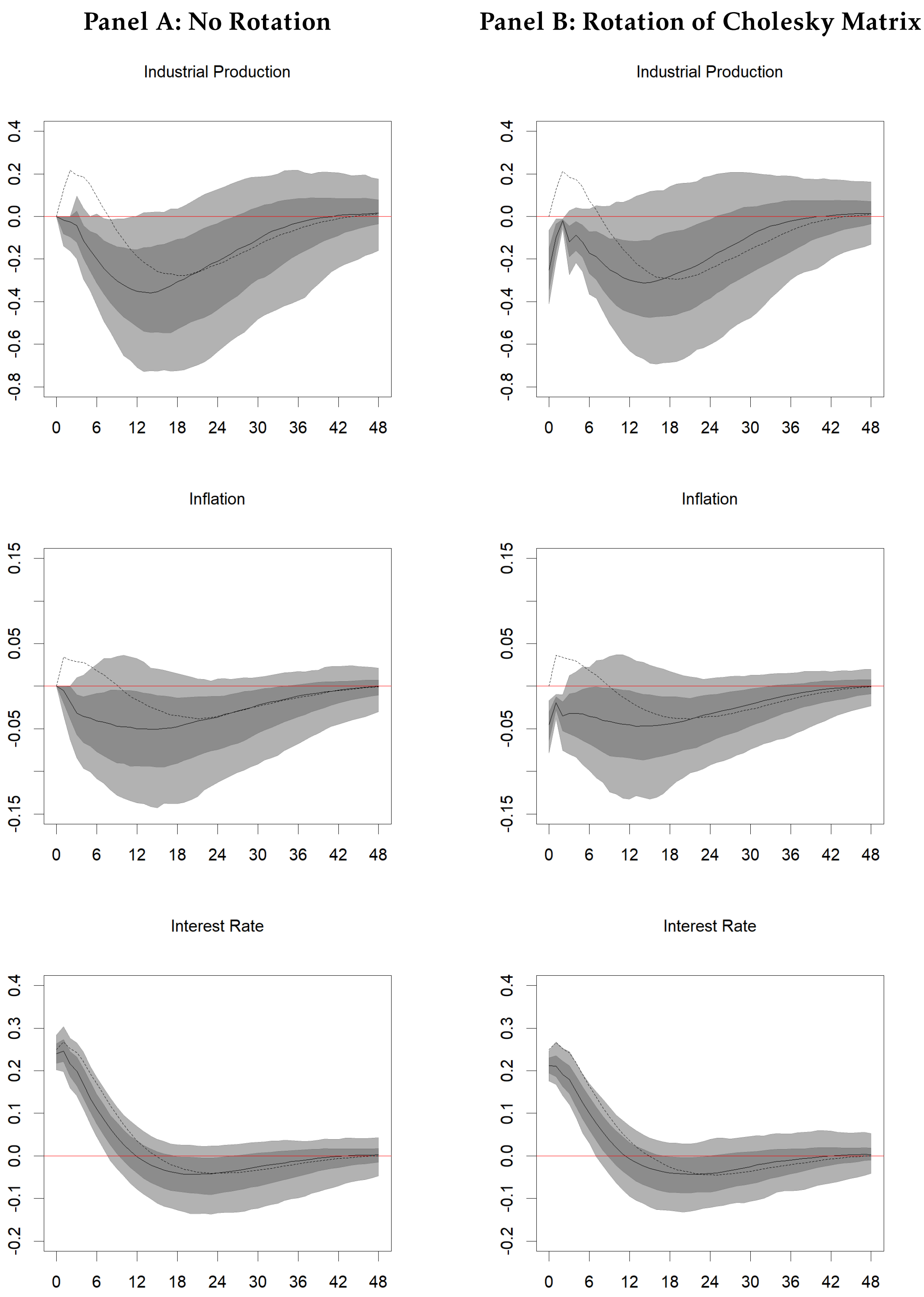
Notes: Solid lines represent median IRFs (in percentage points) to a contractionary monetary policy shock of 25 basis points without (panel A) and with (panel B) a rotation of the Cholesky matrix in addition to the parameter restrictions. Dark gray shaded (light gray shaded) areas indicate 68% (95%) confidence bands derived by bootstrapping and 500 replications. Dashed lines show unrestricted median IRFs taken from Figure 1.

Consequently, we increase the number of restricted months to two and, in a second step, estimate a linear VAR model under the restriction that the IRFs for industrial production and inflation (the interest rate) should be non-positive (non-negative) during the first *two* months after a contractionary monetary policy shock. Figure 4 shows the results with the dashed lines replicating the median IRFs of the unrestricted model in Figure 1.

Again, the left panel shows the results with sign restrictions only, whereas in the right panel we additionally allow for a rotation of the Cholesky factorization. A restriction on the first two months is indeed sufficient for obtaining non-positive median responses for industrial production and inflation in both panels. Figure 4 also documents that the differences between the original IRFs and the restricted ones are only visible during the first 18–24 months after the shock. After two years, the median IRFs are virtually the same for all variables and in both panels.

Figures A2 and A3 in the Appendix show the IRFs with restrictions on the first *three* and *six* months, respectively. In these figures, the dashed lines resemble the median IRFs from Figure 4 with restrictions on the first *two* months. It becomes evident that more persistent restrictions do not induce a substantial difference to the case with a restriction on the first two months, which is the "least invasive" case that yields theory-conform IRFs. Consequently, the following discussion will refer to the unrestricted case and the case with restrictions on the first two months (without and with a rotation of the Cholesky factorization).

Figure 4: Impulse Responses with 2-Month Restriction



Notes: Solid lines represent median IRFs (in percentage points) to a contractionary monetary policy shock of 25 basis points without (panel A) and with (panel B) a rotation of the Cholesky matrix in addition to the parameter restrictions. Dark gray shaded (light gray shaded) areas indicate 68% (95%) confidence bands derived by bootstrapping and 500 replications. Dashed lines show unrestricted median IRFs taken from Figure 1.

4.2.3 Discussion

Table 2 shows the months, during which the responses to a contractionary monetary policy shock are significant when considering 68% confidence bands.

Table 2: Significance of Impulse Responses

	Unrestricted	**Restricted (Two Months)**	
		No Rotation	**Rotation**
Industrial Production	13 – 29	1 – 2; 4 – 26	0 – 24
Inflation	17 – 35	1 – 33	0 – 32
Interest Rate	0 – 11	0 – 9	0 – 8

Notes: Table shows the lags, during which the responses to a contractionary monetary policy shock are significant when considering 68% confidence bands.

The IRFs of industrial production and inflation become significant much earlier in both restricted cases, but also become insignificant three to five months earlier. The IRFs of the interest rate itself is affected to a smaller extent as the ones for the restricted cases become insignificant two to three months earlier; that is, monetary policy shocks are found to be less persistent under parameter restrictions. Finally, it is also worth noting that the almost recursive identification in the right column allows for an instantaneous non-zero response of output and prices to a monetary policy shock and is therefore less restrictive and in line with the implications of the standard New Keynesian model.

Table 3 shows the peak responses to a contractionary monetary policy shock (in percentage points) alongside the month during which the response was found in parentheses.

Table 3: Peak Responses

	Unrestricted	**Restricted (Two Months)**	
		No Rotation	**Rotation**
Industrial Production	–0.28 (18)	–0.36 (14)	–0.31 (14)
Inflation	–0.038 (21)	–0.051 (15)	–0.047 (13)

Notes: Table shows the peak responses to a contractionary monetary policy shock (in percentage points) alongside the month during which the response was found in parentheses.

The size of the peak responses is slightly larger when the estimation restrictions are in place, in particular for the unrotated case. A contractionary monetary policy shock of 25 bps leads to a reduction of 28–36 basis points in industrial production and to a decrease of 3.8–5.1 basis points in the inflation rate. The peak for industrial production is found 18 months after the shock in the unrestricted case and after 14 months in both restricted cases. The corresponding peaks for inflation are found after 21 months (unrestricted) and 13–15 months (restricted cases). Hence, similar to the overall significance, the peaks are four to eight months earlier due to the restriction in the estimation process. All estimated peaks, however, are in line with the usual estimates for the outside lag in the transmission of monetary policy shocks to which central bankers often refer (e.g., the European Central Bank in its monetary policy strategy).

Least squares estimation is naturally providing the best fit to the data in terms of the SSR. Any restriction in the estimation process leads to an increase in this measure. To evaluate the size of the loss in fit, Table 4 shows the decrease in fit (in %) of the restricted models as compared to the unrestricted one. The loss in fit amounts to 2.45% and 1.73% in the industrial production equation and the inflation equation in the case of the restricted, but unrotated model.[11] This corroborates the results of the simulation study where we detect a bias in the unrotated model with negative sign restrictions. In contrast, the loss in fit is negligible in the rotated model.

Table 4: Decrease in Fit (SSR in %)

	No Rotation	**Rotation**
Industrial Production	2.45	0.0009
Inflation	1.73	0.0004
Interest Rate	0.17	0.0003

Notes: Table shows the decrease (in %) in the sum of squared residuals in the restricted models as compared to the unrestricted model.

Finally, we test whether the individual coefficients are different in the restricted cases as compared to the unrestricted model. Table A1 in the Appendix illustrates this for both cases with, again, sign restrictions on the first two months. In the case of the restricted, but unrotated model we find that the first lag of industrial production is

[11] An in-depth analysis reveals that the decrease in fit is driven by the restriction on the first two lags in the case of the IP equation, whereas for inflation the restriction on the first lag is particularly relevant.

significantly different from the parameter of the unrestricted model at the 5% significance level. For the other coefficients of that specification, we find some numerically visible differences that are not statistically significant. However, the coefficients for the rotated model are virtually the same as compared to the unrestricted case, which is also in line with the negligible decrease in fit in Table 4 and the results of the simulation study.

To summarize, a restriction on the first two months in the IRFs of output, prices, and the interest rate is sufficient to get rid of the price puzzle and the counterintuitive response of output for both, the restricted VAR without and with a rotation of the Cholesky factorization. However, the almost recursively identified approach outperforms the Cholesky identification as it features virtually no change in the model parameters and no noteworthy decrease in the model's fit.

5 Conclusions

Recursively identified vector autoregressive models sometimes lead to theoretically counterintuitive impulse responses. As a first step to overcome this problem, we propose to estimate the VAR parameters under the restriction that economic theory is not violated, but stick to a recursive identification of shocks. We solve this optimization problem under non-linear constraints using an augmented Lagrange solution approach and adjust the VAR coefficients so that these meet the theoretical requirements. In a second step, we allow for a (minimal) rotation of the Cholesky matrix in addition to the parameter restrictions.

Using a Monte Carlo simulation, we show that — if the economic parameter constraints are valid for the DGP — the restricted (but unrotated) VAR produces consistent estimates while having a similar small sample performance as the standard VAR. However, if the DGP does not fulfill the economic restrictions, we can only enforce the constraints by allowing for a (substantial) estimation bias. In this case, the "almost recursively identified approach" that slightly rotates the initial propagation of shocks can prevent the bias while generating theory-consistent impulse responses.

As an empirical application, we estimate a standard monetary policy VAR for the euro area and the period January 1999–December 2019. Our results indicate that a restriction on the first two months in the IRFs of output, prices, and the interest rate is sufficient to get rid of the price puzzle and the counterintuitive response of output for both, the restricted VAR without and with a rotation of the Cholesky factorization. However, the almost recursively identified approach outperforms the Cholesky identification as it features virtually no change in the model parameters and no noteworthy decrease in the model's fit. The additional flexibility of this approach is underscored by the possibility of having non-zero responses on impact for output and inflation, which is by definition ruled out in the recursive identification scheme.

The key implication of our paper is to consider the "almost recursively identified approach with parameter restrictions" as a useful complement to the Bayesian sign restriction approach commonly used in empirical macroeconomic studies. This approach leads to a solution without an estimation bias, generates theory-consistent impulse responses, and is as close as possible to the recursive scheme.

References

Burgard, J. P., M. Neuenkirch, and M. Nöckel (2019). "State-Dependent Transmission of Monetary Policy in the Euro Area." *Journal of Money, Credit and Banking* 51 (7), 2053–2070.

Eichenbaum, M. (1992). "Interpreting Macroeconomic Time Series Facts: The Effects of Monetary Policy: Comments." *European Economic Review* 36 (5), 1001–1011.

Faust, J., E. T. Swanson, and J. H. Wright (2004). "Identifying VARs Based on High Frequency Futures Data." *Journal of Monetary Economics* 51 (6), 1107–1131.

Gafarov, B., M. Meier, and J. L. Montiel Olea (2018). "Delta-Method Inference for a Class of Set-Identified SVARs." *Journal of Econometrics* 203 (2), 316–327.

Galí, J. and L. Gambetti (2015). "The Effects of Monetary Policy on Stock Market Bubbles: Some Evidence." American Economic Journal: Macroeconomics 7 (1), 233–257.

Gambacorta, L., Hofmann, B., and Peersman, G. (2014). "The Effectiveness of Unconventional Monetary Policy at the Zero Lower Bound: A Cross-Country Analysis." *Journal of Money, Credit and Banking* 46 (4), 615–642.

Krippner, L. (2015). "*Zero Lower Bound Term Structure Modeling: A Practitioner's Guide*." Palgrave Macmillan: New York.

Krippner, L. (2020). "A Note of Caution on Shadow Rate Estimates." *Journal of Money, Credit and Banking* 52 (4), 951–962.

Ljung, G. M. and G. E. P. Box (1978). "On a Measure of Lack of Fit in Time Series Models." *Biometrika* 65 (2), 297–303.

Lütkepohl, H. (2000). "Bootstrapping Impulse Responses in VAR Analyses." in: J. G. Bethlehem and P. G. M. Heijden (Eds.): "*COMPSTAT: Proceedings in Computational Statistics 2000*." Springer: Berlin/Heidelberg, 109–119.

Lütkepohl, H. (2005). "*New Introduction to Multiple Time Series Analysis*." Springer: Berlin/Heidelberg.

Peersman, G. (2011). "Macroeconomic Effects of Unconventional Monetary Policy in the Euro Area." *ECB Working Paper* 1397.

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

Uhlig, H. (2005). “What are the Effects of Monetary Policy on Output? Results from an Agnostic Identification Procedure.” *Journal of Monetary Economics* 52 (2), 381–419.

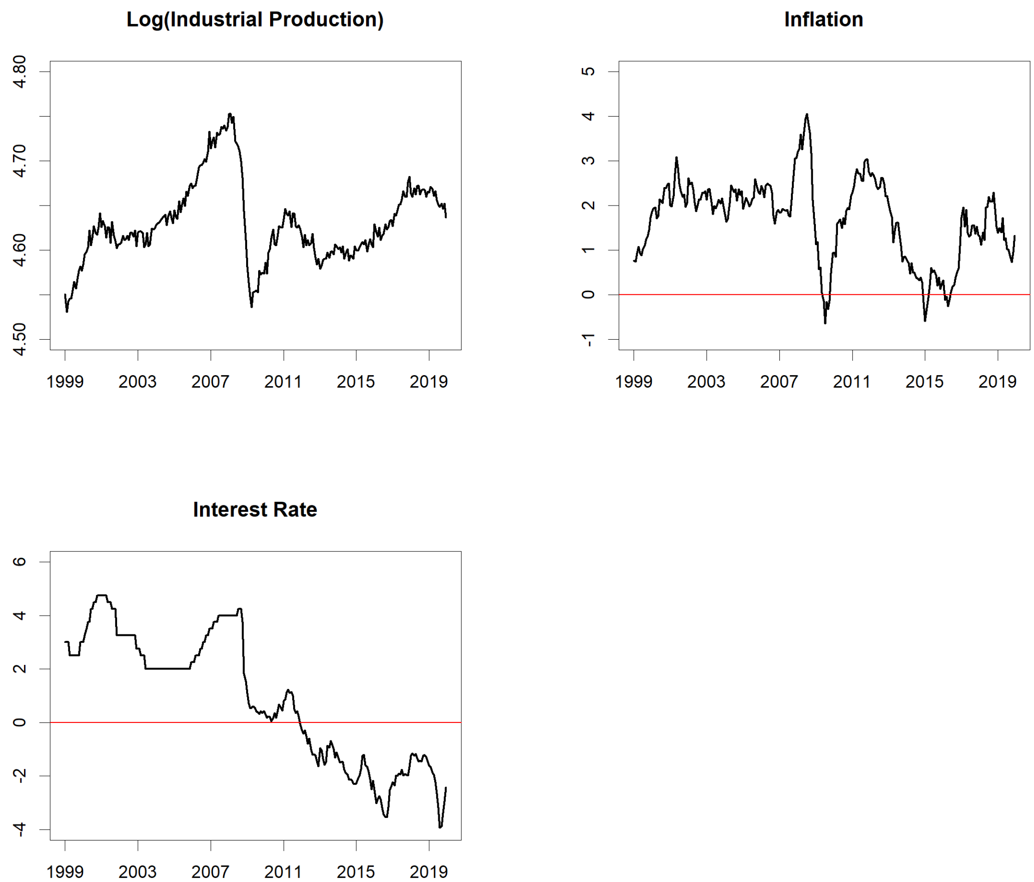
Uhlig, H. (2017). “Shocks, Sign Restrictions, and Identification.” in: B. Honore, A. Pakes, M. Piazzesi, and L. Samuelson (Eds.): “*Advances in Economics and Econometrics*.” Cambridge: Cambridge University Press, 95–127.

Varadhan, R. (2015). “alabama: Constrained Nonlinear Optimization.” R package Version 2015.3-1. `https://CRAN.R-project.org/package=alabama`.

Wu, J. C. and F. D. Xia. (2016). “Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound.” *Journal of Money, Credit and Banking* 48 (2–3), 253–291.

Appendix

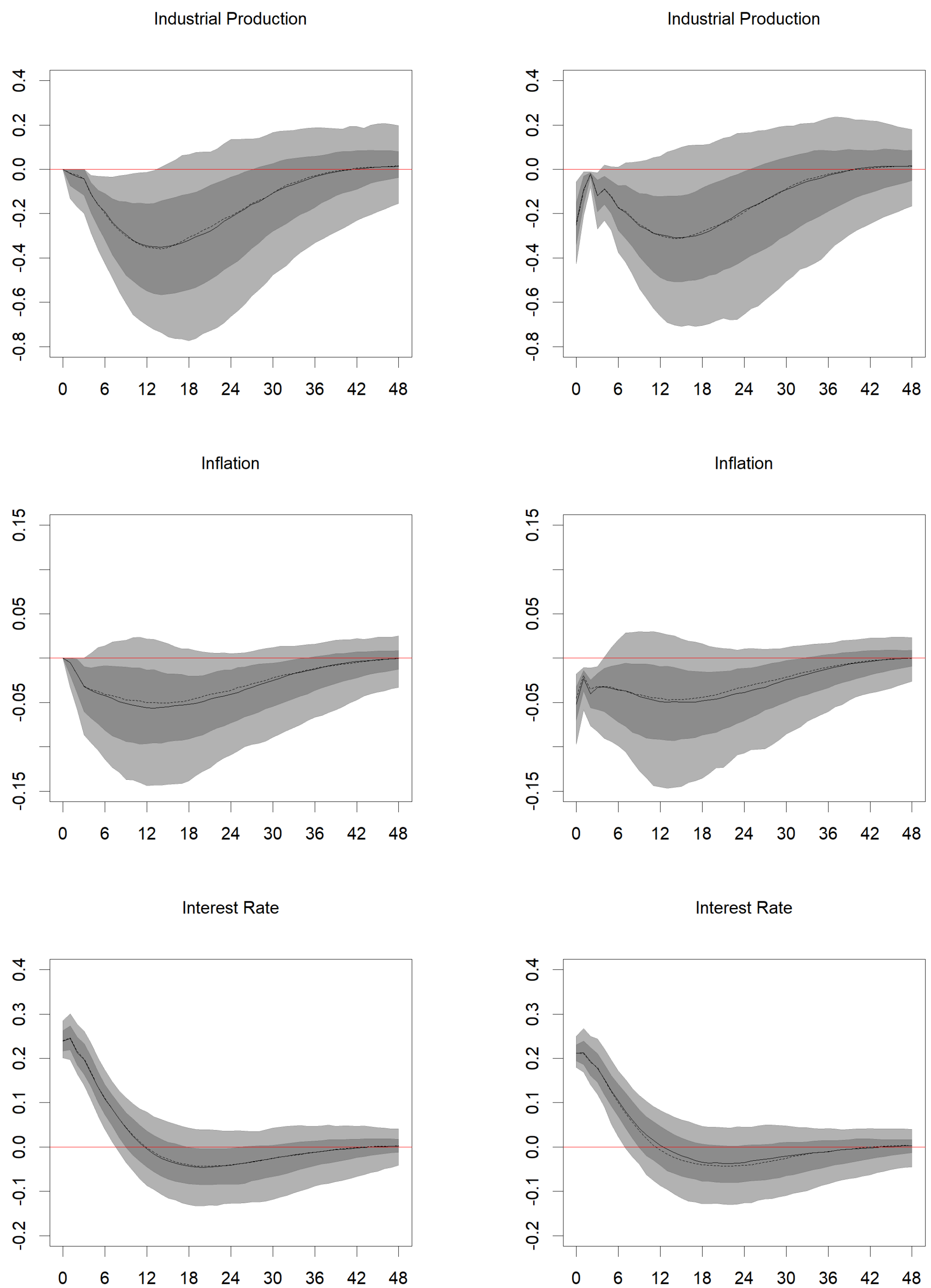
Figure A1: Macroeconomic Variables for the Euro Area



Notes: The "interest rate" is a composite indicator based on the main refinancing rate (until October 2008) and the shadow short rate by Krippner (2015) (from November 2008 onwards) to reflect the European Central Bank's unconventional monetary policy measures since the onset of the Global Financial Crisis. Data are taken from the European Central Bank (industrial production, inflation, and main refinancing operations rate) and Leo Krippner's homepage (shadow short rate, www.ljkmfa.com).

Figure A2: Impulse Responses with 3-Month Restriction

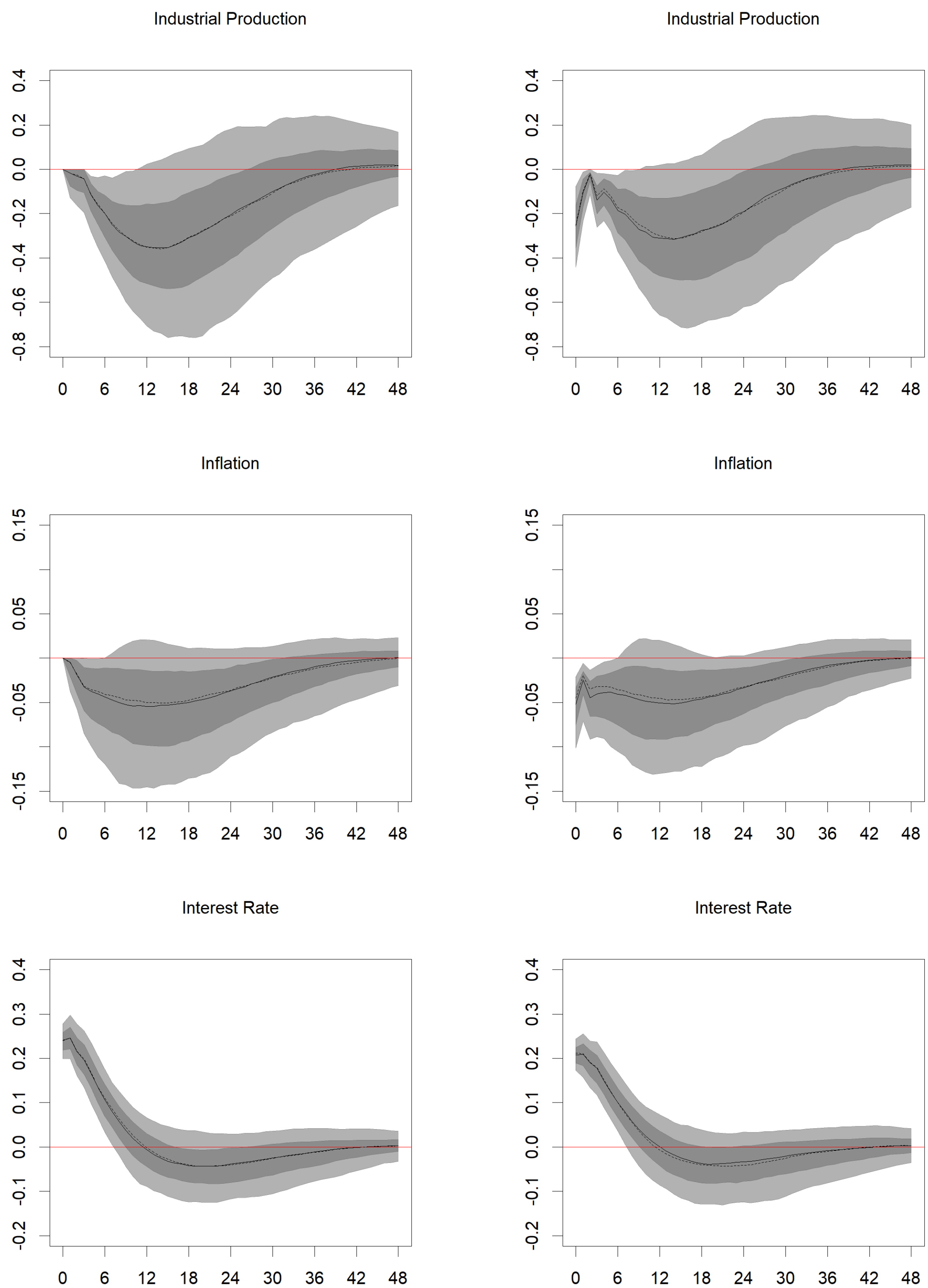
Panel A: No Rotation **Panel B: Rotation of Cholesky Matrix**



Notes: Solid lines represent median IRFs (in percentage points) to a contractionary monetary policy shock of 25 basis points without (panel A) and with (panel B) a rotation of the Cholesky matrix in addition to the parameter restrictions. Dark gray shaded (light gray shaded) areas indicate 68% (95%) confidence bands derived by bootstrapping and 500 replications. Dashed lines show median IRFs taken from Figure 4 with a restriction on the first *two* months only.

Figure A3: Impulse Responses with 6-Month Restriction

Panel A: No Rotation **Panel B: Rotation of Cholesky Matrix**



Notes: Solid lines represent median IRFs (in percentage points) to a contractionary monetary policy shock of 25 basis points without (panel A) and with (panel B) a rotation of the Cholesky matrix in addition to the parameter restrictions. Dark gray shaded (light gray shaded) areas indicate 68% (95%) confidence bands derived by bootstrapping and 500 replications. Dashed lines show median IRFs taken from Figure 4 with a restriction on the first *two* months only.

Table A1: Comparison of Coefficients

	Industrial Production			**Inflation**			**Interest Rate**		
	Unrestr.	Restricted		Unrestr.	Restricted		Unrestr.	Restricted	
		No Rot.	Rotation		No Rot.	Rotation		No Rot.	Rotation
IP_{t-1}	0.732	0.764	0.733	0.012	0.017	0.012	0.046	0.048	0.046
	(0.065)	(0.062)	(0.064)	(0.017)	(0.017)	(0.017)	(0.016)	(0.017)	(0.016)
IP_{t-2}	0.194	0.208	0.194	0.029	0.032	0.029	−0.019	−0.018	−0.019
	(0.078)	(0.080)	(0.078)	(0.021)	(0.021)	(0.021)	(0.020)	(0.019)	(0.019)
IP_{t-3}	0.284	0.278	0.283	−0.018	−0.021	−0.018	−0.005	−0.006	−0.005
	(0.078)	(0.081)	(0.075)	(0.021)	(0.021)	(0.020)	(0.020)	(0.020)	(0.021)
IP_{t-4}	−0.208	−0.233	−0.209	−0.011	−0.014	−0.011	0.000	−0.002	0.000
	(0.064)	(0.065)	(0.063)	(0.017)	(0.018)	(0.016)	(0.016)	(0.016)	(0.016)
$Infl._{t-1}$	0.451	0.574	0.454	1.026	1.050	1.027	0.078	0.086	0.078
	(0.247)	(0.242)	(0.243)	(0.066)	(0.067)	(0.069)	(0.063)	(0.063)	(0.062)
$Infl._{t-2}$	−0.472	−0.511	−0.472	0.022	0.007	0.022	−0.092	−0.096	−0.093
	(0.349)	(0.345)	(0.345)	(0.094)	(0.093)	(0.097)	(0.088)	(0.085)	(0.087)
$Infl._{t-3}$	0.294	0.229	0.291	−0.211	−0.220	−0.211	0.040	0.037	0.040
	(0.349)	(0.358)	(0.356)	(0.094)	(0.095)	(0.091)	(0.088)	(0.090)	(0.088)
$Infl._{t-4}$	−0.491	−0.498	−0.490	0.078	0.079	0.078	−0.049	−0.049	−0.048
	(0.250)	(0.251)	(0.257)	(0.067)	(0.068)	(0.065)	(0.063)	(0.063)	(0.064)
IR_{t-1}	0.493	0.000	0.486	0.143	0.000	0.142	1.095	1.056	1.095
	(0.263)	(0.175)	(0.271)	(0.071)	(0.047)	(0.072)	(0.067)	(0.064)	(0.068)
IR_{t-2}	−0.098	0.000	−0.100	−0.177	−0.039	−0.176	−0.183	−0.155	−0.182
	(0.382)	(0.313)	(0.370)	(0.103)	(0.082)	(0.102)	(0.097)	(0.090)	(0.095)
IR_{t-3}	−0.352	−0.002	−0.344	0.011	0.001	0.011	0.058	0.066	0.058
	(0.384)	(0.353)	(0.370)	(0.103)	(0.091)	(0.098)	(0.097)	(0.093)	(0.090)
IR_{t-4}	−0.175	−0.203	−0.175	0.017	0.020	0.017	−0.070	−0.070	−0.070
	(0.268)	(0.252)	(0.267)	(0.072)	(0.063)	(0.070)	(0.068)	(0.064)	(0.065)
Constant	0.027	0.027	0.027	0.003	0.003	0.003	0.006	0.006	0.006
	(0.056)	(0.076)	(0.078)	(0.015)	(0.021)	(0.023)	(0.014)	(0.020)	(0.019)

Notes: Table shows coefficients (with standard errors in parentheses) of the unrestricted VAR model (see also Figure 1) and the VAR models without and with rotation of the Cholesky matrix where restrictions are in place for the first *two* months (see also Figure 4).