UNIVERSITÄT BERN

Faculty of Business, Economics and Social Sciences

Department of Economics

Impulse Response Analysis at the Zero Lower Bound

Luca Benati, Thomas A. Lubik

23-06

March, 2023

DISCUSSION PAPERS

Schanzeneckstrasse 1
CH-3012 Bern, Switzerland
http://www.vwi.unibe.ch

Impulse Response Analysis at the Zero Lower Bound*

Luca Benati
University of Bern†

Thomas A. Lubik
Federal Reserve Bank of Richmond‡

March 24, 2023

Abstract

We study whether the response of the economy to structural shocks changes at the zero lower bound. Monte Carlo evidence suggests that VARs have a limited ability to detect changes in impulse response functions at the ZLB compared to the standard environment with positive interest rates. This issue is confounded given the short sample lengths that characterize ZLB episodes. This is especially the case for time-varying parameter VARs, whose estimates are two-sided, and therefore tend to smooth changes across regimes. In contrast, fixed-coefficient VARs estimated by sub-sample exhibit greater power. Pooled estimates from panel VARs for six countries based on (long-run and) sign restrictions detect in several instances changes in the IRFs. This evidence is, however, weaker than it appears. Based on (long-run and) sign restrictions we find that prior and posterior IRFs are often close, so that the concern raised by Baumeister and Hamilton (2015) appears to be relevant. Evidence from a multivariate permanent-transitory decomposition of GDP shocks is markedly sharper. It points towards material changes in the IRFs: at the ZLB the IRFs of GDP and unemployment exhibit more inertia, the response of prices is flatter, and the responses of interest rates are weaker.

JEL Classification: C32, C52

Keywords: Zero Lower Bound; Bayesian VARs; structural VARs; monetary policy; sign restrictions

*We wish to thank Federico Ravenna and Juan Rubio-Ramirez for useful comments, Giorgio Primiceri for sharing estimation codes, and Simon Thinggaard Hetland for sharing the shadow rate series for Denmark. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Richmond or the Federal Reserve System.

†Department of Economics, University of Bern, Schanzeneckstrasse 1, CH-3001, Bern, Switzerland. Email: luca.benati@vwi.unibe.ch

‡Research Department, P.O. Box 27622, Richmond, VA 23261. Email: thomas.lubik@rich.frb.org.

1 Introduction

The key response of central banks to the Global Financial Crisis (GFC) was to lower their policy rates to the zero lower bound (ZLB) and pursue additional stimulative measure by means of quantitative easing (QE), that is, the purchase of government and sometimes private assets with newly created money (or reserves). The ZLB period lasted well into the 2010s for most central banks and covered an even longer time span in the case of Japan. The last decade also saw very low inflation, solid, but not spectacular GDP growth and very strong labor markets, with unemployment falling below historical averages in many countries. This unusual constellation has given rise to the notion of "missing inflation" and a "flat Phillips curve". It naturally raises the question whether the dynamics of the economy at the ZLB are different from more standard times and whether QE can be considered a substitute for policy accommodation below the ZLB.

In this paper, we assess whether this is, in fact, the case and to what extent the data can be informative in resolving this question. We estimate a battery of statistical models in the class of Bayesian vector autoregressions (VARs), using a variety of specifications and identifying assumptions, for a group of advanced economies. We highlight the problems posed by the short duration of ZLB episodes in a Monte Carlo analysis and advocate and implement a panel VAR to overcome this issue. In addition, we address the concern that prior choice can materially affect outcomes in Bayesian structural VARs in terms of the impulse response functions (IRFs). We suggest a multivariate permanent-transitory decomposition to ameliorate this issue. Overall, our main finding is that the response of the economy to structural shocks at the ZLB is different from the standard environment with positive interest rates.

We approach the question of changing dynamics at the ZLB by estimating Bayesian VARs that we identify via a combination of long-run and sign restrictions. We provide evidence for five countries that differ in the length of their respective ZLB periods and the extent of QE. Our main conclusion from this exercise is that evidence of changes in the IRFs at the ZLB is weak and inconclusive. We argue that there are two reasons for this. First, and perhaps most importantly, ZLB samples are quite short, which limit the extent how sharp inference can be. The second reason is that a comparison between prior and posterior IRFs shows that the latter are often quite close to the former. This suggests that the concern raised by Baumeister and Hamilton (2015) appears relevant.

We demonstrate by means of a Monte Carlo analysis that the small sample issue is quite prevalent. We show the difficulty that time-varying parameter VARs (TVP-VAR) have in identifying shifts in the dynamic behavior as they happen. In contrast, when we estimate fixed-coefficients VARs on simulated samples they show more cleanly changes in the IRFs, even based on the quite short ZLB sample periods. This evidence is, however, weaker than it appears. Based on our preferred identification using long-run and sign restrictions we find that prior and posterior IRFs are often quite close.

This evidence cannot necessarily be taking at face value because of the concern raised by Baumeister and Hamilton (2015). Specifically, we show that the prior can have undue influence on outcomes, that is, the posterior, when the structural VAR is identified with sign restrictions. This makes the comparison between prior and posterior IRFs as in Arias et al. (2018) and Inoue and Kilian (2022) potentially problematic. As an alternative, we therefore re-estimate the Bayesian VARs using a Blanchard-Quah decomposition for shock identification. Notably, it is not subject to the Baumeister-Hamilton criticism of sign restrictions. The resulting evidence points toward weaker responses of prices and interest rates for ZLB sample periods. We confirm via a Monte Carlo analysis that this is partly due to small-sample issues. Overall, we find that the evidence for changing dynamics at the ZLB is not strong when estimated for individual countries.

In order to address the small-sample concern, we perform joint estimation of the VAR for all countries within a panel VAR framework. We apply the same identification schemes as in the individual country VARs and assess again the presence of small-sample issues by means of a Monte Carlo analysis. Although this solves the small-sample problem and produces sharper inference, it does not eliminate the Baumeister-Hamilton critique, which is still quite widespread as confirmed by a Monte Carlo analysis. In order to remedy this aspect, we implement a multivariate permanent-transitory decomposition of GDP shocks within the panel VAR framework.

The evidence produced by this approach is markedly sharper. Importantly, it underlies our conclusion that there are material changes in the IRFs at the ZLB. In particular, we find that the responses of GDP and unemployment exhibit more inertia than in an environment with positive interest rates. In addition, the response of prices is flatter, which offers a potential explanation for the missing inflation during the 2010s. Finally, we also find that responses of interest rates are weaker. Overall, our main conclusion is therefore that the response of the economy to structural shocks at the ZLB is, in fact, different from more normal times.

The theoretical literature established fairly early, much before the widespread ZLB episodes in the 2010s that there could be qualitatively important changes to the dynamics when the economy is at its ZLB. The simple reason is that the ZLB imposes a non-linear constraint in environments that are (close to) linear so that the dynamics change by necessity when compared to a fully linear setting away from the ZLB. The literature started with the seminal contribution by Benhabib et al. (2001), who showed in a standard monetary model that there is an equilibrium at the ZLB that features different dynamics from what is typically assumed in standard non-ZLB equilibria. Numerous papers build on this insight, develop analytical tools for studying such dynamics and investigate policy responses such as forward guidance, for instance, Benhabib et al. (2002), Eggertsson and Woodford (2003) or Werning (2011) among many others.

During the GFC and the recovery, issues surrounding the ZLB received new attention for obvious reasons. Christiano et al. (2011) show that the government expenditure multiplier at the ZLB can be exceedingly large and thereby a powerful policy instrument. They also shift the focus towards the effects of a variety of shocks on several economic variables in a simulated dynamic stochastic equilibrium (DSGE) model that reveals sizeable dynamic changes. Their findings were confirmed and refined by Fernandez-Villaverde et al. (2015) who emphasize the fundamental non-linear aspect of the ZLB, which materially affects dynamics. Boneva et al. (2016) raise come concern about their conclusion, emphasizing the improtance of the specific solution method for the non-linear system. They conclude that quantitatively the differences between ZLB and non-ZLB influenced dynamics are minor.

Empirical studies of ZLB dynamics are more sparse. An important contribution is Aruoba et al. (2017) who estimate a New Keynesian DSGE model with two steady states as in Benhabib et al. (2001) using likelihood-based methods. They generally confirm the prevalence of different dynamics near the ZLB but highlight the dependence of this finding on model specification. Nevertheless, imposing the ZLB constraint has become a central feature in medium- to large-scale DSGE models that are used for policy analysis at central banks, such as Del Negro et al. (2015).

The closest precursor to our paper is Debortoli et al. (2020). They provide evidence from a time-varying parameter VAR (TVP-VAR) for the US that the IRFs at the ZLB are not significantly different from the IRFs for the period with positive interest rates. They interpret their findings as compatible with the notion that the Federal Reserve's unconventional monetary policies were effective in compensating for the fact that the federal funds rate was held in a narrow range slightly above the ZLB during and for an extended period after the Great Recession.

However, there are some doubts about the strength of the conclusion. For one, the ZLB samples are quite short. In addition, the estimates produced by a TVP-VAR are two-sided and therefore mix the future with the past. This entails smoothing across potentially different monetary policy regimes, which is arguably not appropriate for the question at hand. Moreover, the use of a TVP-VAR when estimated over the full sample does not entirely ameliorate the small-sample problem as we demonstrate in a Monte Carlo exercise.

While TVP-VARs are in principle capable of detecting structural changes even in short samples, they do not entirely avoid this issue, especially when the underlying changes are not large in nature.[1]

The paper is organized as follows. In section 2, we present our model specification, discuss the data, and present the different structural identification methods we use throughout the paper. In the following section we conduct a Monte Carlo analysis to study the ability of TVP-VARs and fixed-coefficient VARs to detect changes in IRFs to structural shocks across regimes. We highlight the problems of TVP-VARs along this dimension. In this section we estimate Bayesian fixed-coefficients VARs by sub-sample, which are identified via a combination of long-run and sign restrictions. We also explore to what extent the differences in IRFs across sub-samples are due to the concern raised by Baumeister and Hamilton (2015) about identification via sign restrictions. In section 4 we attempt to alleviate the small sample concerns by estimating a panel VAR for all countries in the combined sample. Section 5 concludes.

2 Methodology

We study the behavior of the economy off and at the ZLB empirically in a VAR framework. In the following, we describe our VAR specification and discuss our Bayesian approach to inference. We then present several identification methods for isolating monetary policy shocks, the responses to which are our key object of interest. Finally, we present and discuss the data for several countries that had ZLB episodes during the 2010s.

2.1 VAR Specification and Inference

Our baseline specification is the fixed-coefficients VAR:

$$Y_t = B_0 + B_1 Y_{t-1} + ... + B_p Y_{t-p} + u_t, \tag{1}$$

where the notation is standard. Y_t is an $N \times 1$ vector of variables and $u_t \sim N(0, \Sigma)$ collects the reduced-form errors. The B_j are the associated coefficient matrices. We set the lag order p to two for quarterly data and to six for monthly data.

We estimate the VARs as described in Giannone et al. (2015). By defining $\beta \equiv$

[1] Lubik et al. (2016) show in a Monte Carlo study that TVP-VARs have difficulty detecting substantial structural changes even in longer samples, albeit in a specific example of the labor market.

$vec([B_0, B_1, ..., B_p]')$ and $x_t \equiv [1, Y'_{t-1}, ..., Y'_{t-p}]'$, equation (1) can be rewritten as:

$$Y_t = X_t \beta + u_t, \tag{2}$$

where $X_t \equiv I_N \otimes x'_t$. We assume that the prior distribution of the VAR coefficients belongs to the Normal-Wishart family, namely:

$$\Sigma \sim IW(\Psi; d) \tag{3}$$

$$\beta|\Sigma \sim N(b; \Sigma \otimes \Omega), \tag{4}$$

where the elements of Ψ, d, b and Ω are functions of a lower-dimensional vector of hyperparameters.

The degree of freedom of the Inverse-Wishart distribution IW is set to $d = N + 2$, which is the minimum value that guarantees the existence of the prior mean of Σ. We assume that Ψ is a diagonal matrix with the $N \times 1$ vector of hyperparameters ψ on the main diagonal. The conditional Gaussian prior for β is of the Minnesota type. In contrast with the standard specification of random-walk priors, we impose stationarity on the VAR. Consequently, we assume:

$$E\left[(B_s)_{ij}|\Sigma\right] = \begin{cases} \mu \text{ if } i = j \text{ and } s = 1 \\ 0 \text{ otherwise} \end{cases} \tag{5}$$

We set $\mu = 0.25$ for real GDP growth and inflation, which captures the low serial correlation of GDP growth and also of inflation during the Great Moderation period. For the other series we set $\mu = 0.9$. This reflects a prior view that interest rates, the unemployment rate, and hours worked per capita are stationary, yet highly persistent. We set the prior for the second moment as:

$$Cov\left[(B_s)_{ij}, (B_r)_{hm}|\Sigma\right] = \begin{cases} \lambda^2 \frac{1}{s^2} \frac{\Sigma_{ih}}{\psi_j/(d-N-1)} \text{ if } m = j \text{ and } r = s \\ 0 \text{ otherwise} \end{cases} \tag{6}$$

The hyperparameter λ controls the scale of the variances and covariances, thus determining the prior's overall tightness. We set the hyperpriors for λ and ψ as in Giannone et al. (2015).

We estimate the VAR as discussed there, using the MATLAB codes available at Giorgio Primiceri's web page. The only difference to their approach is that we impose stationarity on the VAR. In the MCMC step of the estimation, we move to iteration $i + 1$ if and only if the draw for the VAR parameters associated with iteration i is stationary. Otherwise, we redraw the parameters for iteration i.

2.2 Identification

We consider three alternative identification strategies. Although these are standard in the literature we discuss their implementation in what follows. First, we consider joint long-run and sign restrictions as in Debortoli et al. (2020). Next, we implement a pure sign-restrictions approach using the methodology of Rubio-Ramirez et al. (2010). Finally, we consider a permanent-transitory decomposition to address the concern raised by Baumeister and Hamilton (2015).

Long-run and sign restrictions. We implement joint identification of a permanent shock to real per-capita GDP and of four transitory shocks as in Debortoli et al. (2020) by combining zero long-run restrictions and short-run sign restrictions on the direction of the impulse response functions. Let the structural VAR(p) model be given by $Y_t = B_0 + B_1 Y_{t-1} + ... + B_p Y_{t-p} + A_0 \epsilon_t$, where A_0 is the impact matrix of the structural shocks at $t = 0$, and $\epsilon_t = A_0^{-1} u_t \equiv [\epsilon_t^{PE}, \epsilon_t^{TE}, \epsilon_t^{TA}, \epsilon_t^{MA}, \epsilon_t^{MO}]'$ are the structural shocks. We assume that the shocks have unit variance and are orthogonal to each other.

We can disentangle the shock ϵ_t^{PE} from the other four structural disturbances because it is the only one that is allowed to have a permanent impact on real GDP. The other four disturbances, ϵ_t^{TE}, ϵ_t^{TA}, ϵ_t^{MA}, and ϵ_t^{MO} are, respectively, 'technology', 'taste', 'markup' and 'monetary policy' shocks in the nomenclature of Canova and Paustian (2011). They are identified and separated from each other by imposing sign restrictions on impact. We report the set of restrictions in Table 1, where '+' indicates 'greater than, or equal to zero', and '−' means 'smaller than, or equal to zero'. These restrictions are the same as the 'robust sign restrictions' reported by Canova and Paustian (2011) for their benchmark DSGE model that features sticky prices, sticky wages, and several frictions standard in the literature.[2]

We choose to impose the sign restrictions only on impact. As Canova and Paustian (2011) argue, impact restrictions are in general robust. They hold for the vast majority of sub-classes within a specific class of DSGE models and for the vast majority of plausible parameter configurations. Restrictions at longer horizons, on the other hand, are generally not robust across sub-classes of models. Moreover, implementing multi-horizon restrictions imposes a computational burden to the extent that only a small number of draws from the identification matrix satisfy these restrictions. Since we are limited by the small sample size of the ZLB episodes we thus prefer to remain flexible in that direction.

[2]The differences to the Canova-Paustian restrictions are that their DSGE model is built around the output gap whereas our VARs feature real GDP growth in per-capita terms and, second, that we consider the unemployment rate instead of hours per capita.

Table 1. Impact Sign Restrictions				
	Shock:			
Variable:	ϵ_t^{TE}	ϵ_t^{TA}	ϵ_t^{MA}	ϵ_t^{MO}
Real GDP growth per capita	+	+	−	−
Inflation	−	+	+	−
Unemployment rate	+	−	+	+
Monetary policy rate / shadow rate	−	+	+	+

For each draw from the posterior distribution of the VAR's reduced-form parameters we compute the structural impact matrix A_0 of the shocks using the methodology proposed by Arias et al. (2018). The methodology is described in detail in the online appendix. We set the number of random rotation matrices to 100, and the number of Gibbs-sampling iterations to 10.

Only sign restrictions. Sign restrictions pertain in principle to the infinite long-run. Consequently, the reliability of an identification strategy based on long-run restrictions is a matter of concern, see Faust and Leeper (1997). When we work with the 4-variables VAR we therefore identify the four shocks listed in Table 1 based on a pure sign restrictions approach that we implement as discussed in Rubio-Ramirez et al. (2010). The methodology is described in detail in the online appendix. We set the number of random rotation matrices to 100. Again, for the reasons discussed previously we impose the sign restrictions only on impact.

Permanent-transitory decomposition. One concern about identification via long-run and sign restrictions is that the posterior distribution of the shocks' impulse responses largely reflects the priors implicit in the sign restrictions imposed by the researcher. For instance, Baumeister and Hamilton (2015) argue that uniform priors over the random rotation angles that are the key inputs in the standard algorithms for imposing such restrictions do not map into corresponding uniform priors for the IRFs. Because of the highly non-linear nature of the problem, they imply in general informative priors for the IRFs, sometimes quite remarkably so. In our results below, a comparison of prior and posterior IRFs as in Inoue and Kilian (2022) and Arias et al. (2022) reveals that this concern is, in several instances, materially relevant. This is especially the case for the comparatively short ZLB samples.

We therefore consider a third identification strategy, which is not vulnerable to the Baumeister and Hamilton (2015) criticism and relies on point identification. As in the first identification approach we identify a permanent GDP shock as the only disturbance with a permanent impact on real per-capita GDP. Conditional on the identification of this shock, the remaining transitory GDP disturbances are then rotated as in Uhlig (2005) in order to isolate the most important transitory shock for GDP at the business-cycle frequencies; that is, the shock that explains the largest fraction of the residual variance of GDP within this frequency band. In line with the literature we take the business-cycle frequency band to encompass fluctuations with periods between 2 and 8 years. We find that this identification scheme essentially performs a multivariate permanent-transitory decomposition of GDP shocks since the fraction of forecast error variance of GDP jointly explained by the two identified shocks is very close to 100 percent for all countries, samples, and horizons.[3]

2.3 Data

In the empirical analysis we consider a set of countries that had extended ZLB episodes during and after the global financial crisis. These include Denmark, the Euro area, Japan, Switzerland, the UK, and the US. For all countries and sample periods we study two alternative VARs that are specified for four and five variables, respectively. The 4-variables system features the log-differences of real per-capita GDP and the GDP deflator, a short-term nominal interest rate, where we use the central bank's monetary policy rate for the pre-ZLB period and a 'shadow rate' for the ZLB period, and finally the unemployment rate. The 5-variables system additionally contains a long-term nominal interest rate. The data and their sources are described in detail in the online appendix.

The data used in the estimation are at quarterly frequency. Since the ZLB samples are quite short, less than 10 years in most cases, concerns about inference in small samples is consequently quite valid. We address these concerns by pooling the country data into a panel VAR in section 4. In the case of the US, the relevant series are also available at the monthly frequency. In order to obtain more precise estimates we therefore work with monthly US data and report these as the baseline. The corresponding evidence based on quarterly data is qualitatively the same and is available in the online appendix.

Figure 1 shows the respective central bank's monetary policy rate together with the shadow rate.[4] The graphs reveal the substantial decline of policy rates long before the onset of the GFC, where an underlying factor is certainly the overall decline in real rates of interest. When the financial crisis hit, central banks rapidly dropped rates to zero with the exception

[3] We report respective evidence in the online appendix. A partial exception is Switzerland for which this fraction is between 90 and 100 percent in the pre-ZLB sample.

[4] The shadow rate for Denmark is not depicted because it is an internal estimate of the Danish central bank and has kindly provided to us on a confidential basis.

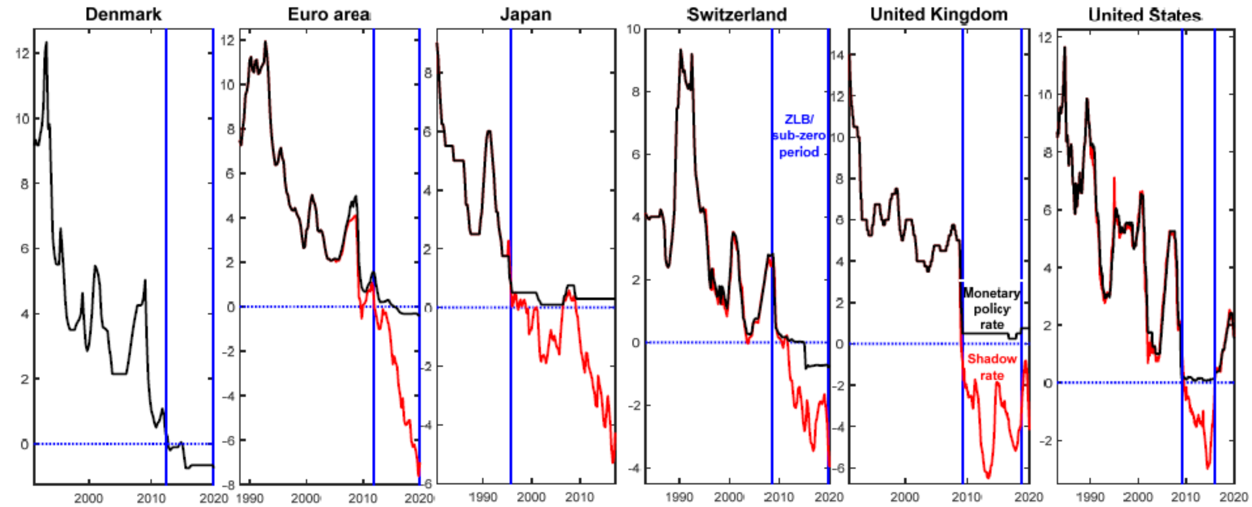

Figure 1: Monetary policy rates and shadow rates

of Japan, which had been at its ZLB since the mid-1990s. It is also notable that Denmark and Switzerland tested out the limits of their interest rate policy by imposing negative rates for an extended period. The figure also shows the shadow rates, which in several cases have gone into substantially negative territory, for instance to almost - 6 percent in the Euro Area.[5]

Table 1 reports the sample periods together with the average values of the policy rate and the shadow rate. During the ZLB period in either the US, the UK or Japan the policy rate has never become negative. It has oscillated, respectively, between 0.07 and 0.24, 0.25 and 0.5, and 0.10 and 0.75 percent. In the other three countries it fell below zero, reaching a minimum of -0.75 percent in Denmark, -0.85 percent in Switzerland, and -0.40 percent in the Euro area. Although the peak value of the policy rate during the ZLB period has been above zero in the UK and Japan, we choose to include both countries in the analysis because of the sharp contrast between the average values over the ZLB/sub-zero sample (0.47 and 0.34 per cent, respectively) and the corresponding values for the pre-ZLB period (7.57 and 4.19, respectively). While not ideal, the evidence provided by either country should still be informative in assessing whether the transmission mechanism of structural shocks at the ZLB is different.

[5]The gap between actual and shadow rates signal the extent of accommodation that central banks provided during and after the recession. It also hints at the severity of the constraint that the ZLB poses in pursuing policy. From a different perspective, QE is seen as providing a substitute for interest rate policy to the tune of $100 billion in asset purchases being the equivalent of a 25 basis points drop in the policy rate. See Krishnamurthy and Vissing-Jorgensen (2011) for extended discussion on this point.

Table 2. Sample periods and average monetary policy and shadow rates			
Country	Sample (quarters)	Average policy rate	Average shadow rate
	Pre-ZLB		
Denmark	1983Q2-2010Q2 (109)	5.94	–
Euro area	1980Q2-2012Q2 (129)	6.58	–
Japan	1980Q3-1995Q3 (61)	4.19	–
Switzerland	1983Q2-2008Q4 (103)	3.40	–
United Kingdom	Feb. 1983-Feb. 2009 (104)	7.57	–
United States	Feb. 1983-Nov. 2008 (103)	5.36	–
	ZLB/sub-zero		
Denmark	2010Q3-2019Q4 (38)	-0.23	–
Euro area	2012Q3-2019Q4 (30)	-0.10	-3.55
Japan	1995Q4-2017Q1 (86)	0.34	-1.32
Switzerland	2009Q1-2019Q4 (44)	-0.27	-1.65
United Kingdom	Mar. 2009-Jul. 2018 (43)	0.47	-3.74
United States	Dec. 2008-Dec. 2015 (28)	0.13	-1.21

3 Monte Carlo Evidence

We now turn to the Monte Carlo evidence on the ability of the fixed-coefficients Bayesian VAR estimator of Giannone et al. (2015) and of the TVP-VAR estimator of Del Negro and Primiceri (2015) used by Debortoli et al. (2020) to allow the corresponding SVARs to identify changes in the shocks' IRFs across regimes. We first describe the set-up of our Monte-Carlo procedure and the data-generating process we use for the question at hand. We then report results from the estimation on simulated data and then estimate VARs on the set of countries discussed above.

3.1 Monte Carlo Framework

Our data-generating process (DGP) is a standard New Keynesian model with both forward- and backward-looking components in the Euler-equation and the Phillips curve. It is based on the model estimated in Benati (2008). The main difference in our set up is that we assume that the log of real GDP has a unit root component that evolves as a random-walk with drift as in Watson (1986). The model contains three variables, real GDP, inflation, and a short-term interest rate. The economy is driven by three shocks: a monetary shock $\epsilon_{R,t}$, a permanent shocks to real GDP v_t, and a transitory supply shock u_t. We discuss the model and the Bayesian estimation methodology in the online appendix. The left-hand panels of 2 show the IRFs of the DSGE model we use as data generation process (DGP) under two alternative parameterizations.

For our Monte Carlo simulations we draw 10,000 times from the posterior distribution of the structural model parameters. Each draw represents a statistical 'model'. We select the two models, i.e., draws i and j, for which the following measure of distance between the two set of IRFs is maximized:

$$\sum_{v=1}^{N}\sum_{s=1}^{N}\sum_{h=0}^{H}\left|IRF^{i}_{v,s,h} - IRF^{j}_{v,s,h}\right|, \tag{7}$$

where v, s, and h index, respectively, the variable, the shock and the horizon. We set H equal to 10 years, which is roughly the average duration of the ZLB episodes of the countries in our sample. In most cases there is a sharp difference between the IRFs under the two parameterizations, which suggests that this is an appropriate DGP for exploring the ability of a VAR to detect differences in the IRFs across sub-periods.

We conduct the following Monte Carlo experiment. We generate 1000 artificial samples of length $T = T_1 + T_2$ for inflation, the short rate, and real GDP from the DGP. The first sub-sample of length T_1 is generated under parameterization 1, whereas the second with length T_2 is generated under parameterization 2. We adapt the size of the simulated sample to its empirical counterpart under investigation. Our frame of reference is Debortoli et al. (2020), who use the TVP-VAR specification of Del Negro and Primiceri (2015). Their overall estimation sample runs from 1953Q2-2015Q4, so that we use $T_1 = 222$ and $T_2 = 28$. We also estimate the fixed-coefficients specification of Giannone et al. (2015) and set the sub-sample lengths corresponding to the minimum lengths (in quarters) of the pre-ZLB and ZLB samples reported in Table 2. This implies $T_1 = 61$ and $T_2 = 28$. This choice is designed to assess the ability of fixed-coefficients SVARs estimated by sub-sample to detect differences in the IRFs within the specific empirical settings used in the present work under a worst-case scenario.[6]

We jointly identify the permanent output shock and the two transitory shocks by combining long-run and sign restrictions as described in 2.2. In case of the fixed-coefficients

[6]In our estimation of the fixed-coefficient VARs we follow Giannone et al. (2015) as described in 2.1. Similarly, the implementation of the TVP-VAR is as in Debortoli et al. (2020) with the difference that we impose stationarity on a quarter-by-quarter basis as in, e.g., Cogley and Sargent (2005). We are grateful to Giorgio Primceri for providing us with the Matlab codes.

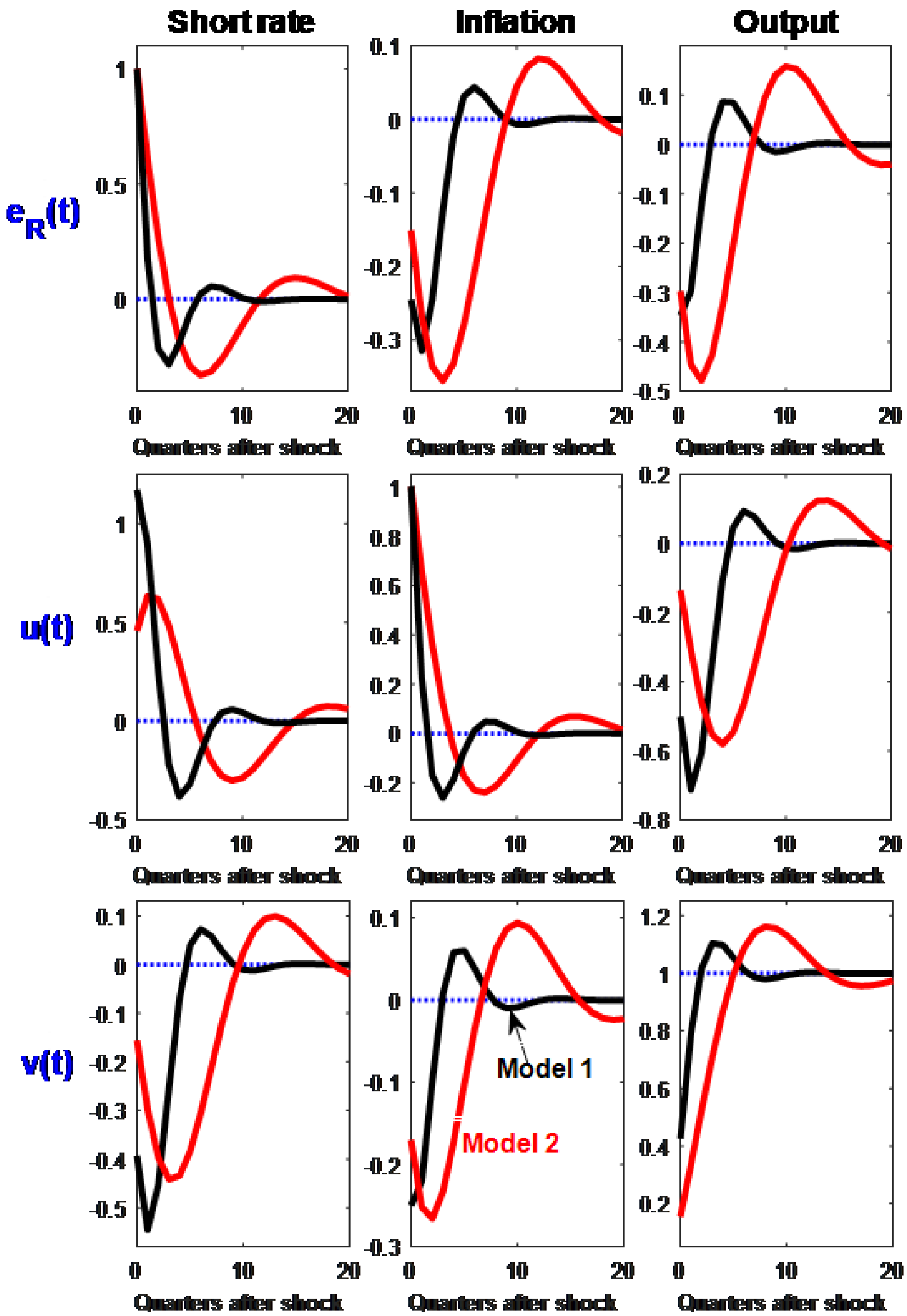

Figure 2: Impulse responses of the data-generating process

VAR this is done for each draw from the posterior distribution reduced-form VAR estimates based on the two sub-samples. The successful draws, that is, those for which the sign restrictions are satisfied, are retained. We compute IRFs to the three shocks for each them and then store the median IRFs and the 16th and 84th percentiles of their respective posterior distributions. In case of the TVP-VAR estimation of the simulated data we replicate the empirical setup of Debortoli et al. (2020). As before, we identify the three shocks, compute the IRFs and extract their medians and 16th and 84th percentiles. This is done for each draw from the posterior and each quarter for both the second sub-sample of length T_2 and the corresponding sub-sample of length T_2 ending at $T_2 - 1$. Just as in their paper, the ZLB sub-sample for which we compute the IRFs is of length T_2.

The results of the Monte Carlo analysis are shown in Figure 3. We report the averages across simulations of the median and the 16th and 84th percentiles of the posterior distributions of the IRFs. We find that the TVP-VAR does not systematically detect differences in the IRFs across sub-periods. The Monte Carlo averages of the median estimated IRFs and the 16-84 percent credible sets based on the two parameterizations are consistently very close to each other for nearly all shocks and all variables. The immediate interpretation of this finding is that the estimates produced by TVP-VARs are two-sided and therefore automatically tend to smooth changes across regimes. If both sub-samples were sufficiently long, TVP-VARs would likely detect differences in the IRFs across parameterizations.[7] In our case, the ZLB sample is quite short at 28 quarters, however. It is perhaps not surprising that the TVP-VARs do not seem to detect differences in the IRFs because the estimates for the second sub-sample end up being dominated by the first one.

Evidence for the fixed-coefficients VAR estimated by sub-sample is just slightly better. In particular, the main difference compared to the TVP-VAR are the responses of the short rate and output to u_t and the responses of the short rate and inflation to v_t. In all other instances the difference between the estimated IRFs for the two parameterizations is small or almost negligible.

[7] Lubik et al. (2016) show, however, via simulation for a parameter shift in a labor market search-and-matching model that sample size would have to be considerably longer than the typical macroeconomic post-World War II sample.

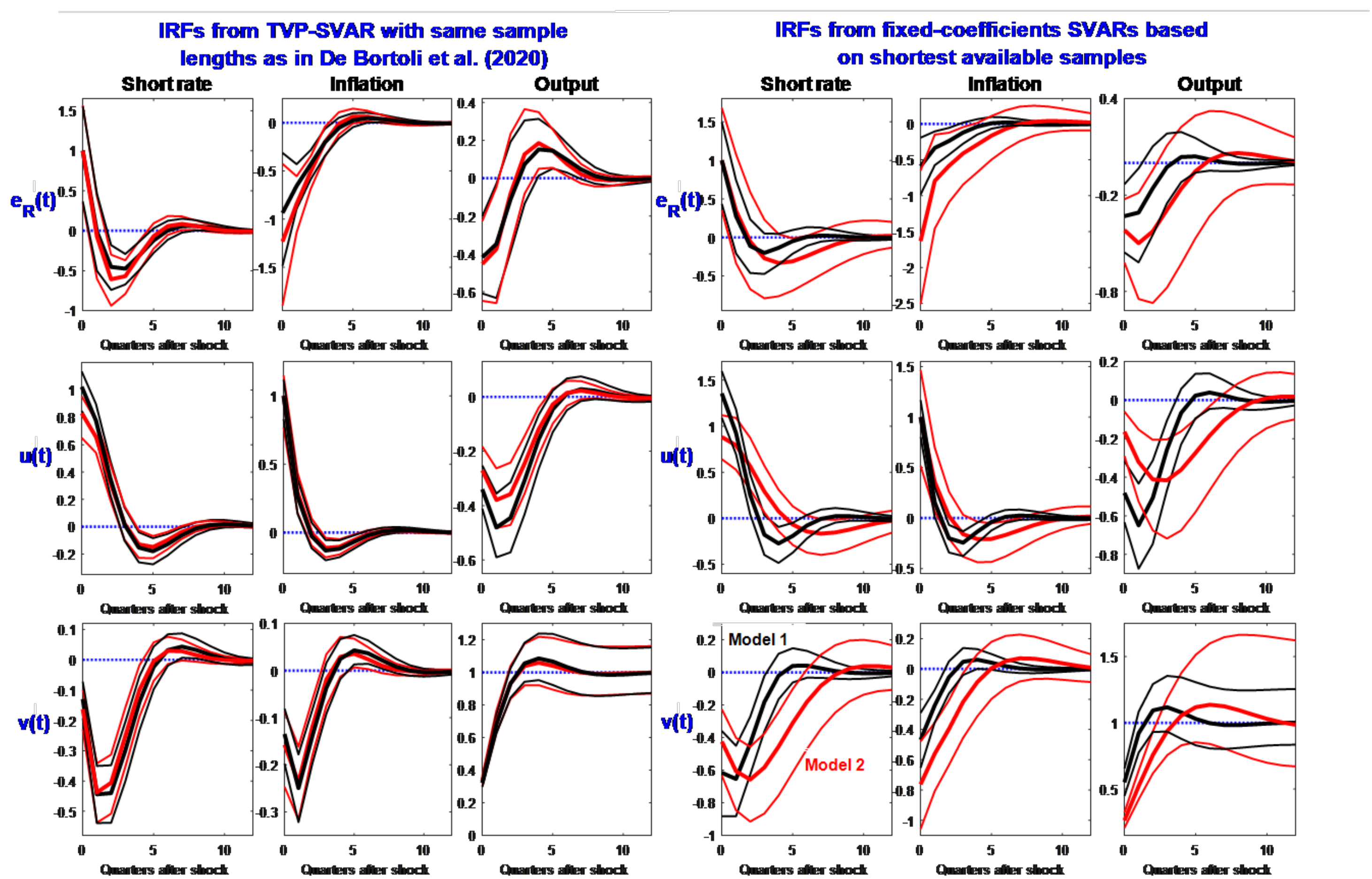

Figure 3: Monte Carlo simulation: Estimated IRFs based on TVP-VAR and fixed-coefficient VAR estimated by sub-sample.

Overall, the evidence in Figure 3 suggests that VARs as employed by Debortoli et al. (2020) have a limited statistical ability to detect changes in IRFs at the ZLB compared to the standard environment with positive interest rates, even when such changes are sizeable. This issue is compounded by the short sample lengths that characterize ZLB episodes. We find that this is especially the case for time-varying parameter VARs, whose estimates are two-sided and therefore tend to smooth changes across regimes. This leads us to conclude that within this specific context the evidence for structural changes in the dynamic behavior at the ZLB is likely unreliable.

In principle, one avenue that can be pursued is to perform pooled estimation of all countries jointly. Assuming that the responses of individual countries to the structural shocks are sufficiently similar, this allows to effectively increase the sample size and therefore improve the precision of the estimates. When pooling all of the six countries together, the sum of their individual sample sizes is 290 quarters. Figure 4 shows evidence from the same exercise performed in Figure 3 for fixed-coefficients VARs for $T_1 = T_2 = 290$. The evidence seems straightforward: with such a large sample size, the fixed-coefficient specification can in principle detect differences in the IRFs across periods. In particular, for either u_t or v_t differences are notably detected for all variables, whereas for $\epsilon_{R,t}$ they are still detected, but less clearly.

3.2 Evidence from Individual Country Estimates

For completeness, we report the estimates for individual countries, using different specifications of the VAR and different identification approaches. We discuss the main insights from this exercise here, but the figures are relegated to the online appendix. The figures in appendix A show evidence for individual countries from long-run and sign restrictions based on Arias et al. (2018). The first half of the graphs depicts the estimated impulse-response functions to individual shocks, while the remainder shows the prior and posterior IRFs to individual shocks for the pre-ZLB and ZLB samples. The other figures in that section report the same evidence based on the pure sign restrictions approach of Rubio-Ramirez et al. (2010).

The main message emerging from the two groups of figures is that evidence of changes in the IRFs at the ZLB is weak and inconclusive. There are two reasons for this. First, ZLB samples are quite short. Second, a comparison between prior and posterior IRFs shows that the latter are often quite close to the former, so that the problem highlighted by Baumeister and Hamilton (2015) appears to be relevant. In order to address the first problem, in the next section we perform joint estimation of the VAR for all countries within a panel VAR

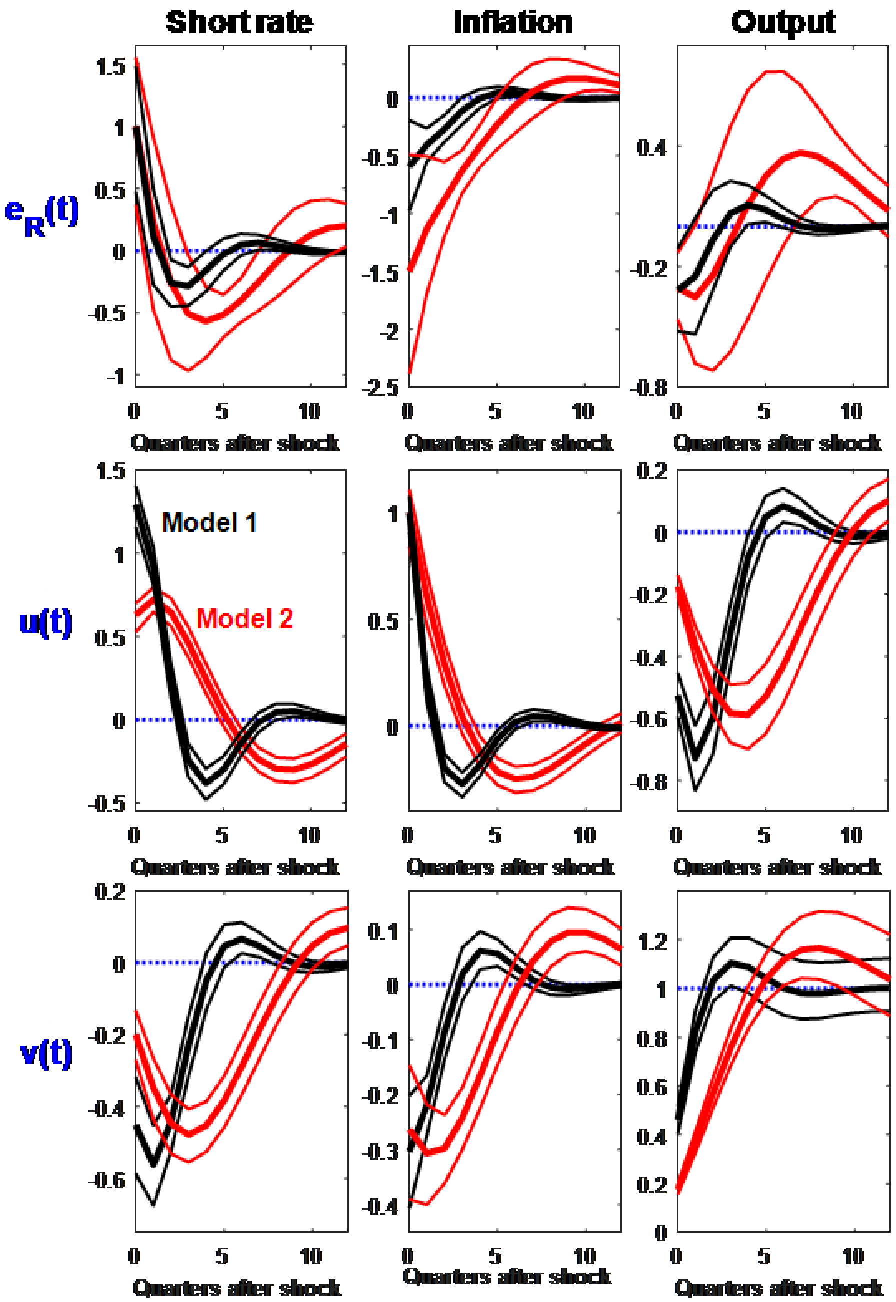

Figure 4: Monte Carlo simulation: Estimated IRFs based on fixed-coefficient VAR. Pooled Sample

framework. Although this solves the small-sample problem and produces sharper inference, it does not eliminate the Baumeister-Hamilton problem, which is still quite widespread. This will motivate our next step, i.e. implementing a multivariate permanent-transitory decomposition of GDP shocks within the panel VAR framework.

4 Panel VAR Evidence

Any empirical study of the dynamic behavior of macroeconomic aggregates during the ZLB period has to contend with the fact that across countries these episodes were of short duration, especially when compared with the generally short length of macroeconomic time series. Debortoli et al. (2020) attempt to circumvent this aspect by estimating a TVP-VAR, which is designed to capture structural changes in the underlying time series, while using the information from the full sample and not just the ZLB sub-sample. Our results in the previous section indicate, however, that this small-sample issue also affects their inference. Moreover, our own sub-sample estimates with three different identification assumptions reveal weak inference, so that any conclusions drawn are likely to be unreliable.

In this section, we address the small-sample issue by implementing joint estimation over all countries in the sample. Since individual countries' sample periods are quite short we pooled the countries together and estimated a panel VAR. Naturally, the countries in our sample are quite different, especially in terms of size, economic structure and international openness. Nevertheless, we would expect that shocks at the ZLB would induce broadly similar behavior for the key aggregates despite these differences. In order to sharpen inference, we allow a limited degree of heterogeneity in the specification of the panel VAR.

Specifically, we estimate a fixed-coefficients VAR as before:

$$Y_t = B_0 + B_1 Y_{t-1} + ... + B_p Y_{t-p} + u_t, \tag{8}$$

where the Y_t vector now contains the individual country series stacked on top of each other. The B_i matrices are adapted conformably, as is the vector u_t of reduced-form errors. We allow for the fact that different countries have different unconditional sample means of real GDP growth and other variables so that the intercept vector B_0 is different for each country. The rest of the panel specification is identical for each country, namely the coefficient matrices B_1, B_2, ..., B_p as is the covariance matrix of of shocks.

These are strong restrictions, which are arguably not exactly satisfied. The trade-off involved is that a completely unrestricted panel specification likely results in no precision gain, while our specification is potentially misspecified. However, our focus is on the information gain from joint estimation of the pooled countries. As our results below indicate is that the qualitative picture that emerges from the restricted panel specification is quite similar to the individual countries, while at the same time alleviating the small-sample issues.

Figure 5 shows the impulse responses based on long-run and sign restrictions for the ZLB period and for the previous period. There are notable differences in several cases, for instance, the impact of a transitory taste shock on the price level. However, in many cases the differences are minor. At any rate, in order to properly assess these results we need to take into account that samples for the pre-ZLB period are much longer than for the second sub-sample. Moreover, the Baumeister-Hamilton critique that we identified as potentially relevant for the question at hand is likely at play here, too. That is, we need to assess whether the posterior IRFs are sufficiently different from the prior IRFs.

We first address the issue of different sample sizes. Figure 6 shows the posterior IRFs for the ZLB period (in red) together with results from a Monte-Carlo simulation. We depict the means of the percentiles of the IRFs computed under the null hypothesis that the data-generating process for the ZLB period is the same as for the previous period. We implement the Monte-Carlo procedure as follows.

Joint estimation over the pre-ZLB period produces a posterior distribution for the SVAR in equation (8). where $u_t = A_0 e_t$. We take N draws of the VAR coefficient matrices and the SVAR's impact matrix of the structural shocks A_0. For each draw $j = 1, 2, ..., N$, we simulated the SVAR (8) for a sample length equal to the overall sample in the ZLB period, i.e., the sum of the individual countries' sample lengths. Innovations are drawn from the standard normal distribution for e_t. Based on each simulated sample we estimated and identified the SVAR using the same long-run and sign restrictions as for the actual data. This process produces a set of N percentiles for the IRFs. We compute the mean of the percentiles of the IRFs across all of the N Monte Carlo simulations, which we show in Figure 6 (in black).

By construction, the difference between these IRFs and the corresponding black IRFs in Figure 5 stems from the fact that those in Figure 6 are based on a shorter sample period, equal to the actual overall length of the ZLB episode. The difference between the two sets of black IRFs in Figures 5 and 6 is quite small, however. At the same time, they paint the same picture qualitatively. We therefore conclude that the difference between the IRFs for the two periods in Figure 5 does likely not originate from the fact that the ZLB periods are shorter.

We address the concerns raised by Baumeister and Hamilton (2015) in Figure 7. It depicts the same IRFs (in red) for the ZLB period, that is, the posterior IRFs obtained from the actual data. In addition, the graph shows the prior IRFs (in black). These are computed

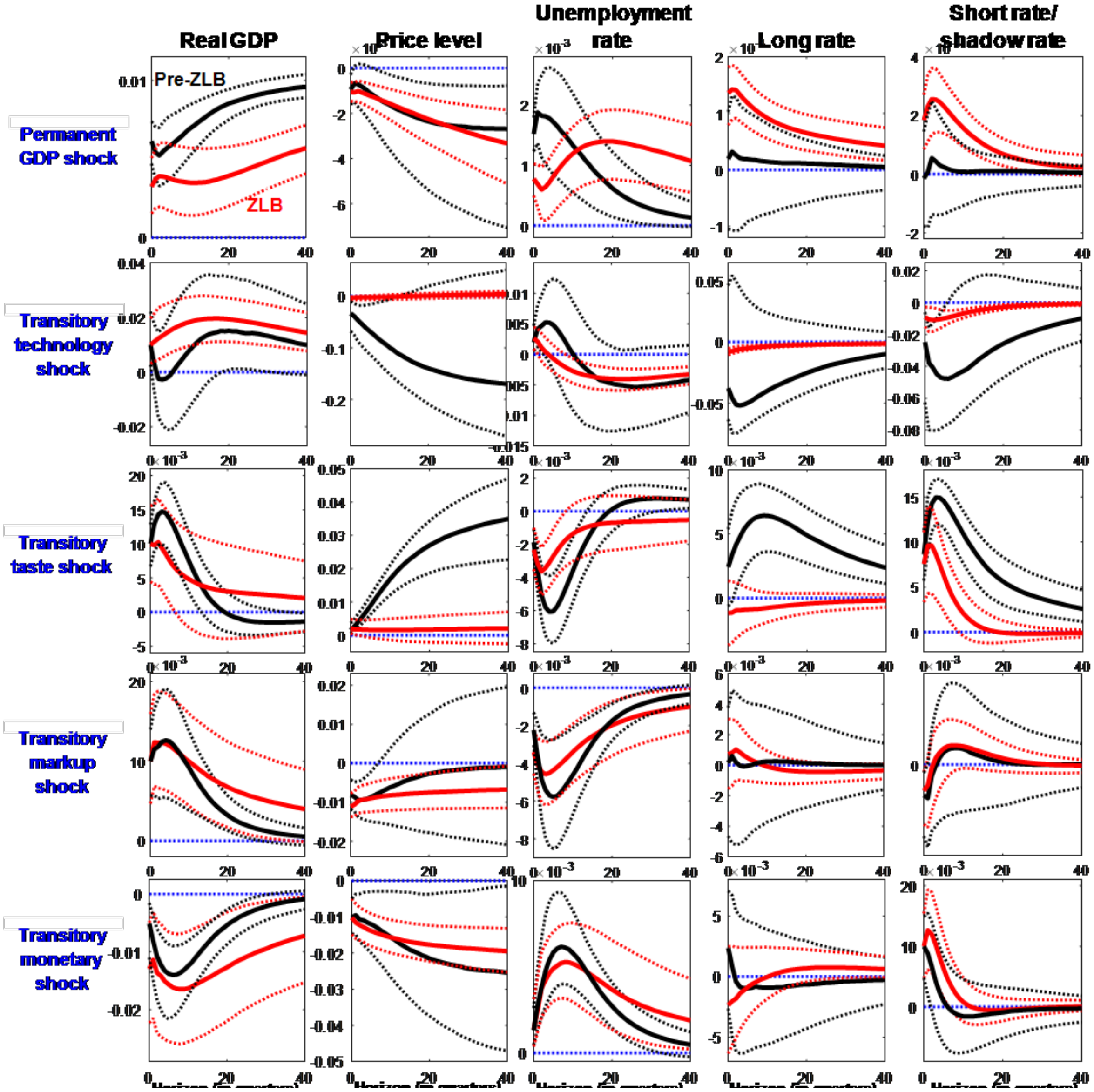

Figure 5: Evidence from long-run and sign restrictions based on joint estimation: median and 16th and 84th percentiles of the posterior distributions of the estimated impulse-response functions to the structural shocks

by first drawing from the prior distribution of the reduced-form coefficients of the VAR in equation (8) based on the estimator of Giannone et al. (2015). We then use the Arias et al. (2018) methodology to impose the sign restrictions.[8]

The key insight from Figure 7 is that in several cases the Baumeister-Hamilton critique

[8]Arguably, this is the correct way of computing the prior IRFs as discussed by Inoue and Kilian (2022). It differs from the procedure followed by Watson (2020), who applied the second step to the maximum likelihood point estimate of the VAR, rather than to each of the draws from a prior distribution of the VAR's coefficients.

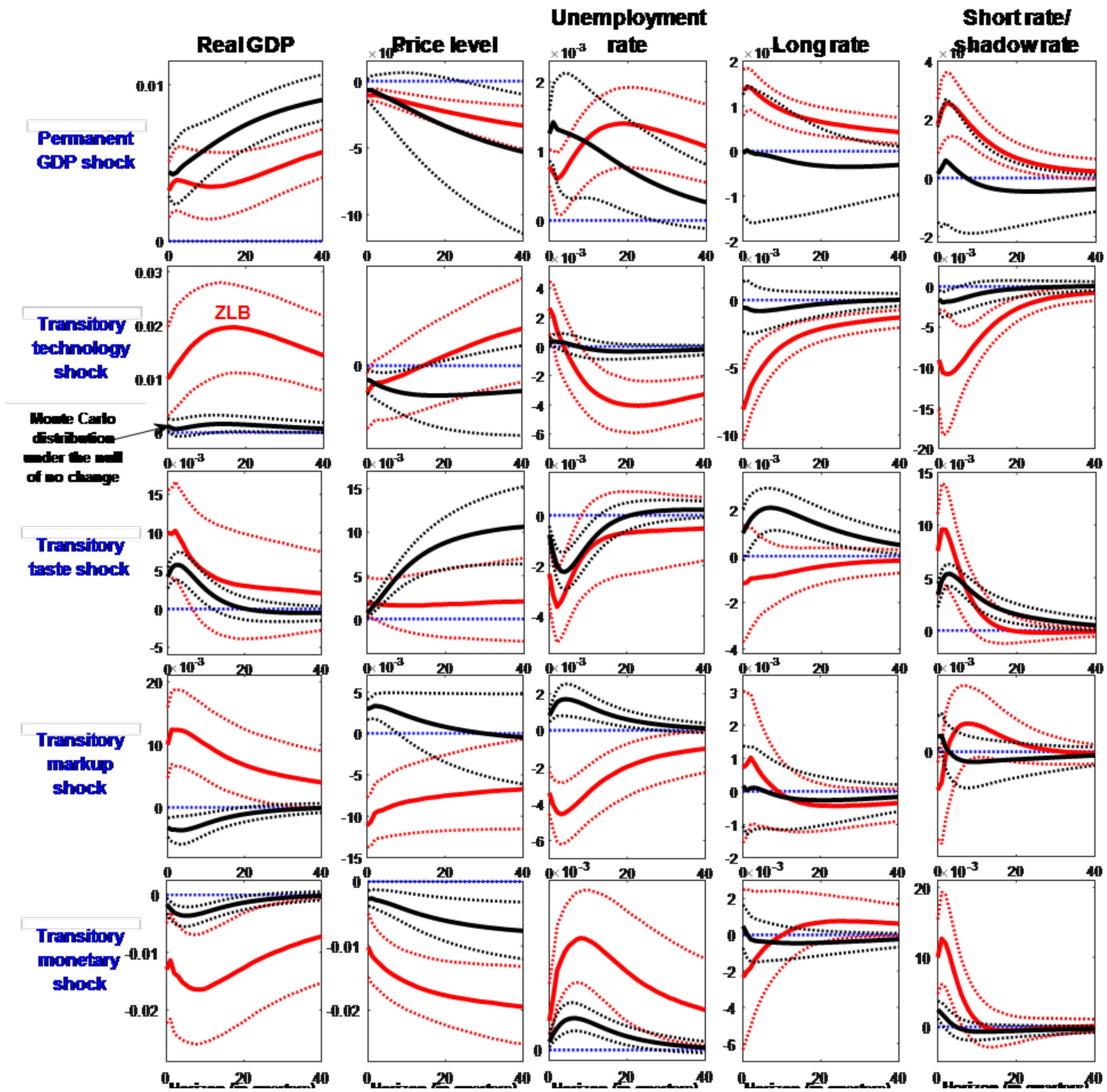

Figure 6: Evidence from long-run and sign restrictions based on joint estimation for the ZLB/sub-zero period: median and 16th and 84th percentiles of the posterior distributions of the estimated impulse response functions to the structural shocks, and comparison with the Monte Carlo distribution under the null of no change.

appears to be relevant, for instance, the effect of monetary and markup shocks on short rates and shadow rates; or the impact of taste shocks on GDP and the price level. Only in a few cases are prior and posterior IRFs clearly different. Contrary to the examples provided by Inoue and Kilian (2022) this shows that the Baumeister-Hamilton criticism seems to bite in this specific example.

This observation thus motivates an estimation strategy based on point identification to avoid vulnerability to this aspect of Bayesian estimation. Specifically, we pursue a multi-

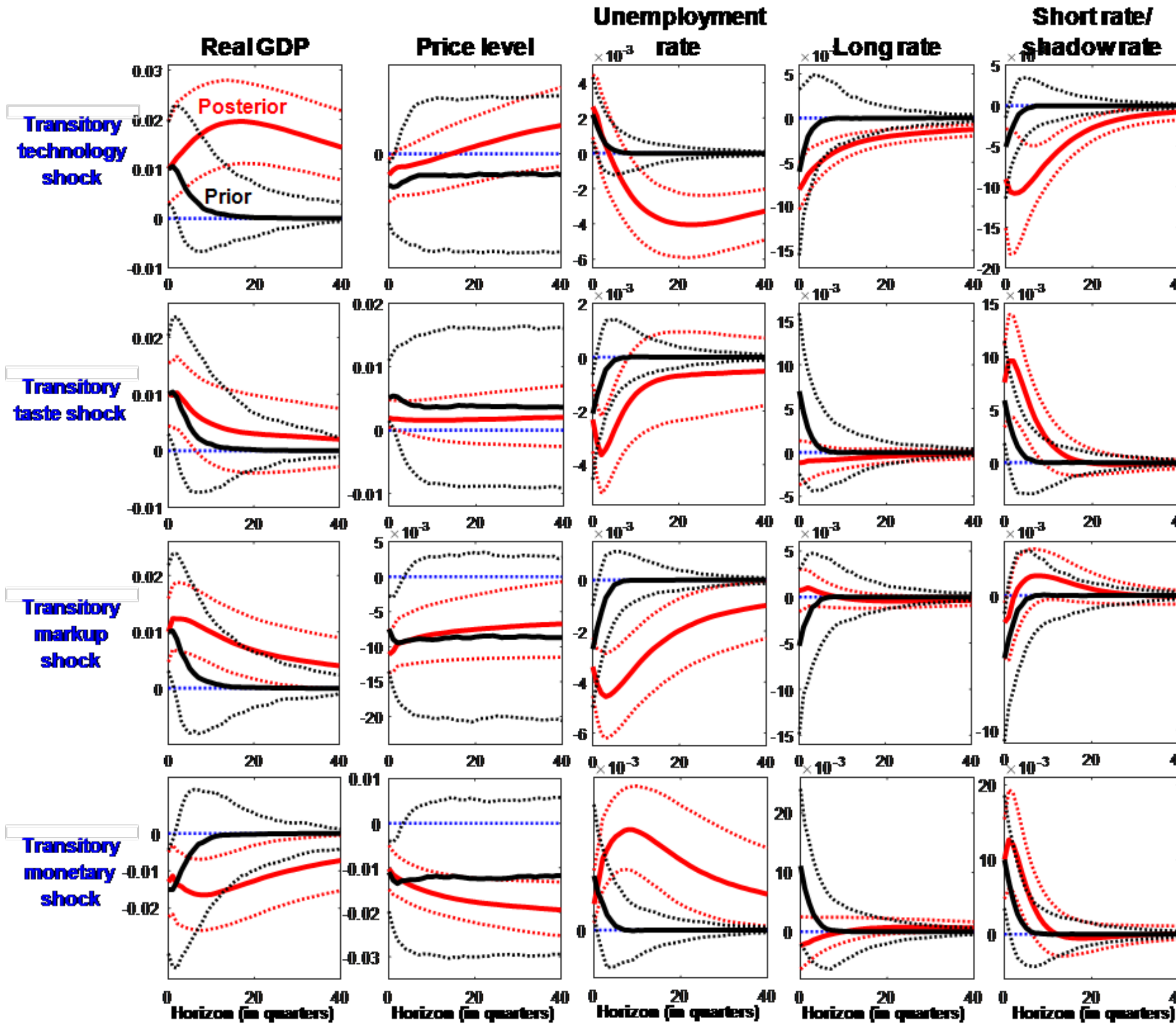

Figure 7: Comparison between prior and posterior IRFs for the ZLB sub-zero period based on joint estimation: median and 16th and 84th percentiles of the distributions of the impulse response functions to the structural shocks.

variate permanent-transitory decomposition as in the individual country examples. The first two columns of Figure 8 show the results, respectively, for the ZLB and pre-ZLB periods. In the last two columns we depict the ZLB IRFs together together with the means from an exercise analogous to that shown in Figure 6.[9]

A comparison between the results from the first two and the last two columns shows that

[9]We took 10,000 draws from the posterior distribution of the SVAR identified using the multivariate permanent-transitory decomposition. The panel VAR is estimated on the pre-ZLB sample. We then simulate the VAR for a length equal to the actual joint sample length for the ZLB period. This is then estimated based on Giannone et al. (2015), and we compute the identified IRFs. For each draw this produces a set of percentiles for the IRFs to the 2 shocks. The means of the percentiles across all of the N simulations are shown in Figure 8 in black.

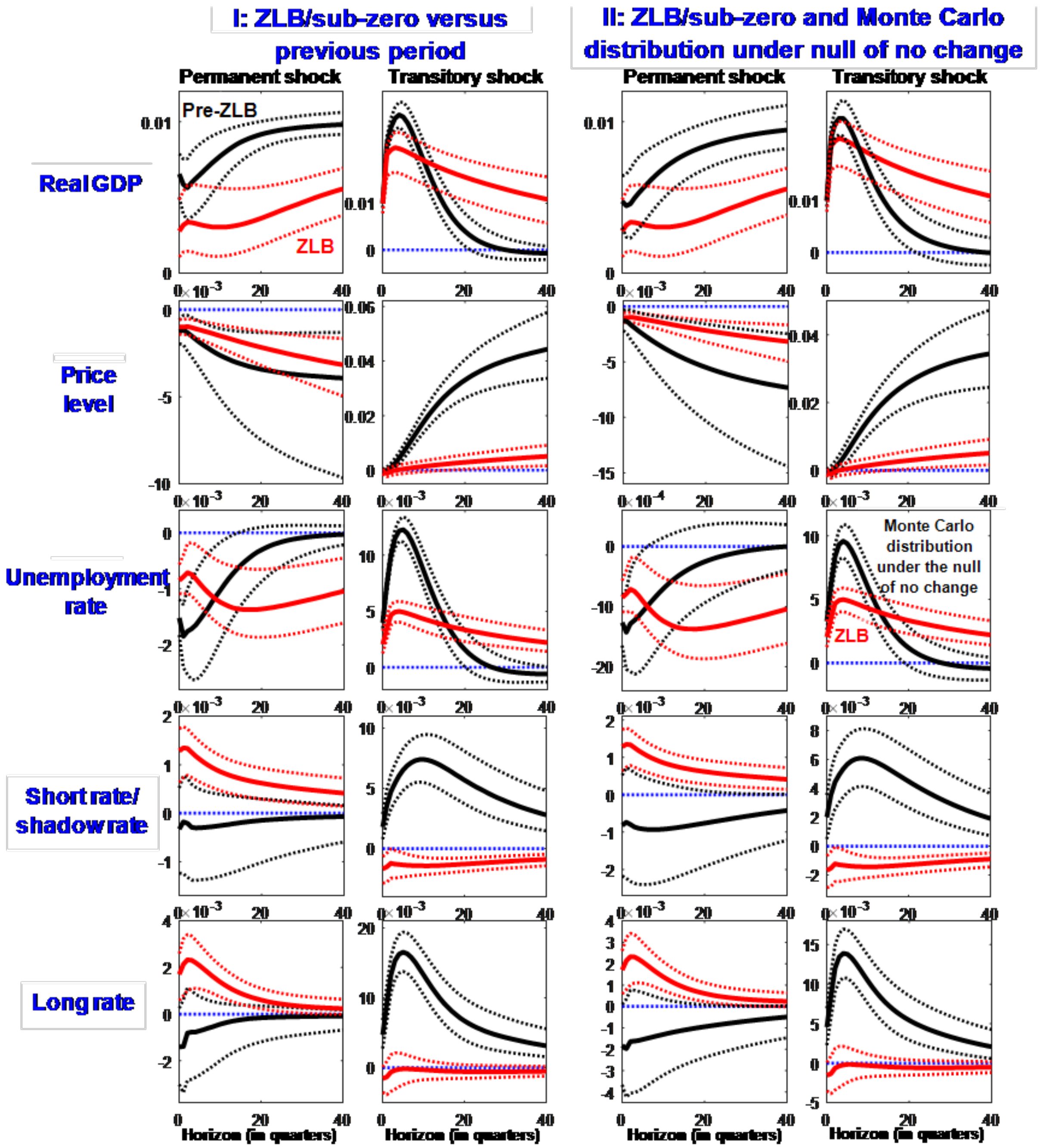

Figure 8: Evidence from a multivariate permanent-transitory decomposition of GDP shocks based on joint estimation: median and 16th and 84th percentiles of the posterior distributions of the estimated impulse-response functions to GDP shocks.

small-sample issue do not play much of a role here. The posterior IRFs for the pre-ZLB period are very close to the means of the Monte Carlo distributions computed under the null hypothesis of no change in the DGP (i.e., the DGP for the ZLB is the same as for the pre-ZLB period). As the figure shows, the evidence produced by this approach is markedly sharper, especially when compared to earlier results. It points towards material changes in the IRFs at the ZLB. In particular, the responses of GDP and unemployment exhibit more inertia than in the standard environment with positive interest rates. Second, the response of prices is flatter. This is especially clear for transitory GDP shocks and slightly less so for permanent shocks. Finally, the responses of interest rates are weaker. Overall, this suggests that there are differences between the ZLB period and the previous period of interest rates away from the lower bound.

5 Conclusion

The financial crisis of the late 2000s and the ensuing recovery during the 2010s is a remarkable period in economic history that has raised a host of questions and issues, not the least of which are the effects of extremely accommodative monetary policy during this time. Central banks held their policy rates at or even below the zero lower bound for extended periods in combination with massive asset purchases. The key question in this context is whether the behavior of the economy is, in fact, different at the zero lower bound not just in terms of the effectiveness of monetary policy but also in terms of the transmission of other shocks.

In this paper we address this issue from a purely empirical perspective. While economic theory delivers a fairly unequivocal answer, namely that dynamics change because of the non-linearity imposed on the economy at the ZLB, finding empirical support is more complicated. The main issue is that the length or ZLB episodes is considerably shorter than the sample size that is typical in macroeconomic data. Consequently, inference is fraught with considerable uncertainty. A related issue is that the changes in the dynamics may be too small to be reliably picked up statistical models.

We confirm this suspicion by means of a Monte Carlo analysis of VAR models with alternative identification schemes. No approach can reliably detect differences between ZLB periods and non-ZLB periods in terms of the response to structural shocks, least of all TVP-VARs. Subsequently, we show that pooling information from several countries that underwent ZLB episodes can sharpen the inference. When we estimate a panel VAR some differences between episodes emerge, but not as sharply as perhaps hoped. We argue that this can be partially alleviated by being mindful of the prior choice by addressing the concerns raised in Baumeister and Hamilton (2015). We show that a specific identification scheme based on a permanent-transitory decomposition is able to hone in on some differences, although they are not large.

The obvious resolution to small-sample issues is to get more data. We considered the case of the US ZLB period where we had monthly data available for the key variables but neither qualitatively nor quantitatively we could identify differences with the quarterly case. It is possible to go to higher-frequency data for financial variables and price variables, but there is a limit to what can be effected with real quantity data. Nevertheless, a mixed-frequency implementation of our study seems a fruitful direction to pursue.

References

An, Sungbae and Frank Schorfheide (2007), "Bayesian analysis of DSGE models." *Econometric Reviews*, 26, 113–172.

Arias, Jonas E., Juan F. Rubio-Ramirez, and Daniel F. Waggoner (2018), "Inference based on structural vector autoregressions identified with sign and zero restrictions: Theory and applications." *Econometrica*, 86, 685–720.

Arias, Jonas E., Juan F. Rubio-Ramirez, and Daniel F. Waggoner (2022), "Uniform priors for impulse responses." *Working Paper*.

Aruoba, Borağan S., Pablo Cuba-Borda, and Frank Schorfheide (2017), "Macroeconomic dynamics near the ZLB: A tale of two countries." *The Review of Economic Studies*, 85, 87–118.

Baumeister, Christiane and James D. Hamilton (2015), "Sign restrictions, structural vector autoregressions, and useful prior information." *Econometrica*, 83, 1963–1999.

Benati, Luca (2008), "Investigating inflation persistence across monetary regimes." *The Quarterly Journal of Economics*, 123, 1005–1060.

Benati, Luca (2015), "The long-run Phillips curve: A structural VAR investigation." *Journal of Monetary Economics*, 76, 15–28.

Benhabib, Jess, Stephanie Schmitt-Grohe, and Martín Uribe (2001), "The perils of Taylor rules." *Journal of Economic Theory*, 96, 40–69.

Benhabib, Jess, Stephanie Schmitt-Grohé, and Martín Uribe (2002), "Avoiding liquidity traps." *Journal of Political Economy*, 110, 535–563.

Boneva, Lena Mareen, R. Anton Braun, and Yuichiro Waki (2016), "Some unpleasant properties of loglinearized solutions when the nominal rate is zero." *Journal of Monetary Economics*, 84, 216–232.

Canova, Fabio and Matthias Paustian (2011), "Business cycle measurement with some theory." *Journal of Monetary Economics*, 58, 345–361.

Christiano, Lawrence, Martin Eichenbaum, and Sergio Rebelo (2011), "When is the government spending multiplier large?" *Journal of Political Economy*, 119, 78 – 121.

Cogley, Timothy and Thomas J. Sargent (2005), "Drifts and volatilities: Monetary policies and outcomes in the post WWII US." *Review of Economic Dynamics*, 8, 262–302.

Corana, A., M. Marchesi, C. Martini, and S. Ridella (1987), "Minimizing multimodal functions of continuous variables with the "simulated annealing" algorithm." *ACM Transactions on Mathematical Software*, 13, 262–280.

Debortoli, Davide, Jordi Galí, and Luca Gambetti (2020), "On the empirical (ir)relevance of the zero lower bound constraint." *NBER Macroeconomics Annual*, 34, 141–170.

Del Negro, Marco, Marc P. Giannoni, and Frank Schorfheide (2015), "Inflation in the great recession and new keynesian models." *American Economic Journal: Macroeconomics*, 7, 168–96.

Del Negro, Marco and Giorgio Primiceri (2015), "Time varying structural vector autoregressions and monetary policy: A corrigendum." *Review of Economic Studies*, 82, 1342–1345.

Eggertsson, Gauti B. and Michael Woodford (2003), "The zero bound on interest rates and optimal monetary policy." *Brookings Papers on Economic Activity*, 2003, 139–211.

Faust, Jon and Eric M. Leeper (1997), "When do long-run identifying restrictions give reliable results?" *Journal of Business and Economic Statistics*, 15, 345–353.

Fernandez-Villaverde, Jesus, Grey Gordon, Pablo Guerron-Quintana, and Juan F. Rubio-Ramirez (2015), "Nonlinear adventures at the zero lower bound." *Journal of Economic Dynamics and Control*, 57, 182–204.

Giannone, Domenico, Michele Lenza, and Giorgio E. Primiceri (2015), "Prior selection for vector autoregressions." *The Review of Economics and Statistics*, 97, 436–451.

Goffe, William L., Gary D. Ferrier, and John Rogers (1994), "Global optimization of statistical functions with simulated annealing." *Journal of Econometrics*, 60, 65–99.

Inoue, Atsushi and Lutz Kilian (2022), "Joint Bayesian inference about impulse responses in VAR models." *Journal of Econometrics*, 231, 457–476. Special Issue: The Econometrics of Macroeconomic and Financial Data.

Krishnamurthy, Arvind and Annette Vissing-Jorgensen (2011), "The effects of quantitative easing on interest rates: Channels and implications for policy." *Brookings Papers on Economic Activity*, 2011, 215–265.

Lubik, Thomas A., Christian Matthes, and Andrew Owens (2016), "Beveridge curve shifts and time-varying parameter vars." *Federal Reserve Bank of Richmond Economic Quarterly*, 102, 197–223.

Ohanian, Lee E. and Andrea Raffo (2012), "Aggregate hours worked in OECD countries: New measurement and implications for business cycles." *Journal of Monetary Economics*, 59, 40–56. Carnegie-NYU-Rochester Conference Series on Public Policy.

Rubio-Ramirez, Juan F., Daniel F. Waggoner, and Tao Zha (2010), "Structural vector autoregressions: Theory of identification and algorithms for inference." *The Review of Economic Studies*, 77, 665–696.

Uhlig, Harald (2005), "What are the effects of monetary policy on output? Results from an agnostic identification procedure." *Journal of Monetary Economics*, 52, 381–419.

Watson, Mark W. (1986), "Univariate detrending methods with stochastic trends." *Journal of Monetary Economics*, 18, 49–75.

Watson, Mark W. (2020), "Comment." *NBER Macroeconomics Annual*, 34, 182–193.

Werning, Ivan (2011), "Managing a liquidity trap: Monetary and fiscal policy." Working Paper 17344, National Bureau of Economic Research.

A Data Appendix

A.1 Euro area

All data are obtained from the European Central Bank, with the exception of a monthly series for the Wu-Xia 'shadow rate', which is from Cynthia Wu's website, at: https://sites.google.com/site/jingcynthiawu/shadowrate_ECB.xls. The shadow rate has been converted to quarterly frequency by taking averages within the quarter. The sample periods used for estimation are 1988Q2-2009Q4 and 2010Q1-2019Q4.

A.2 Denmark

Quarterly seasonally adjusted series for nominal and real GDP are from Statistic Denmark. A monthly series for the monetary policy rate is from Statistic Denmark. It has been converted to quarterly frequency by taking averages within the quarter. In 2012Q3 the monetary policy rate fell below zero. For the subsequent periods, we thus use Wu-Xia's shadow rate for the Euro area, as proxy for the shadow rate. The rationale for doing so is that the monetary policy of Danmarks National Bank aims at keeping the exchange rate of the Danish Krona relative to the Euro essentially fixed, within a $\pm 0.15\%$ tolerance band. The implication is that the monetary policy stance of Danmarks National Bank should be regarded as essentially the same as that of the European Central Bank. A monthly series for the yield on long-term Danish government bonds is from Kim Abildgren's database, available at https://sites.google.com/view/kim-abildgren, and it has been converted to the quarterly frequency by taking averages within the quarter. A monthly seasonally adjusted series for the harmonized unemployment rate ('Total: All Persons for Denmark, Percent, Quarterly, Seasonally Adjusted') is from the OECD Main Economic Indicators. The series has been converted to the quarterly frequency by taking averages within the quarter. The sample periods used for estimation are 1991Q2-2010Q2 and 2010Q3-2019Q4.

A.3 Japan

A quarterly seasonally adjusted series for real GDP is from the Department of National Accounts of the Economic and Social Research Institute, Cabinet Office, Government of Japan. A quarterly seasonally adjusted series for the GDP deflator ('GDP Implicit Price Deflator in Japan, Main Economic Indicators, JPNGDPDEFQISMEI, Index 2010=100') is from the OECD Main Economic Indicators. A quarterly seasonally adjusted series for hours worked per capita is from the Ohanian and Raffo (2012) dataset. The series is available for the period 1960Q2-2017Q1, which dictates the end of the ZLB estimation period. A quarterly seasonally unadjusted series for the monetary policy rate has been constructed by linking the monthly Bank of Japan's discount rate series, up until 1994Q4, and since 1995Q1 the monthly 'shadow rate' series produced by Leo Krippner, available at his website (https://www.ljkmfa.com/test-test/international-ssrs/). The resulting monthly linked series has been converted to the quarterly frequency by taking averages within the quarter. A monthly seasonally unadjusted series for the 10 year government bond yield ('Japan - YIELD, SECOND.MKT, INT.-BEARING GOVT.BONDS,10 YRS,(O.T.C.),M-END BISM.M.HGCA.JP.01') is from the Bank for International Settlements. The series has been converted to the quarterly frequency by taking averages within the quarter. The sample periods used for estimation are 1980Q3-1995Q3 and 1995Q4-2017Q1.

A.4 Switzerland

Quarterly seasonally adjusted series for real GDP ('Gross domestic product, expenditure approach, seasonally and calendar adjusted data, in Mio. Swiss Francs, at prices of the preceding year, chained values, reference year 2010') and the GDP deflator ('Gross domestic product, expenditure approach, seasonally and calendar adjusted data, implicit chain price indexes') are from SECO, the Swiss Statistical Agency. A quarterly series for the monetary policy rate has been constructed by linking the monthly discount rate of the Swiss National Bank for the period 1960Q1-1988Q1; the monthly 3 months LIBOR rate from 1989Q2 until 1994Q4; and since 1995Q1 the monthly 'shadow rate' series produced by Leo Krippner, available at his website (https://www.ljkmfa.com/test-test/international-ssrs/). The resulting monthly linked series has been converted to the quarterly frequency by taking averages within the quarter. A quarterly seasonally unadjusted series for a long rate ('IRLTLT01CHM156N: Long-Term Government Bond Yields: 10-year: Main (Including Benchmark) for Switzerland, Percent, Monthly, Not Seasonally Adjusted') is from the St. Louis FED's internet data portal, FRED II. A quarterly seasonally adjusted series for the unemployment rate ('Registered Unemployment Rate for Switzerland') is from the OECD Main Economic Indicators. A quarterly seasonally unadjusted population series ('POPTOTCHA647NWDB: Population, Total for Switzerland, Persons, Not Seasonally Adjusted') is from the St. Louis FED's internet data portal, FRED II. The sample periods used for estimation are 1983Q2-2008Q4 and 2009Q1-2019Q4.

A.5 United Kingdom

Monthly data A monthly seasonally unadjusted series for the core CPI ('CPIH Index: Excluding Energy, food, alcoholic beverages & tobacco 2015=100', acronym is L5KB) is from the Office for National Statistics (ONS), and it has been seasonally adjusted via ARIMA X-12 as implemented in Eviews. A monthly series for the 20-year government bond yield is from the ONS (the series' code is AJLX). A monthly series for the monetary policy rate has been constructed by linking the *Bank of England*'s rate (i.e., the 'Bank rate') and Wu-Xia's 'shadow rate' from Cynthia Wu's website. Specifically, the resulting monetary policy rate series is equal to the Bank rate up until the collapse of Lehman Brothers (September 2008), and it is equal to Wu-Xia's 'shadow rate' after that. A monthly seasonally adjusted series for real GDP is from the National Institute for Economic and Social Research (NIESR), and it has been kindly provided by Garry Young. A monthly seasonally adjusted series for the unemployment rate has been constructed by linking the series 'Monthly administrative unemployment rates and levels 1881-2015', from the spreadsheet of very long-run statistics millenniumofdata_v3_final.xls, which is available at the Bank of England's website (until December 2015), and the series 'Unemployment rate (aged 16 and over, seasonally adjusted)' (the series' acronym is MGSX) since then. The sample periods used for estimation are January 1983-March 2009 and April 2009-December 2019.

Quarterly data Quarterly seasonally adjusted series for real GDP ('ABMI: Real GDP at market prices, £ million at chained volume measures') and the GDP deflator ('L8GG: Implied GDP deflator at market prices: SA Index') are from the Office for National Statistics (ONS). A monthly series for the 20-year government bond yield is from the ONS (series code AJLX). The series has been converted to quarterly frequency by taking averages within the quarter. A monthly series for the monetary policy rate has been constructed by linking the Bank of England's rate (i.e., the 'Bank rate') and the Wu-Xia 'shadow rate' from Cynthia Wu's website. Specifically, the resulting monetary policy rate series is equal to the Bank rate up until the collapse of Lehman Brothers (September 2008), and it is equal to Wu and Xia's 'shadow rate' after that. The resulting linked series has been converted to the quarterly frequency by taking averages within the quarter. A quarterly seasonally adjusted series for total hours worked has been constructed by linking the series from Ohanian and Raffo (2012) and the series 'HOUR03: Average actual weekly hours of work by industry sector2: People all in employment seasonally adjusted' from the ONS (the resulting linked series is equal to the one from Ohanian and Raffo (2012) until 2016Q4, and to that from the ONS since then). Over the period of overlapping the two series are identical, which justifies their linking. The sample periods used for estimation are 1983Q1-2009Q1 and 2009Q2-2019Q4.

A.6 United States

Monthly data A monthly seasonally adjusted series for the core PCE deflator ('Personal Consumption Expenditures Excluding Food and Energy (Chain-Type Price Index)') is from the US Bureau of Economic Analysis. A seasonally adjusted monthly series for real is from Stock and Watson (2012) until 2010, and from IHS Markit, after that (available at: https://ihsmarkit.com/products/us-monthly-gdp-index.html) IHS Markit's production notes for its monthly real GDP series states:

> '**Note**: IHS Markit's index of Monthly GDP (MGDP) is a monthly indicator of real aggregate output that is conceptually consistent with real Gross Domestic Product (GDP) in the NIPA's. The consistency is derived from two sources. First, MGDP is calculated using much of the same underlying monthly source data that is used in the calculation of GDP. Second, the method of aggregation to arrive at MGDP is similar to that for official GDP. Growth of MGDP at the monthly frequency is determined primarily by movements in the underlying monthly source data, and growth of MGDP at the quarterly frequency is nearly identical to growth of real GDP.'

A monthly series for the monetary policy rate has been constructed by linking the federal funds rate series from FRED II (series name FEDFUNDS) and the Wu-Xia 'shadow rate' from Cynthia Wu's website. Specifically, the resulting monetary policy rate series is equal to the federal funds rate up until the collapse of Lehman Brothers (September 2008), and it is equal to Wu-Xia's 'shadow rate' after that. A monthly seasonally adjusted series for the unemployment rate is the 'Civilian Unemployment Rate' series from the Bureau of Labor Statistics. The series is available from FRED II, and the series name is UNRATE. A monthly seasonally unadjusted series for the 10-year Treasury constant maturity rate is from the Board of Governors of the Federal Reserve System, and is available from FRED II (series name is GS10). A monthly seasonaly unadjusted series for the civilian non-institutional population is from the Bureau of Labor Statistics, and is available from FRED II (series name is CNP16OV).

Quarterly data A monthly series for the monetary policy rate has been constructed by linking the federal funds rate from St. Louis FED ('FEDFUNDS: Effective Federal Funds Rate, Monthly, Not Seasonally Adjusted, Percent') and Wu-Xia's 'shadow rate' from Cynthia

Wu's website. Specifically, the resulting monetary policy rate series is equal to the federal funds rate up until the collapse of Lehman Brothers (September 2008), and it is equal to Wu-Xia's 'shadow rate' after that. The resulting linked series has been converted to the quarterly frequency by taking averages within the quarter. A monthly seasonally unadjusted series for the 10-year government bond yield ('GS10: 10-Year Treasury Constant Maturity Rate, Percent, Monthly, Not Seasonally Adjusted') is from St. Louis FED. It has been converted to the quarterly frequency by taking averages within the quarter. A quarterly seasonally adjusted series for real GDP ('GDPC1: Real Gross Domestic Product, Billions of Chained 2012 Dollars, Quarterly, Seasonally Adjusted Annual Rate') are from the Bureau of Economic Analysis. A quarterly seasonally adjusted series for the GDP deflator ('GDPCTPI: Gross Domestic Product: Chain-type Price Index') are from the Bureau of Economic Analysis. A quarterly seasonally adjusted series for working age population (i.e., aged 15-64) has been constructed by linking the series from Ohanian and Raffo (2012) and the series 'Working Age Population, Aged 15-64, Noninstitutional, non-armed forces Population for the United States, Persons, Quarterly, Seasonally Adjusted' from the OECD's Main Economic Indicators database (the resulting linked series is equal to the one from Ohanian and Raffo (2012) until 1999Q4, and to that from the OECD since then). Over the period of overlapping the two series are identical, which justifies their linking. A quarterly seasonally adjusted series for hours of all persons in the non-farm business sector is from the U.S. Bureau of Labor Statistics. The sample periods used for estimation are 1983Q1-2008Q3 and 2008Q4-2015Q4.

B Computing the VAR's Structural Impact Matrix of the Shocks

B.1 Long-run and sign restrictions

When identifying the VARs by combining zero and sign restrictions we compute the VAR's structural impact matrix of the shocks using the methodology proposed by Arias et al. (2018). The approach is exactly the same as in Benati (2015), with the only difference that we are here imposing a different set of long-run and sign restrictions on a different set of variables.

Let B_0^j, B_1^j, ..., B_p^j, and Ω^j be the j-th draw from the posterior distribution for the intercept, the VAR matrices, and the covariance matrix of reduced-form innovations of the VAR, for j = 1, 2, 3, ..., J. Let $P_j D_j P_j'$ be the eigenvalue-eigenvector decomposition of Ω^j. We start by computing an initial estimate of A_0^j, denoted $\tilde{A}_0^j$, as $\tilde{A}_0^j = P_j D_j^{\frac{1}{2}}$ with the corresponding matrix of long-run impacts of the structural shocks $\tilde{L}_j = [I_N - B^j(1)]^{-1}\tilde{A}_0^j$, where N is the number of series in the VAR, and $B^j(1) = B_1^j + B_2^j + ... + B_p^j$. Based on the Gibbs-sampling algorithm described in Arias et al. (2018), we then draw Z random orthonormal matrices of dimension $N \times N$ from the uniform distribution, conditional on the zero restrictions on the long-run impacts on real GDP of the four transitory shocks. Let Q_z^j, $z = 1, 2, 3, ..., Z$, be the z-th random orthonormal matrix, with $Q_z^j(Q_z^j)' = I_N$. we then combine each of the Z random orthonormal matrices with the initial estimate of the long-run impact of the structural shocks, $\tilde{L}_j$, in order to obtain a randomly rotated long-run impact matrix, $L_j^z = \tilde{L}_j Q_z^j$. By construction, each L_j^z, $z = 1, 2, 3, ..., Z$, satisfies the zero long-run restrictions pertaining to the four transitory shocks. By applying the Q_z^j, $z = 1, 2, 3, ...,$ Z, to the initial estimate of the VAR's structural impact matrix of the shocks, $\tilde{A}_0^j$, we then obtain the corresponding candidate estimates of the structural impact matrix, $A_{0,j}^z = \tilde{A}_0^j Q_z^j$. Finally, we check whether $A_{0,j}^z$ satisfies the sign restrictions reported in Table 1 in the main text of the paper, and out of the Z candidate structural impact matrices we only keep, for draw j, those satisfying the sign restrictions. For each draw from the posterior we consider 100 random rotation matrices. Finally, we set the number of Gibbs-sampling iterations in the algorithm to $L = 10$.

B.2 Pure sign restrictions

When identifying the VARs based on a pure sign restrictions approach we compute the structural impact matrix of the shocks *via* the methodology proposed by Rubio-Ramirez et al. (2010).

Let B_0^j, B_1^j, ..., B_p^j, and Ω^j be the j-th draw from the posterior distribution for the intercept, the VAR matrices, and the covariance matrix of reduced-form innovations of the VAR, for $j = 1, 2, 3, ..., J$. Let $P_j D_j P_j'$ be the eigenvalue-eigenvector decomposition of Ω^j. We start by computing an initial estimate of A_0^j—let's call it $\tilde{A}_0^j$—as $\tilde{A}_0^j = P_j D_j^{\frac{1}{2}}$. Given $\tilde{A}_0^j$, we then draw $z = 1, 2, ..., Z$ $(N \times N)$ matrices M_z from a $Normal(0, 1)$ distribution, where N is the number of series in the VAR, and we take the the QR decomposition of M_z, i.e. thus obtaining matrices Q_z and R_z such that $M_z = Q_z R_z$, with Q_z being an orthogonal matrix and R_z being upper triangular. Finally, we compute the the candidate structural impact matrix as $A_{0,j}^z = \tilde{A}_0^j Q_z'$. Finally, we check whether $A_{0,j}^z$ satisfies the sign restrictions reported in Table 1 in the main text of the paper, and out of the Z candidate structural impact matrices we only keep, for draw j, those satisfying the sign restrictions. For each draw from the posterior we consider 100 random rotation matrices.

C The New Keynesian Model Used as a Data Generating Process

The DSGE model we use as data generation process (DGP) in the Monte Carlo exercise of Section 3 is the standard New Keynesian model with both forward- and backward-looking components in either the IS or the Phillips curve estimated, for instance, in Benati (2008), with the only difference that log real GDP, $y_t = \ln(Y_t)$, is here assumed to feature a unit root component that evolves as a random-walk with drift as in Watson (1986),

$$y_t^P = y_{t-1}^P + \delta + v_t,\ v_t \sim WN(0, \sigma_v^2). \tag{C.1}$$

The rest of the model is standard, and once log-linearized takes the form

$$\pi_t = \frac{\beta}{1+\alpha\beta}\pi_{t+1|t} + \frac{\alpha}{1+\alpha\beta}\pi_{t-1} + \kappa\hat{y}_t + u_t,\ u_t \sim WN(0, \sigma_u^2) \tag{C.2}$$

$$\hat{y}_t = \gamma\hat{y}_{t+1|t} + (1-\gamma)\hat{y}_{t-1} - \sigma^{-1}(R_t - \pi_{t+1|t}) - (1-\gamma)\Delta y_t^P \tag{C.3}$$

$$R_t = \rho R_{t-1} + (1-\rho)[\phi_\pi \pi_t + \phi_y \hat{y}_t] + \epsilon_{R,t},\ \epsilon_{R,t} \sim WN(0, \sigma_R^2) \tag{C.4}$$

where π_t and R_t are inflation and the nominal interest rate. Real GDP, Y_t, is stationarized by rescaling it by its unit root component Y_t^P, so that $\hat{y}_t \equiv \ln(Y_t/Y_t^P)$ is the log-deviation of GDP from its stochastic trend. The rest of the notation is standard: $\alpha \in [0, 1]$ is price setters' extent of indexation to past inflation; $\gamma \in [0, 1]$ is the forward-looking component in the intertemporal IS curve; κ and σ are the slope of the Phillips curve and the elasticity of intertemporal substitution in consumption, respectively; and ρ, ϕ_π, and ϕ_y are the smoothing parameter and the coefficients on inflation and the output gap in the monetary rule, respectively.

We estimate the model based on U.S. data for the Federal Funds rate, GDP deflator inflation and log real GDP for the period 1954Q3-2008Q4. Specifically, we calibrate $\delta = 0.0074$, which is the average value taken by the log-difference of real GDP over the sample period, and we estimate all of the other parameters via standard Bayesian methods exactly as in Benati (2008). Table D.1 reports, for each parameter, the chosen prior density, its mode, and its standard deviation, together with the median and 16th and 84th percentiles of the posterior distribution, which we generate via Random-Walk Metropolis.

Table D.1 Prior distributions for the structural parameters, and median and 16th and 84th percentiles of the posterior distribution					
			Prior		*Posterior*
Parameter	Domain	Density	Mode	Standard deviation	Median and 16th and 84th percentiles
σ_R^2	$\mathbb{R}^+$	Inverse Gamma	5×10^{-3}	5×10^{-3}	6.0×10^{-4} [5.0 7.0]$\times10^{-4}$
σ_u^2	$\mathbb{R}^+$	Inverse Gamma	10^{-2}	10^{-2}	1.3×10^{-3} [1.1 1.5]$\times10^{-3}$
σ_v^2	$\mathbb{R}^+$	Inverse Gamma	10^{-2}	10^{-2}	3.0×10^{-3} [2.4 3.4]$\times10^{-3}$
κ	$\mathbb{R}^+$	Gamma	0.1	0.1	0.089 [0.060 0.131]
σ	$\mathbb{R}^+$	Gamma	1	2	11.736 [9.281 15.066]
α	[0, 1]	Beta	0.9	0.05	0.852 [0.772 0.913]
γ	[0, 1]	Beta	0.5	0.25	0.486 [0.474 0.498]
ρ	[0, 1)	Beta	0.75	0.1	0.695 [0.633 0.747]
ϕ_π	$\mathbb{R}^+$	Gamma	1.5	0.25	2.386 [2.085 2.706]
ϕ_y	$\mathbb{R}^+$	Gamma	0.5	0.15	0.255 [0.190 0.335]

We numerically maximise the log posterior, defined as $lnL(\theta|Y) + lnP(\theta)$, where θ is the vector collecting the model's structural parameters, $lnL(\theta|Y)$ is the likelihood of θ conditional on the data, and $P(\theta)$ is the prior through simulated annealing. Following Goffe et al. (1994) we implement simulated annealing *via* the algorithm proposed by Corana et al. (1987), setting the key parameters to T_0=100,000, r_T=0.9, N_t=5, N_s=20, ϵ=10^{-6}, N_ϵ=4, where T_0 is the initial temperature, r_T is the temperature reduction factor, N_t is the number of times the algorithm goes through the N_s loops before the temperature starts being reduced, N_s is the number of times the algorithm goes through the function before adjusting the stepsize, ϵ is the convergence (tolerance) criterion, and N_ϵ is number of times convergence is achieved before the algorithm stops. Finally, initial conditions were chosen stochastically by the algorithm itself, while the maximum number of functions evaluations, set to 1,000,000, was never achieved.

In implementing the Random-Walk Metropolis (RWM) algorithm we exactly follow An and Schorfheide (2007), with the exception of the method we use to calibrate the covariance matrix's scale factor (the parameter c below), for which we follow the methodology we describe in the next paragraph. Let $\hat{\theta}$ and $\hat{\Sigma}$ be the mode of the maximized log posterior and its estimated Hessian, respectively (we compute $\hat{\Sigma}$ numerically as in An and Schorfheide (2007). We start the Markov chain of the RWM algorithm by drawing $\theta^{(0)}$ from $N(\hat{\theta}, c^2\hat{\Sigma})$. For s = 1, 2, ..., N we then draw $\tilde{\theta}$ from the proposal distribution $N(\theta^{(s-1)}, c^2\hat{\Sigma})$, accepting the jump (i.e., $\theta^{(s)} = \tilde{\theta}$) with probability $min[1, r(\theta^{(s-1)}, \theta|Y)]$, and rejecting it (i.e., $\theta^{(s)} =$

$\theta^{(s-1)}$) otherwise, where

$$r(\theta^{(s-1)}, \theta|Y) = \frac{L(\theta|Y)\ P(\theta)}{L(\theta^{(s-1)}|Y)\ P(\theta^{(s-1)})}$$

A key problem in implementing Metropolis algorithms is how to calibrate the covariance matrix's scale factor in order to achieve an acceptance rate of the draws close to the ideal one (in high dimensions) of 0.23. Typically the problem is tackled by starting with some 'reasonable' value for c, and adjusting it after a certain number of iterations during the initial burn-in period. Specifically, given that the draws' acceptance rate is decreasing in c, c gets increased (decreased) if the initial acceptance rate was too high (low). A problem with this approach is that it does not guarantee that after the adjustment the acceptance rate will be reasonably close to the ideal one. The approach for calibrating c used in this paper, on the other hand, is based on the idea of estimating a reasonably good approximation to the inverse relationship between c and the acceptance rate by running a pre-burn-in sample. Specifically, let C be a grid of possible values for c. In what follows, we consider a grid over the interval [0.1, 1] with increments equal to 0.05. For each single value of c in the grid, call it c_j, we run n draws of the RWM algorithm, storing, for each c_j, the corresponding fraction of accepted draws, f_j. we then fit a third-order polynomial to the f_j's *via* least squares, and letting $\hat{a}_0$, $\hat{a}_1$, $\hat{a}_2$, and $\hat{a}_3$ be the estimated coefficients, we choose c by solving numerically the equation $\hat{a}_0+\hat{a}_1 c+\hat{a}_2 c^2+\hat{a}_3 c^3=0.23$. The fraction of accepted draws from RWM, equal to 0.2288, testifies to the good performance of this procedure.