# Time Varying IV-SVARs based on kernel estimators *

Robin Braun
Federal Reserve Board
robin.a.braun@frb.gov

George Kapetanios
King's College, London
george.kapetanios@kcl.ac.uk

Massimiliano Marcellino
Bocconi University, IGIER and CEPR
massimiliano.marcellino@unibocconi.it

September 3, 2024

**Abstract**

This paper studies the estimation and inference of time-varying impulse response functions in structural vector autoregressions (SVARs) identified with external instruments. Building on kernel estimators that allow for nonparametric time variation, we derive the asymptotic distributions of the relevant quantities. Our estimators are simple and computationally trivial and allow for potentially weak instruments. Simulations suggest satisfactory empirical coverage even in relatively small samples—as long as the underlying parameter instabilities are sufficiently smooth. We illustrate the methods by studying the time-varying effects of global supply news shocks on US industrial production, and monetary policy shocks on financial variables in the UK.

**JEL classification**: C14, C32, C53, C55
**Keywords**: Time-varying parameters, Nonparametric estimation, Structural VAR, External instruments, Weak instruments, Monetary shocks, Impulse response analysis

*We thank seminar participants at the European Central Bank, Bocconi University, Bank of England, King's College, the Friendly Faces workshop, University of Konstanz, University of East Anglia, the Philadelphia Fed, the IAAE 2022, ESEM 2022, the CFE, and a conference at Harvard University for useful comments. We also thank Jonas Arias, Ralf Brüggemann, Ambrogio Cesa-Bianchi, Thorsten Drautzburg, Sophocles Mavroeidis, Aaron Metheny, Silvia Miranda-Agrippino, Pascal Paul, Mikkel Plagborg-Moller, Jim Stock and Mark Watson for insightful comments. Marcellino thanks Ministero dell Istruzione, Universita e Ricerca - PRIN Bando 2017 prot. 2017TTA7TYC for financial support. The views expressed in this paper are those of the authors and do not necessarily reflect those of the Board of Governors of the Federal Reserve System.

# 1 Introduction

Instrumental variable (IV) identification of structural vector autoregressions (IV-SVARs) has become increasingly popular in empirical macroeconomics. For example, with respect to the effects of monetary policy, many recent influential studies rely on this identification strategy, including Gertler and Karadi (2015), Caldara and Herbst (2019), Jarociński and Karadi (2020), Miranda-Agrippino and Ricco (2021). At the same time, important refinements to the methodology have been developed, building on the pioneering contributions of Stock (2008), Stock and Watson (2018) and Mertens and Ravn (2013). This includes the conduct of Bayesian inference (Arias et al., 2021; Giacomini et al., 2022), the establishment of the connection to local projections and the robustness to noninvertibility (Plagborg-Møller and Wolf, 2021, 2022; Forni et al., 2023), and, importantly, robust inference under weak identification (Montiel-Olea et al., 2021). Furthermore, Paul (2020) introduced the possibility of allowing for time-varying parameters in a Bayesian setting.

In this paper we contribute to the literature in various ways. First, from a theoretical point of view, we introduce slowly changing parameters in the IV-SVAR, which are meant to capture instabilities in macroeconomic relationships—see, for example, Stock and Watson (1996). While we take no stance on what causes the parameter changes, institutional modifications, technological developments, economic trends such as globalization, and an evolving policy toolkit are potential causal factors that are often discussed.

Our paper complements and extends previous work on time-varying IV-SVARs by Paul (2020) in many ways. First, instead of spelling out a law of motion for the model coefficients, we take a nonparametric approach that relies on persistence and smoothness assumptions on the pattern of parameter evolution. Formally, we build on classical kernel-based estimators introduced by Giraitis et al. (2014) and adapted for IV estimation in Giraitis et al. (2021). Besides its nonparametric nature, our frequentist inference procedure for the model parameters and structural impulse response functions (IRFs) is computationally trivial and scales easily to

larger dimensions and sample sizes. Furthermore, unlike the Bayesian alternative, inference can be robustified to account for potentially weak identification.

Second, we study two estimators that cater to different needs of the researcher—namely, (1) the classical IV-SVAR and (2) the internal instrument VAR estimator proposed by Plagborg-Møller and Wolf (2021). Under shock invertibility of the model, the IV-SVAR may be a powerful device, as it allows the researcher to back out the structural shocks up to a known constant. Hence, it is straightforward to construct time-varying IRFs that remain comparable over time — for example, by considering shocks of unit standard deviation.

Shock invertibility can be tested within time-varying IV-SVARs. If rejected, it is still possible to rely on the internal instrument estimator, which allows one to estimate relative IRFs consistently in the absence of invertibility. However, given that the shock's scale remains unknown in the internal instrument VAR, a stronger time-invariance assumption is required about the covariance between the shock of interest and the external instrument. In this case, one can obtain a comparable shock size across time by fixing the variation induced by the instrument to a constant (Paul, 2020).

Our main object of interest is IRFs, and in order to conduct inference for the relevant quantities, we proceed in two steps. First, we derive the asymptotic theory for the corresponding reduced-form parameters that characterize the joint dynamics of the endogenous time series and the external instrument. We then either rely on an application of the delta method or follow Montiel-Olea et al. (2021) in constructing confidence sets via an inversion of the Anderson–Rubin test statistic. The latter has the advantage of providing confidence-set robustness to a situation in which the instrument is only weakly correlated with the shock of interest—see, for example, Staiger and Stock (1997). This feature can be particularly important when a smaller bandwidth of the kernel estimator lowers the effective sample size considerably, even in larger sample sizes. Both methods are accompanied by closed-form solutions that allow for computationally efficient implementation.

In order to understand the finite sample properties of the proposed method, we proceed with a Monte Carlo exercise. Following Montiel-Olea et al. (2021), we calibrate data-generating processes that estimate time-varying versions of the global oil market VAR model of Kilian (2009). We are able to obtain satisfactory empirical coverage if the evolution of parameters is sufficiently smooth. When selecting the bandwidth with a data-driven method that targets out-of-sample model fit, we document only a small deterioration in empirical coverage.

We then apply our method in two different set-ups. First, we study the time-varying effects of monetary policy on financial variables in the United Kingdom. Based on daily data from mid-2004 to early 2023, we estimate a series of rich time-varying parameter (TVP) IV-SVARs. The model includes interest rates at different maturities, the exchange rate, stock prices, (non-financial) corporate bond spreads and yields, and market measures of inflation compensation. The sample we consider covers important structural events that are likely to have altered the way monetary policy transmits to financial conditions; these include a period at the zero lower bound (ZLB) in the aftermath of the financial crisis, the introduction of quantitative easing (QE) and the explicit use of forward guidance, and policies to support liquidity and macroeconomic conditions during the pandemic. Based on an external instrument approach that builds on the monetary event study literature pioneered by Kuttner (2001) and Cochrane and Piazzesi (2002), we document significant time variation in the dynamic effect of monetary policy on financial variables. While monetary policy acted primarily through short-term interest rates in the earlier part of the sample—and to some extent again after 2021—we document consistently strong responses of 2- and 10-year rates even during the ZLB. Furthermore, we find evidence suggesting that monetary policy has become more effective at steering financial variables, particularly stock market prices, corporate financing conditions, and medium-term inflation expectations.

In a second application, we study the time-varying transmission of supply news shocks on US industrial production. Building on Känzig (2021), we study a time-varying IV-SVAR that

includes monthly macroeconomic variables for the global crude-oil market and key US aggregates. For identification, the model relies on an external instrument that leverages financial price movements around OPEC production quota announcements. A constant parameter model suggests that these shocks transmit as supply type of shocks, where US industrial production declines with increasing real oil prices. However, our methodology reveals strong time-variation in the estimated impulse response functions which aligns well with the shale-oil revolution. US mining output, which includes extraction of oil- and gas, reacts more strongly and quickly nowadays than in the past. Furthermore, US manufacturing output no longer declines, challenging that positive oil price news still transmit as clear cost-push shocks to the US manufacturing sector. Our finding complement recent evidence presented in Bjørnland and Skretting (2024) on the time-varying impact of oil-price shocks. However, in contrast to that paper, we identify an oil-market specific shock by external instruments instead of exclusion restrictions, and use kernel based methods instead of Bayesian techniques.

## Related Literature

Our paper builds on seminal work by Cogley and Sargent (2005) and Primiceri (2005), who introduced TVP into VARs, or TVP-VARs, allowing for coefficients to evolve according to a random walk. Paul (2020) extends this framework to identification of IRFs via external instruments. Our methodological approach is distinct from these papers along various dimensions. First, we rely on an entirely frequentist, nonparametric approach that leverages simple kernel estimators. This approach has several practical advantages. First, it avoids a parametric choice for the law of motion underlying the parameters and instead relies on nonparametric smoothness conditions. Second, it is computationally straightforward and can handle large datasets, for example financial variables available on a daily frequency. Moreover, the framework allows for the leveraging of robust inference methods for IRFs that remain valid under weak instruments. Finally, unlike previous papers, we also cover the internal IV estimator of Plagborg-Møller and Wolf (2021), which can be practically important if there is evidence

against invertibility of the target shock in the IV-SVAR.

Our theory for TVP kernel estimators in IV-SVARs largely builds on earlier work of Giraitis et al. (2014), Giraitis et al. (2018) and Giraitis et al. (2021). However, our paper provides additional results that are required to accommodate identification via external instruments, including the asymptotic distribution of the covariance matrix estimator, the construction of confidence sets, and the joint distribution of neighboring estimators. The joint distribution allows for inference of IRFs that are comparable over time when studying relative IRFs in the internal instrument VAR.

### Outline

The paper is organized as follows. Section 2 develops the methodology for kernel-based inference in TVP IV-SVARs. Section 3 presents results of Monte Carlo simulations. Section 4 studies the effects of U.K. monetary policy on financial variables, while section 5 studies the transmission of oil supply news shocks on the US economy. Section 6 recapitulates. Proofs of the theoretical results and additional empirical results are gathered in a set of appendices.

## 2 Methodology

In this section, we start revisiting instrumental variable identification of Impulse Response Functions in a constant-parameter SVARs. We then generalize the model towards time-varying coefficients, discuss normalization of shock sizes across time, and inference of reduced form quantities via kernel based methods. Finally, we show how those results can be leveraged to compute confidence sets of impulse response functions, which is the ultimate object of interest in this paper.

## 2.1 Identification of VAR impulse response functions via external instruments

Throughout this paper, we consider the $n$-variate SVAR($p$) model given by:

$$y_t = \nu + A_1 y_{t-1} + A_2 y_{t-2} + \ldots + A_p y_{t-p} + u_t, \qquad u_t \sim \mathcal{WN}(0, \Sigma) \tag{1}$$

$$u_t = B\varepsilon_t, \qquad \varepsilon_t \sim \mathcal{WN}(0, I), \tag{2}$$

where $y_t = (y_{1t}, \ldots, y_{nt})'$ is a $n \times 1$ vector of endogenous time series, $\nu$ is a $n \times 1$ vector of intercepts and $A_i, i = 1, \ldots, p$ are $n \times n$ matrices of autoregressive coefficients. Equation (1) describes the reduced form VAR dynamics of $y_t$ as function of lagged realizations and a vector of $n \times 1$ white noise (WN) residuals $u_t$ with full covariance matrix $\Sigma$. Equation (2) relates the prediction errors $u_t$ to $n \times 1$ structural shocks $\varepsilon_t$ whose elements are orthogonal and standardized to unit variance. The $n \times n$ matrix $B$ is the contemporaneous impact matrix and reflects the immediate responses of the variables $y_t$ to the structural shocks $\varepsilon_t$. For the moment, we assume that the model is stable, which implies that the SVAR($p$) has a MA($\infty$) representation given by $y_t = \mu_y + \sum_{j=0}^{\infty} C_j(A) B \varepsilon_{t-j} = \mu_y + \sum_{j=0}^{\infty} \Theta_j \varepsilon_{t-j}$, where $\mu_y = \mathrm{E}(y_t)$ and the $n \times n$ coefficient matrices $\Theta_j = C_j(A)B$, are the structural impulse response functions (IRFs). The reduced form MA($\infty$) matrices $C_j(A)$ can be computed recursively from $C_j(A) = \sum_{i=1}^{j} C_{j-1}(A) A_i$ with $C_0(A) = I_n$ and $A_i = 0$ for $i > p$.

The main focus of this paper is the computation of impulse responses to a single shock. Without loss of generality, let this shock be ordered first in the system ($\varepsilon_{1,t}$) and call it the *target shock*. Corresponding IRFs are then given by picking the corresponding element in the MA($\infty$) matrices:

$$\frac{\partial Y_{i,t+k}}{\partial \varepsilon_{1,t}} = \lambda_{k,i} = e_i' C_k(A) B e_1, \tag{3}$$

where $e_i$ denotes the $i$th column of the identity matrix $I_n$. Hence, equation (3) defines the IRFs $\lambda_{k,i}$ as the dynamic effect a unit standard deviation shock in $\varepsilon_{1t}$ on variable $i$, $k$ periods into the future.

It is important to note that, without further assumptions, these IRFs are not identified. The reason is that the same reduced form dynamics of the VAR forecast errors $u_t = B\varepsilon_t$ are obtained for any alternative structural model $\tilde{B} = BQ$ where $Q$ is an orthogonal rotation matrix ($\{Q : Q'Q = I_n, Q' = Q\}$). To see this, note that both models imply the same reduced form covariance matrix $\Sigma_u = BB' = BQQ'B = \tilde{B}\tilde{B}'$. Throughout this paper, we rely on an identification strategy that involves an instrumental variable $z_t$ for the target shock (Stock and Watson, 2012; Mertens and Ravn, 2013).

**Assumption 1** (External Instrument)**.** *Let $z_t$ be an instrument for the first shock. The stochastic process $\{(\varepsilon_t, z_t)\}_{t=1}^{\infty}$ satisfies:*

1. *$E(z_t\varepsilon_{1t}) = \alpha \neq 0$,*
2. *$E(z_t\varepsilon_{jt}) = 0$ for $j \neq 1$.*

Assumption 1 allows to identify $b_1 = Be_1$ up to scale and sign normalization, since:

$$\Gamma = \mathrm{E}\left(u_t z_t\right) = B \begin{pmatrix} \mathrm{E}[\varepsilon_{1t} z_t] \\ \mathrm{E}[\varepsilon_t^o z_t] \end{pmatrix} = \left[\begin{array}{c|c} b_1 & b_2 \end{array}\right] \begin{pmatrix} \alpha \\ 0 \end{pmatrix} = \alpha b_1,$$

where $\varepsilon_t^o = [\varepsilon_{2t}, \ldots, \varepsilon_{nt}]'$. In words, the correlation between the external instrument and reduced form prediction error is proportional to $b_1$.

In order to get rid of $\alpha$ and obtain interpretable magnitudes, there are two popular approaches to normalize IRFs in IV-SVARs. The first is known as unit shock standardization (Stock and Watson, 2016) and is particularly simple to compute. Here, the shock variance is re-normalized to yield IRFs that increase the first variable by unity on impact (say $\tilde{b}_{11} = 1$), and hence it is often referred to as *relative* IRFs. In that case, it holds that $\Gamma_{11} = \mathrm{E}(z_t u_{1,t}) = \alpha$, implying that $\tilde{b}_1 = \Gamma / e_1'\Gamma$ is the first column of the rescaled impact matrix measuring the response to a target shock with unidentified standard deviation $\mathrm{Var}(\tilde{\varepsilon}_{1t}) = b_{11}^2$. Corresponding IRFs as function of reduced form parameter are then given by:

$$\lambda_{k,i}^1 = e_i' C_k(A)\Gamma / e_1'\Gamma. \tag{4}$$

The second approach is known as IRFs in response to a shock of unit standard deviation, that

is $\varepsilon_{1t} \sim (0,1)$ as we stated in equation (3). Here, one is required to incorporate additional information of the reduced form covariance matrix $\Sigma$, allowing to recover $\alpha$ and hence $b_1$ up to sign. Given that the shock variance is fixed to a known number (unity), the result is also described as *absolute* IRFs. Exploiting invertibility of the model $\Sigma = BB'$ yields the following quadratic form:

$$\Gamma'\Sigma^{-1}\Gamma = (\alpha B e_1)'(BB')^{-1}(\alpha B e_1) = \alpha^2.$$

Normalizing $\alpha > 0$, one can back out $b_1 = \Gamma/\alpha = \Gamma/\sqrt{\Gamma'\Sigma^{-1}\Gamma}$ and define the corresponding IRFs as the following function of reduced form parameters:

$$\lambda_{k,i} = e_i' C_k(A)\Gamma/\sqrt{\Gamma'\Sigma^{-1}\Gamma}. \tag{5}$$

At this point, it is worth discussing a key difference between the two definitions of impulse response functions: $\lambda^1_{k,i}$ does not necessarily rely on invertibility of the model, which is the assumption that structural shocks can be recovered as a function of the VAR prediction errors $\varepsilon_t = B^{-1}u_t$. As shown in Plagborg-Møller and Wolf (2021), augmenting the VAR with the external instrument $z_t$ allows for consistent estimation of relative impulse response functions $\lambda^1_{k,i}$, even if invertibility does not hold. Specifically, for $\tilde{y}_t = [z_t, y_t']'$, the resulting "internal instrument VAR" model reads:

$$\tilde{y}_t = \tilde{A}_1\tilde{y}_{t-1} + \tilde{A}_2\tilde{y}_{t-2} + \ldots + \tilde{A}_p\tilde{y}_{t-p} + \tilde{u}_t, \qquad \tilde{u}_t \sim \mathcal{WN}(0, \tilde{\Sigma}), \tag{6}$$

and robust relative IRFs are obtained by $\lambda^1_{k,i} = e_{1+i}' C_k\left(\tilde{A}\right)\tilde{\Gamma}/(e_2'\tilde{\Gamma}_t)$ for $\tilde{\Gamma} = e_1'\text{chol}(\tilde{\Sigma})$.
On the other hand, without invertibility it is no longer possible to identify absolute IRFs $(\lambda_{k,i})$.[1] As we will discuss in the next subsection, the ability to recover the shock up to a known constant will be an important advantage when it comes to studying time-varying impulse response functions that remain comparable over time.

[1] See Plagborg-Møller and Wolf (2022) for detailed analysis on how the shock can be set-identified, however.

## 2.2 Introducing time-varying coefficients

In this paper, we conduct inference for impulse response functions when parameters of the SVAR are slowly evolving over time. Introducing time-varying coefficients into the SVAR reads:

$$y_t = A_{1t}y_{t-1} + A_{2t}y_{t-2} + \ldots + A_{pt}y_{t-p} + B_t\varepsilon_t, \tag{7}$$

where $\mathrm{E}_t(\varepsilon_t) = 0$, $\mathrm{E}_t(\varepsilon_t\varepsilon_t') = I_n$ and $\mathrm{E}_t(u_tu_t') = \Sigma_t = B_tB_t'$. Also, let $\Gamma_t = \mathrm{E}_t(z_tu_t)$. At this point, one approach would be to impose a specific assumption about how time variation is generated, e.g. via a random walk, allowing for likelihood based inference using the Kalman filter (Primiceri, 2005; Paul, 2020). Instead, in this paper we follow a non-parametric approach along the lines of Giraitis et al. (2014, 2018), which assumes a bound on the degree of time variation that can be allowed for in order to conduct valid asymptotic inference via kernel-based estimators:

**Assumption 2.** *Let* $\beta_t = \mathrm{vec}\,(A_t)$ *for* $A_t = [A_{1t}, \ldots, A_{pt}]$, $\sigma_t = \mathrm{vech}(\Sigma_t)$, *and* $\theta_t = [\beta_t', \Gamma_t', \sigma_t']'$. *Then:*

$$\sup_{j\leq h} ||\theta_t - \theta_{t+j}||^2 = O\left(\frac{h}{t}\right), ||\theta_t|| < \infty, \text{ for all } t.$$

Assumption 2 states that the model parameters are bounded and that changes to those parameters are restricted to be small. The rate is assumed, for simplicity, to be of the order $T^{-1}$ but in previous work (see, e.g. Giraitis et al. (2018)), a relaxation to an order given by $T^{-\gamma}$, $0 < \gamma \leq 2$, has been shown to be feasible. Such an order is equivalent to a mild Lipschitz condition on the smoothness of the parameters and is much milder then existing conditions in the time-varying literature. Note that unlike most other existing work, it is not assumed that parameters are smooth deterministic functions of time but, instead, we place a restriction on their differences. For simplicity we assume that parameters are a sequence of deterministic constants and that $\gamma = 1$.

Under Assumption 2, Giraitis et al. (2018) show that the MA($\infty$) representation can be ex-

pressed as:

$$y_t = \sum_{k=0}^{\infty} C_k(A_t) B_t \varepsilon_{t-k} + o(1). \tag{8}$$

Equation (8) states that, under assumption 2, the MA($\infty$) representation of the TVP-SVAR is asymptotically given by that of a fixed-coefficient model, but replacing $A$ and $B$ with their time-varying counterparts. Under the instrumental variables assumption 1, (time-varying) IRFs to a shock of size one standard deviation are then given by:[2]

$$\lambda_{k,i,t} = e_i' C_k(A_t) b_{1t} = e_i' C_k(A_t) \Gamma_t / \sqrt{\Gamma_t' \Sigma_t^{-1} \Gamma_t}, \tag{9}$$

where $b_{1t} = B_t e_1$ is the first column of $B_t$. For the unit shock normalization, the corresponding time-varying IRFs are:

$$\lambda^1_{k,i,t} = e_i' C_k(A_t) \Gamma_t / e_1' \Gamma_t, \tag{10}$$

effectively measuring IRFs to a re-normalized target shock with a standard deviation $b^2_{11,t}$.

It becomes clear that in a time-varying setting, it is difficult to interpret differences in (relative) IRFs under the unit shock standardization. The reason is that if $b_{11,t}$ varies over time, standardizing the impulse responses to increase the first variable by one may require the shock size to adjust. This stands in contrast to $\lambda_{k,i,t}$, for which the shock variance is constant and normalized to unity throughout the sample.

Following the discussion in section 2.1, the computation of $\lambda_{k,i,t}$ relies on shock invertibility and hence, may not always be an option for the researcher. However, when invertibility is a concern, it is still possible to obtain comparable relative IRFs under stronger assumption about the relationship between the external instrument and the target shock (Paul, 2020). Specifically, assuming that $E(z_t \varepsilon_{1t}) = \alpha$ is constant, all the time-variation observed in $\Gamma_{1,t} = \alpha b_{11,t}$ can be attributed to differences in the scale of the shock ($b_{11,t}$). In that case, normalizing the IRFs to increase the first variable by unity at a fixed time point $t_b$ is sufficient to obtain responses to

[2] Alternatively, one might pursue a simulation based approach to obtain a more accurate picture as advocated in Koop et al. (1996), which is based on the exact MA($\infty$) representation.

a constant shock size over time:

$$\bar{\lambda}_{k,i,t} = e_i' C_k(A_t)\Gamma_t / e_1'\Gamma_{t_b}. \tag{11}$$

Specifically, $\Gamma_t / e_1'\Gamma_{t_b} = \alpha b_{1t}/(\alpha b_{11,t_b}) = b_{1t}/b_{11,t_b}$ measures the impact effect of the target shock normalized to have variance $b_{11,t_b}^2$. Hence, although the shock volatility is still unidentified, it is constant throughout the sample and hence comparable across time. Note that this is generally not the case if $\alpha_t$ itself was subject to time-variation.

Summing up our discussion, once time-variation is introduced into the model, a trade-off arises. Assuming invertibility allows to identify the scale of the shock throughout the sample and study (absolute) IRFs $\lambda_{k,i,t}$ in response to a fix shock size. Whenever invertibility does not hold, unit shock IRFs may be a useful alternative. However, they requires stronger assumptions on $\alpha_t$ to obtain (relative) IRFs $\bar{\lambda}_{k,i,t}$ that remain comparable across time. In practice, we therefore recommend to pre-test for shock-invertibility. If there is no evidence against invertibility of the target shock in a given application, studying $\lambda_{k,i,t}$ may be preferable as it requires minimal assumptions on $\alpha_t$. However, if shock invertibility is rejected, $\bar{\lambda}_{k,i,t}$ c an yield informative results under stronger assumptions on $\alpha_t = \alpha$. A pre-test for shock invertibility that can be readily applied in a time-varying set-up is described in Plagborg-Møller and Wolf (2022). Specifically, the testable prediction is that under invertibility of the shock, $z_t$ should not Granger cause $y_t$ in an instrument augmented VAR.

## 2.3 Joint inference for the reduced form parameters

In order to conduct inference for $\lambda_{k,i,t}$ and $\bar{\lambda}_{k,i,t}$, we proceed in two steps. We start deriving the joint asymptotic distribution of kernel-based estimators of the reduced form parameters $A_t$, $\Gamma_t$, $\Sigma_t$ and $\Sigma_{t_b}$. In a second step, we construct confidence sets for the impulse responses either by the Delta method or an inversion of the Anderson and Rubin test statistic.

In order to estimate the reduced form parameters, we cover two popular estimation approaches. The first is the IV-SVAR estimator of Stock and Watson (2012) and Mertens and Ravn (2013),

where $z_t$ is treated as external. This estimator leverages shock invertibility, and hence allows to compute absolute IRFs $\lambda_{k,i,t}$ under minimal assumptions about $\alpha_t$. The second estimator we consider is the internal instrument VAR estimator by Plagborg-Møller and Wolf (2021), treating $z_t$ as internal variable in the VAR model. Here, robustness to invertibility is achieved but relative IRFs $\bar{\lambda}_{k,i,t}$ can only be compared over time under the stronger assumptions of $\alpha_t = \alpha$.

Starting with the IV-SVAR given in equation (7), let $\beta_t = \text{vec}(A_t)$ and $x_t = [y'_{t-1}, y'_{t-2}, \ldots, y'_{t-p}]$. Then, the kernel estimator is given by:

$$\hat{\beta}_t = \left[ I_n \otimes \sum_{j=1}^{T} w_{t,j}\left(H\right) x_j x'_j \right]^{-1} \left[ \sum_{j=1}^{T} w_{t,j}\left(H\right) \text{vec}(x_j y'_j) \right], \tag{12}$$

$$\hat{\Gamma}_t = \frac{1}{H} \sum_{j=1}^{T} w_{t,j}\left(H\right) \hat{u}_j z_j, \tag{13}$$

$$\hat{\Sigma}_t = \frac{1}{H} \sum_{j=1}^{T} w_{t,j}\left(H\right) \hat{u}_j \hat{u}'_j, \tag{14}$$

where $\hat{u}_j = y_j - (I_n \otimes x'_j)\hat{\beta}_t$ and $w_{t,j}\left(H\right) = K(|t-j|/H)$ is a Kernel function to ensure more distant observations get discounted when forming the estimate at time $t$. To establish theoretical properties of the estimator, we make the following two assumptions:

**Assumption 3.** *$\varepsilon_t = (\varepsilon_{1t}, \cdots, \varepsilon_{nt})'$ is an iid process such that $E[\varepsilon_{i1}^4] < \infty$. $z_t$ is a stationary, $\alpha$-mixing process with exponentially declining mixing coefficients, such that $E[z_1^4] < \infty$. Further, $E[y_{i0}^4] < \infty$ for $i = 1, \cdots, n$.*

**Assumption 4.** *$K$ is a non-negative bounded function with a piecewise bounded derivative $\dot{K}(x)$ such that $\int K(x)dx = 1$. If $K$ has unbounded support, we assume in addition that*

$$K(x) \leq C\exp(-cx^2), \quad |\dot{K}(x)| \leq C(1+x^2)^{-1}, \quad x \geq 0, \qquad \text{for some } C > 0,\ c > 0.$$

In this paper, we simply rely on a Gaussian kernel $K_{j,t}(H) \propto \exp\left[-\frac{1}{2}\left(\frac{j-t}{H}\right)^2\right]$, further normalized such that $\sum_j w_{t,j} = H$. In Appendix A we show that:

**Theorem 1.** *[joint asymptotic normality of reduced form parameters in the TVP-IV-SVAR] Under Assumption 2-4 and $H = o(T^{\frac{1}{2}})$ it holds that:*

$$\sqrt{H} \begin{pmatrix} \hat{\beta}_t - \beta_t \\ \hat{\Gamma}_t - \Gamma_t \\ \text{vech}(\hat{\Sigma}_t) - \sigma_t \end{pmatrix} \xrightarrow{d} \mathcal{N}(0, V_{\theta_t}),$$

*for* $V_{\theta_t} = S_t \Pi_{ww,t} S_t'$ *and*

$$S_t = \begin{pmatrix} I_n \otimes \Pi_{x,t}^{-1} & 0 & 0 \\ -\left(I_n \otimes \Pi_{xz,t} \Pi_{x,t}^{-1}\right) & I & 0 \\ 0 & 0 & S_\sigma \end{pmatrix},$$

*for* $\Pi_{x,t} = plim_{T\to\infty} \frac{1}{H} \sum_{j=1}^{T} w_{j,t} x_j x_j'$, $\Pi_{xz,t} = plim_{T\to\infty} \frac{1}{H} \sum_{j=1}^{T} w_{j,t} z_j x_j$, $\Pi_{ww,t} = plim_{T\to\infty} \frac{1}{H} \sum_{j=1}^{T} w_{j,t}^2 \xi_j \xi_j'$, $\xi_j = [\mathrm{vec}(x_j u_j)', (z_j u_j - \Gamma)', \mathrm{vec}\left(u_j' u_j - \Sigma_t\right)']'$, *and* $S_\sigma$ *such that* $\mathrm{vech}(\Sigma_t) = S_\sigma \,\mathrm{vec}(\Sigma_t)$.

It is worth mentioning that the bandwidth $H$ assumed in Theorem 1 is strictly related to the amount of time-variation permitted in Assumption 2. If a different rate is assumed in Assumption 2, say $T^{-\gamma}$ rather than $T^{-1}$, a different $H$ should be used (the more time variation, the smaller the optimal $H$). From a practical point of view, since $\gamma$ is not known, a cross validation approach can be used to select $H$. Towards the end of this section, we describe a simple procedure that targets out-of-sample model fit for IV identified impulse response functions.

A second important comment is that, as mentioned, the way we model non-parametric time variation follows GKY and it is different from the more common approach of assuming that the parameters change as a function (with at least a bounded derivative) of $t/T$ e.g. as in Dahlhaus (1997). Hence, the proof of the results and the conditions on the optimal bandwidth differ from the usual ones.

Equivalent results can be obtained for the reduced form parameters of the time-varying internal instrument VAR. For $\tilde{y}_t = [z_t, y_t']'$, the underlying model reads:

$$\tilde{y}_t = \tilde{A}_{1t}\tilde{y}_{t-1} + \tilde{A}_{2t}\tilde{y}_{t-2} + \ldots + \tilde{A}_{pt}\tilde{y}_{t-p} + \tilde{u}_t, \qquad \tilde{u}_t \sim (0, \tilde{\Sigma}_t), \tag{15}$$

where $\tilde{\beta}_t = \mathrm{vec}\left(\left[\tilde{A}_{1t}, \ldots, \tilde{A}_{pt}\right]\right)$ and $\tilde{P}_t = \mathrm{chol}(\tilde{\Sigma}_t)$ is the Cholesky decomposition such that $\tilde{P}_t \tilde{P}_t' = \tilde{\Sigma}_t$. As discussed above, the main object of interest based on the internal instrument VAR is $\bar{\lambda}_{k,i,t} = e_{1+i}' C_k\left(\tilde{A}_t\right) \tilde{P}_{\bullet 1,t} / (e_2' \tilde{P}_{\bullet 1,t_b})$, that is the relative IRF standardized to increase the first variable by unit on date $t_b$. We start with results for the following estimator of time

$t$ reduced form coefficients:

$$\hat{\tilde{\beta}}_t = \left[ I_{n+1} \otimes \sum_{j=1}^{T} w_{t,j}\left(H\right) \tilde{x}_j \tilde{x}_j' \right]^{-1} \left[ \sum_{j=1}^{T} w_{t,j}\left(H\right) \mathrm{vec}(\tilde{x}_j \tilde{y}_j') \right] \tag{16}$$

$$\hat{\tilde{\Sigma}}_t = H^{-1} \sum_{j=1}^{T} w_{t,j}(H) \hat{\tilde{u}}_j \hat{\tilde{u}}_j', \tag{17}$$

where $\hat{\tilde{u}}_j = \tilde{y}_j - (I_{n+1} \otimes x_j') \hat{\tilde{\beta}}_t$. Joint asymptotic normality between the reduced form parameters are then given as follows:

**Theorem 2.** *[joint asymptotic normality of reduced form parameters in the TVP internal instrument VAR.] Under Assumption 2-4 and $H = o(T^{\frac{1}{2}})$: define $\tilde{\Pi}_{x,t} = plim_{T\to\infty} \frac{1}{H} \sum_{j=1}^{T} w_{j,t} \tilde{x}_j \tilde{x}_j'$, $\tilde{\Pi}_{ww,t} = plim_{T\to\infty} \frac{1}{H} \sum_{j=1}^{T} w_{j,t}^2 \tilde{x}_j \tilde{x}_j', \Pi_{uu,uu,t} = plim_{T\to\infty} \frac{1}{H} \sum_{j=1}^{T} w_{j,t} \mathrm{vec}(\tilde{u}_j \tilde{u}_j') \mathrm{vec}(\tilde{u}_j \tilde{u}_j')'$ , $\tilde{\sigma}_t = \mathrm{vech}(\tilde{\Sigma}_t)$ and $L_n$ be the $n(n+1)/2 \times n^2$ elimination matrix such that $\mathrm{vech}(A) = L_n \mathrm{vec}(A)$. Then, the estimators $\hat{\tilde{\beta}}_t$ and $\hat{\tilde{\sigma}}_t$ are asymptotically independent and:*

$$\sqrt{H}\left(\hat{\tilde{\beta}}_t - \tilde{\beta}_t\right) \xrightarrow{d} \mathcal{N}\left(0, \tilde{\Sigma}_t \otimes \left(\tilde{\Pi}_{x,t}\right)^{-1} \tilde{\Pi}_{ww,t} \left(\tilde{\Pi}_{x,t}\right)^{-1}\right),$$

$$\sqrt{H}\left(\hat{\tilde{\sigma}}_t - \tilde{\sigma}_t\right) \xrightarrow{d} \mathcal{N}\left(0, L_{n+1}\Pi_{uu,uu,t}L_{n+1}' - \tilde{\sigma}_t\tilde{\sigma}_t'\right).$$

Under an additional normality assumption for the errors, the asymptotic variance of $\hat{\tilde{\sigma}}_t$ further reduces to $2D_{n+1}^{+}\left(\tilde{\Sigma}_t \otimes \tilde{\Pi}_{uu,t}\right) D_{n+1}^{+\prime}$, where $\tilde{\Pi}_{uu,t} = \mathrm{plim}_{T\to\infty} \frac{1}{H} \sum_{j=1}^{T} w_{j,t}^2 \tilde{u}_j \tilde{u}_j'$ and $D_{n+1}^{+} = (D_{n+1}'D_{n+1})^{-1}D_{n+1}'$ for $D_{n+1}$ the duplication matrix such that $\mathrm{vec}(\tilde{\Sigma}_t) = D_{n+1}\,\mathrm{vech}(\tilde{\Sigma}_t)$.

Given that estimates of $\bar{\lambda}_{k,i,t}$ are based on reduced form parameters at time $t$ and $t_b$, the construction of corresponding confidence sets requires an expression for their joint distribution. This is particularly relevant when $t_b$ and $t$ are close, and estimates are highly correlated by construction. Given that VAR slope and covariance parameters are asymptotically uncorrelated, and given that only covariance estimates of $t_b$ are used to construct $\bar{\lambda}_{k,i,t}$, it's sufficient to focus on the joint distribution of $\hat{\tilde{\sigma}}_t$ and $\hat{\tilde{\sigma}}_{t_b}$. The following Corollary gives their asymptotic joint distribution.

**Corollary 3.** *Let Assumptions 2-4 hold and $H = o(T^{\frac{1}{2}})$. Define $w_t(H) = [w_{t,1}(H), \ldots, w_{t,T}(H)]'$ and $\tilde{\sigma}_{t,t_b} = \mathrm{vech}(\tilde{\Sigma}_{t,t_b})$. Let $\Pi_{uu,uu,t,t_b} = plim_{T\to\infty} \frac{1}{H} \sum_{j=1}^{T} \mathrm{vec}(\tilde{\xi}_{wj}\tilde{\xi}_{wj}') \mathrm{vec}(\tilde{\xi}_{wj}\tilde{\xi}_{wj}')'$ for*

$$\tilde{\xi}_{wj} = \left[ w_{t,j}^{1/2}(H)\left(\tilde{y}_j - \tilde{x}_j\tilde{\Theta}_t\right), w_{t_b,j}^{1/2}(H)\left(\tilde{y}_j - \tilde{x}_j\tilde{\Theta}_{t_b}\right)\right].$$

*Under these definitions, it follows that:*

$$\sqrt{H}\left(\hat{\tilde{\sigma}}_{t,t_b} - \tilde{\sigma}_{t,t_b}\right) \xrightarrow{d} \mathcal{N}\left(0, L_{2(n+1)}\Pi_{uu,uu,t,t_b}L'_{2(n+1)} - \tilde{\sigma}_{t,t_b}\tilde{\sigma}'_{t,t_b}\right)$$

As above, under normality assumption of the errors, the asymptotic variance of $\hat{\tilde{\sigma}}_t$ further simplifies to $2D^{+}_{2(n+1)}\left(\tilde{\Sigma}_{t,t_b} \otimes \tilde{\Pi}_{uu,t,t_b}\right) D^{+'}_{2(n+1)}$, where $\tilde{\Pi}_{uu,t,t_b} = \text{plim}_{T\to\infty}\frac{1}{H}\sum_{j=1}^{T}\tilde{U}_2\tilde{U}'_2$ for

$$\tilde{U}_2 = \left[w_t(H) \otimes (\tilde{Y} - \tilde{X}\tilde{\Theta}_t), w_{t_b}(H) \otimes (\tilde{Y} - \tilde{X}\tilde{\Theta}_{t_b})\right]$$

and $D^{+}_{2(n+1)} = (D'_{2(n+1)}D_{2(n+1)})^{-1}D'_{2(n+1)}$ for $D_{2(n+1)}$ the duplication matrix such that $\text{vec}(\tilde{\Sigma}_{t,t_b}) = D_{2(n+1)}\,\text{vech}(\tilde{\Sigma}_{t,t_b})$.

## 2.4 Inference for impulse response functions

Based on asymptotic results for the estimators of the reduced form parameters, we can rely on standard methods to construct confidence sets for the object of interest, that are the estimates of time-varying structural impulse response functions:

$$\hat{\lambda}_{k,i,t} = e'_i C_k(\hat{A}_t)\hat{\Gamma}_t / \sqrt{\hat{\Gamma}'_t\hat{\Sigma}_t^{-1}\hat{\Gamma}_t},$$
$$\hat{\tilde{\lambda}}_{k,i,t} = e'_{1+i} C_k\left(\hat{\tilde{A}}_t\right)\hat{\tilde{P}}_{\bullet 1,t}/(e'_2\hat{\tilde{P}}_{\bullet 1,t_b}),$$

where $\hat{A}_t$, $\hat{\Gamma}$ and $\hat{\Sigma}_t$ are based on the IV-SVAR, while $\hat{\tilde{A}}_t$, $\hat{\tilde{P}}_{\bullet 1,t}$ and $\hat{\tilde{P}}_{\bullet 1,t_b}$ are based on the internal instrument VAR.

In this paper, we discuss two approaches to construct appropriate confidence sets, either via the classical Delta method or an inversion of the Anderson Rubin (AR) test statistic as in Montiel-Olea et al. (2021). The latter fixes $\alpha$ under the null hypothesis and hence remains valid even under asymptotically weak instruments, that is if $\alpha \to 0$.

Starting with the Delta Method, an application yields that $\sqrt{H}\left(\hat{\lambda}_{k,t} - \lambda_{k,t}\right) \xrightarrow{d} \mathcal{N}(0, \Omega_{k,t})$, where $\Omega_{k,t} = J_k(\beta_t, \Gamma_t, \sigma_t)\, V_{\theta_t} J_k(\beta_t, \Gamma_t, \sigma_t)'$. Here, $V_{\theta_t}$ denotes the joint distribution of the IV-SVAR reduced form parameters which we give in Theorem 1. Furthermore, $J_k(\beta_t, \Gamma_t, \sigma_t)$ denotes the derivative of $\lambda_{k,t}$ with respect to the reduced form parameters. Similarly, for relative

IRFs we get $\sqrt{H}\left(\hat{\bar{\lambda}}_{k,t} - \bar{\lambda}_{k,t}\right) \xrightarrow{d} \mathcal{N}\left(0, \bar{\Omega}_{k,t}\right)$ where $\bar{\Omega}_{k,t} = \bar{J}_k\left(\tilde{\beta}_t, \tilde{\sigma}_{t,t_b}\right) \bar{V}_{\theta_t} \bar{J}_k\left(\tilde{\beta}_t, \tilde{\sigma}_{t,t_b}\right)'$ for $\bar{V}_{\theta_t}$ being elements of the joint asymptotic covariance matrix stated in Theorem 2 and Corollary 3, and $\bar{J}_k()$ is the gradient of $\bar{\Omega}_{k,t}$ with respect to the reduced form parameters. Analytical formulas for both gradients are derived in Appendix B.

As documented in Montiel-Olea et al. (2021) for relative IRFs, empirical coverage rates of the Delta method can deteriorate quickly when the instrument is only weakly correlated with the target shock. Therefore, within $\bar{\lambda}_{k,t}$ it is important to assess whether there are changes in the instrument strength over time, and normalize the relative IRFs at times $t_b$ when $z_t$ is a strong predictor for the first reduced form residual. Therefore, we suggest to run a simple Wald test for the element in the denominator of $\bar{\lambda}_{k,t}$, that is $e_2'\hat{\tilde{P}}_{\bullet 1,t_b}$. Choosing a time point $t_b$ for which the Wald statistic is particularly high makes inference about $\bar{\lambda}_{k,t}$ less prone to the weak instrument problem.

Given that no such strategy can be pursued for $\lambda_{i,t}$, we also consider weak identification robust confidence sets. In the following, consider the $(n+1)\times 1$ vectors $L$ and the $(n+2)\times 1$ vector $\bar{L}_{k,t}$:

$$L_{k,t} = \begin{pmatrix} C_k(A_t)\Gamma_t \\ \sqrt{\Gamma_t'\Sigma_t^{-1}\Gamma_t} \end{pmatrix}, \qquad \bar{L}_{k,t} = \begin{pmatrix} C_k\left(\tilde{A}_t\right)\tilde{P}_{\bullet 1,t} \\ e_2'\tilde{P}_{\bullet 1,t_b} \end{pmatrix},$$

for which it holds that $\lambda_{k,i,t} = (e_i'L_{k,t})/(e_{n+1}'L_{k,t})$ and $\bar{\lambda}_{k,i,t} = (e_{1+i}'\bar{L}_{k,t})/(e_{n+2}'\bar{L}_{k,t})$. To derive the AR confidence set, first note that an application of the Delta Method implies that $\sqrt{H}\left(\hat{L}_{k,t} - L_{k,t}\right) \xrightarrow{d} \mathcal{N}(0, \Omega_{k,t}^L)$ where $\Omega_{k,t}^L$ depends on the covariance matrix given in Theorem 1 and the gradient of $L_{k,t}$ with respect to the reduced form parameters. Similarly, a statement can be obtained for $\sqrt{H}\left(\hat{\bar{L}}_{k,t} - \bar{L}_{k,t}\right) \xrightarrow{d} \mathcal{N}(0, \Omega_{k,t}^{\bar{L}})$ based on the reduced form results for the internal VAR estimator.

Without loss of generality, let us focus on the confidence set for $\lambda_{k,i,t}$. The null hypothesis $\lambda_{k,i,t} = \lambda_0$ implies $e_i'L_{k,t} - \lambda_0 e_{n+1}'L_{k,t} = 0$, a linear restriction on $L_{k,t}$ (see also Fieller (1944)). Following Montiel-Olea et al. (2021), a Wald Test statistic can be set up as $q(\lambda_0) =$

$\frac{H(e_i'\hat{L}_{k,t}-\lambda_0 e_{n+1}'\hat{L}_{k,t})^2}{\hat{\omega}_{ii}-2\lambda_0\hat{\omega}_{i,n+1}+\lambda_0^2\hat{\omega}_{n+1,n+1}}$ where $\hat{\omega}_{ij}$ is the $ij$th element of $\hat{\Omega}_{k,t}$. Further inversion yields the AR confidence set of coverage $1-a$, given by $\text{CS}^{AR}\{\lambda_{k,i,t}|q(\lambda_{k,i,t}) \leq \chi^2_{1,1-a}\}$. The inequality $q(\lambda_{k,i,t}) \leq \chi^2_{1,1-a}$ is quadratic in $\lambda_{k,i,t}$ and can be solved in closed form. For details, including the gradients necessary to obtain $\Omega^L_{k,t}$, we refer to Appendix B.

A few properties are worth mentioning at this point. First, even in a weak instrument case where $\alpha_H = a/\sqrt{H}$ for some fixed $a$, the AR CS remains valid. The reason is that the Wald statistic fixes $\lambda_{k,i,t}$ under the null hypothesis and hence does not require consistent estimates thereof for its validity. Second, the CS may be infinite, but is guaranteed to be finite whenever a Wald test for $e_{n+1}'L_{k,t} = 0$ can be rejected at the $1-a$ confidence level. Third, in the strong instrument case, Montiel-Olea et al. (2021) prove that the AR confidence set converges to Delta Method implied confidence intervals.

## 2.5 Bandwidth selection

Finally, the user is required to set a bandwidth $H$ at the time of estimating the model parameters. We acknowledge that there are several ways this can be done. First, similar to a Bayesian approach, one might take on off-model information about how much time-variation is reasonable to see in Impulse Response Functions over a certain time span, and select a bandwidth accordingly.

In this paper, however, we pursue a purely data-driven approach selecting a bandwidth $H$ that provides the best out-of-sample model fit. Acknowledging that IRFs are essentially conditional forecasts, we propose to evaluate the models out-of-sample predictive performance comparing data realizations $y_{t+h}$ to conditional predictions $\hat{y}_{t+h}(H) = E_t[y_{t+h}|\tilde{y}_1, \ldots, \tilde{y}_t, z_{t+1}, \ldots, z_{t+h}]$. Just like the IRFs identified by IV, this is a function of both the VAR slope parameters $\tilde{A}_t$ and covariance matrix $\tilde{\Sigma}_t$ in the instrument-augmented VAR (Waggoner and Zha, 2003).[3]

[3]Note that to evaluate the conditional expectations based on the information set up to time $t$, we rely on modified estimators $\hat{\tilde{A}}_{t|t}(H)$ and $\hat{\tilde{\Sigma}}_{t|t}(H)$ defined as in equation (16)-(17), but based on truncated kernels for which we set $w_{t,j}(H) = 0, j > t$.

Based on those forecasts, we propose the following objective function to choose the bandwidth:

$$\min_{H} \sum_{i=1}^{n} \mathrm{w}_i \sum_{t=T_s}^{T-h} \sum_{h=1}^{h_m} \left(y_{i,t+h} - \hat{y}_{i,t+h}(H)\right)^2,$$

where $T_s$ is the time in the sample where the pseudo out-of-sample exercise starts, $h_m$ is the maximum forecast horizon to be included in the evaluation, and $\mathrm{w}_i$ is a variable specific weight to account for differences in scale. For the latter, we simply use the inverse variance of AR(1) residuals in each variable.

# 3 Monte Carlo Simulations

In the following, we study the finite sample properties of the proposed inference procedure for time-varying impulse response function. As we expect, the performance of confidence sets will depend on the effective sample size, the speed of time-variation underlying the SVAR coefficients and the instrument strength. Overall, our findings suggest that the asymptotic theory provides a reasonable approximation in finite samples.

## 3.1 Data Generating Process

In order to simulate from a practically relevant Data Generating Process (DGP), we follow Montiel-Olea et al. (2021) and calibrate a time-varying parameter VAR model based on actual macroeconomic data. Building on the oil market literature (Kilian, 2008, 2009), we fit VAR kernel estimators for a monthly trivariate dataset of size $T = 377$. The dataset includes the change in (log) global crude oil production, an index for real economic activity and the log of real oil price. A total of three lags are considered, yielding the following estimated structural VAR model:

$$y_t = \hat{A}_t(H)x_t + \mathrm{chol}\left(\hat{\Sigma}_t^{1/2}(H)\right) Q\varepsilon_t, \qquad \varepsilon_t \sim N(0, I),$$

where the bandwidth is set to $H = 100$ and $Q$ is a rotation matrix set in a way that $b_1 \propto [1, 1, -1]'$ resembles a supply shock in a fixed parameter model ($H \to \infty$). Finally,

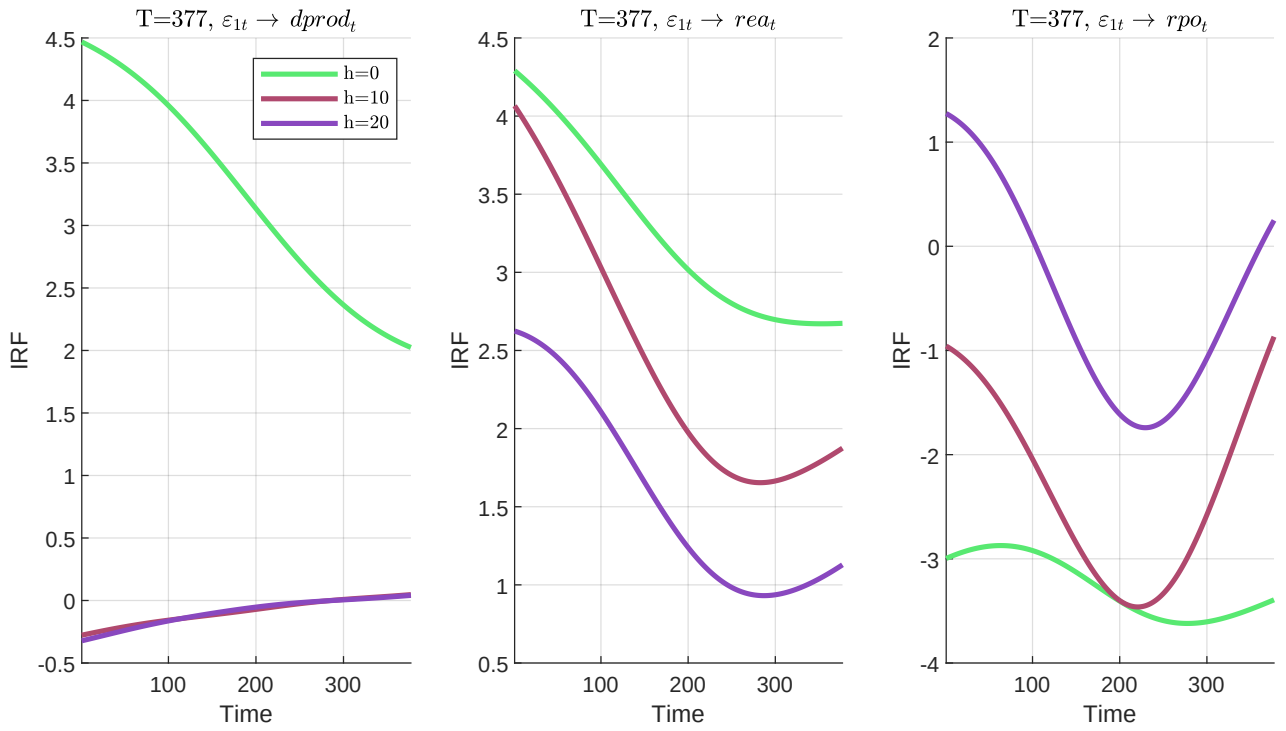


Figure 1: True impulse response functions $\lambda^{\sigma}_{h,i,t}$.

an instrument $z_t$ is generated by the following measurement error equation:

$$z_t = \phi_z \varepsilon_{1t} + \sigma_z \eta_t, \qquad \eta_t \sim \mathcal{N}(0,1),$$

where we consider $\theta^{strong} = \{\phi_z = 0.86, \sigma_z = 0.06\}$ and $\theta^{weak} = \{\phi_z = 0.48, \sigma_z = 0.71\}$ following two parameter constellations proposed in Montiel-Olea et al. (2021) that yield a strong- and weak instrument for relative IRFs $\lambda^{1}_{k,i}$ (unit shocks) in a fixed parameter case. However, as we will document, these parameter constellations don't necessarily translate to strong and weak instrument dynamics for absolute IRFs $\lambda_{k,i}$, which generally rely on more information coming from $\Gamma$ and $\Sigma$ exploiting underlying shock invertibility assumption.

The true impulse response functions for the resulting DGP are given in Figure 1 for three horizons: $h = 0, 10, 20$. As visible from the chart, they display substantial time-variation over the sample period. For example, the impact effect ($h = 0$) of the supply shock on the first variable ($dprod_t$) halves over the sample, while at $h = 10$ and $h = 20$ the sign changes at around $t = 300$. Similar patterns for magnitudes and signs are present in the IRFs of the other two variables, although with an increased persistence.

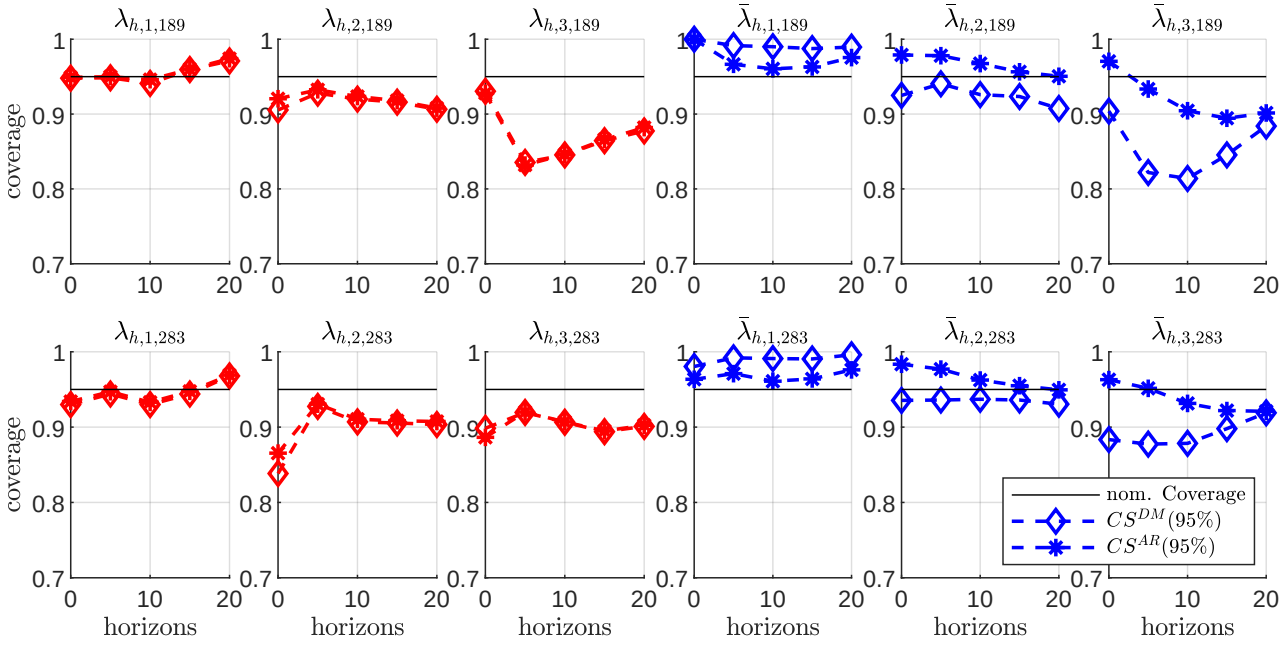


Figure 2: Estimated empirical coverage at 95% confidence level obtained for $\lambda_{h,i,t}$ (red) and $\bar{\lambda}_{h,i,t}$ (blue) at $t = 1/4T = 180$ (first row) and $t = 3/4T = 283$ (second row). $\theta^{strong} = \{\phi_z = 0.86, \sigma_z = 0.06\}$ and $H$ is known. Confidence Sets (CS) based on the Delta Method (DM) are highlighted by diamonds, while Anderson Rubin confidence sets (AR) by stars.

## 3.2 Empirical coverage

We proceed simulating a total of 2000 datasets from the DGP, each of sample size $T = 377$. Empirical coverage at 95% confidence level is then computed for $\lambda_{h,i,t}$ using the TVP-IV-SVAR estimator, and for $\bar{\lambda}_{h,i,t}$ based on the internal instrument TVP-VAR estimator.[4] While we assume the lag length to be known during the Monte Carlo exercise, we explore empirical coverage under both the true bandwidth and the simple data-driven selection method described in section 2.4. Finally, for increased readability, we focus on the empirical coverage at only two points of time, $t = 1/4 \times T$ and $t = 3/4 \times T$.

Figure 2 shows simulation results under the strong instrument parameters setting and known bandwidth. Focusing on $\lambda_{h,i,t}$ (in red) we document that the Delta Method (DM) and Anderson Rubin (AR) confidence sets largely coincide, suggesting that the instrument remains strong. Empirical coverage for $\lambda_{h,1,t}$ is very close to the nominal confidence level (95%), while that of $\lambda_{h,2,t}$ and $\lambda_{h,3,t}$ is still reasonable with values of more than 90% at most horizons. The worst

[4]For the latter, IRFs are re-standardized to increase the first variable by one at a fixed time point $t_b = T/2$.

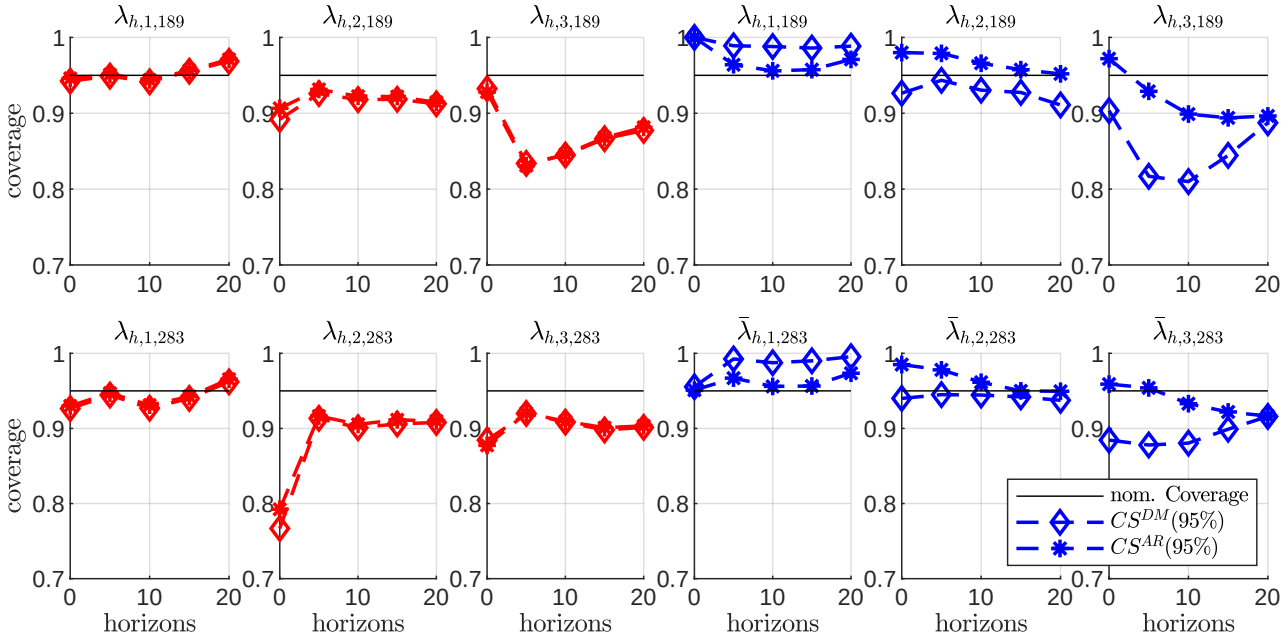


Figure 3: Estimated empirical coverage at 95% confidence level obtained for $\lambda_{h,i,t}$ (red) and $\bar{\lambda}_{h,i,t}$ (blue) at $t = 1/4T = 180$ (first row) and $t = 3/4T = 283$ (second row). $\theta^{strong} = \{\phi_z = 0.86, \sigma_z = 0.06\}$ and $H$ is estimated. Confidence Sets (CS) based on the Delta Method (DM) are highlighted by diamonds, while Anderson Rubin confidence sets (AR) by stars.

performance we document is for $\bar{\lambda}_{h,3,t}$ at $t = 1/4 \times T = 189$ for horizons $h = 5$ and $h = 10$, where coverage is about 85%.

With respect to $\bar{\lambda}_{h,i,t}$ (in blue), we document that the DM- and AR confidence sets provide generally different empirical coverage. This suggests that unlike for $\lambda_{h,i,t}$ a weak instrument problem has likely arisen. This is not unsurprising, given that relative to constant parameter setup, the kernel estimator leverages information only locally and hence is subject to a lower effective sample size. Given the resulting weak instrument problem, the Delta Method generally provides worse coverage than the AR CS. For the IRFs of the first variable, $\bar{\lambda}_{h,1,t}$, the DM provides considerable over-coverage throughout horizons, while the AR confidence sets are close to the nominal level. On the other hand, for $\bar{\lambda}_{h,2,t}$ and $\bar{\lambda}_{h,3,t}$ the DM provides coverage that is generally too low, whereas coverage by the weak-IV robust confidence sets performs better and stays at or above 90%.

Figure 3 shows equivalent simulations when $H$ is chosen by the data-driven method we describe in 2.5. We document very similar coverage rates with exception of $\lambda_{h,2,283}$ where we see some

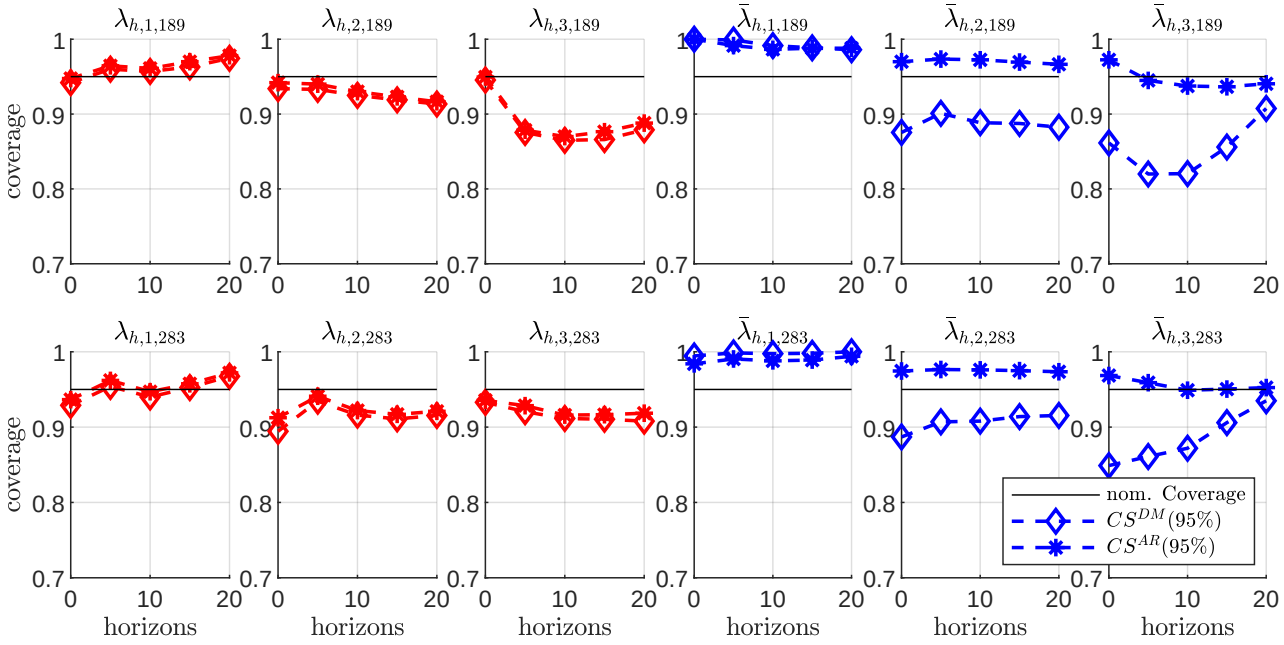


Figure 4: Estimated empirical coverage at 95% confidence level obtained for $\lambda_{h,i,t}$ (red) and $\bar{\lambda}_{h,i,t}$ (blue) at $t = 1/4T = 180$ (first row) and $t = 3/4T = 283$ (second row). $\theta^{weak} = \{\phi_z = 0.48, \sigma_z = 0.71\}$. Confidence Sets (CS) based on the Delta Method (DM) are highlighted by diamonds, while Anderson Rubin confidence sets (AR) by stars.

deterioration to levels of about 80%.

For the parameter constellation $\theta^{weak}$, results obtained under the known bandwidth are reported in Figure 4. Starting with $\lambda_{h,i,t}$ (in red), we find very similar results than reported previously in the strong instrument case. Generally, the coverage remains satisfactory near the nominal value of 95%. The worst coverage is obtained for $\lambda_{h,3,t}$ with rates slightly above 85%. Interestingly, note that only marginal differences arise between the DM- and AR confidence sets, suggesting that the weak instrument problem created by Montiel-Olea et al. (2021) for relative IRFs does not translate to absolute IRFs, despite the lower effective sample size.

With respect to $\bar{\lambda}_{h,i,t}$ (blue), the performance of both confidence sets deteriorates as one would expect when the instruments become even weaker. Still, the AR confidence sets perform better than the DM, remaining closer to the nominal 95% level. However, one starts to observe some over-coverage, particularly for $\bar{\lambda}_{h,1,t}$.

Similar to the first parameter constellation for the instrument, choosing $H$ by a data-driven method yields broadly similar results, with only minor deterioration in coverage rates for some

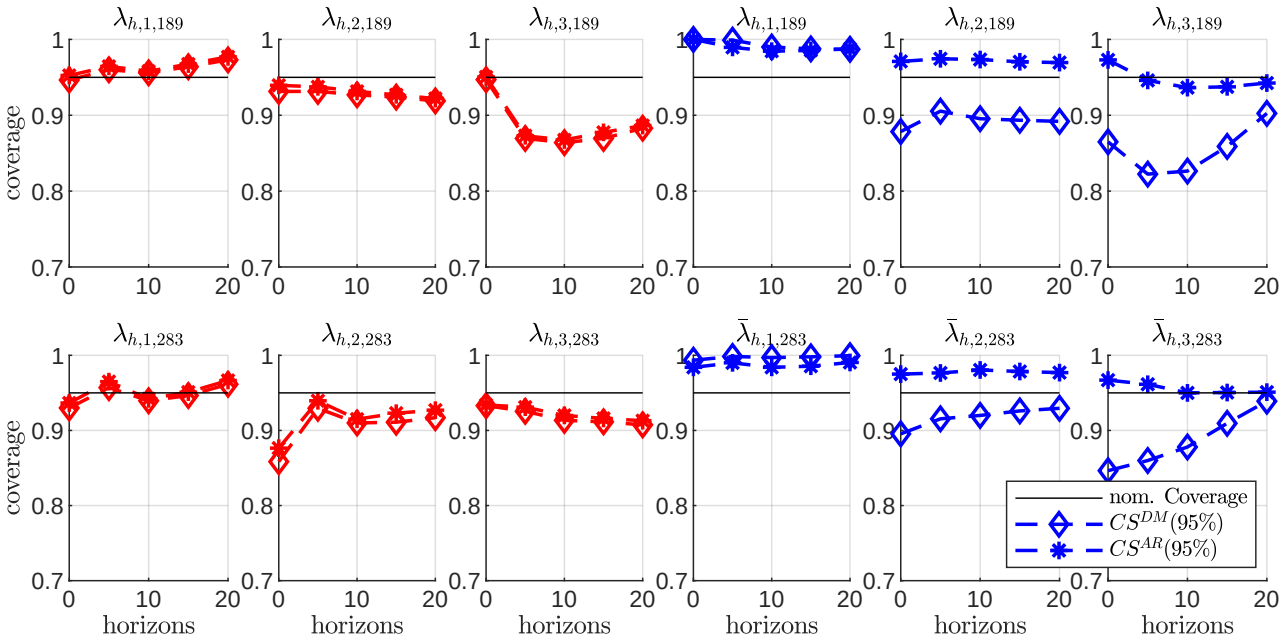


Figure 5: Estimated empirical coverage at 95% confidence level obtained for $\lambda_{h,i,t}$ (red) and $\bar{\lambda}_{h,i,t}$ (blue) at $t = 1/4T = 180$ (first row) and $t = 3/4T = 283$ (second row). $\theta^{weak} = \{\phi_z = 0.48, \sigma_z = 0.71\}$. and $H$ is estimated. Confidence Sets (CS) based on the Delta Method (DM) are highlighted by diamonds, while Anderson Rubin confidence sets (AR) by stars.

of the IRFs (see Figure 5).

In Appendix C, we provide supplementary Monte Carlo results obtained in a large sample environment. To obtain the larger model, we interpolate coefficients linearly to obtain an equivalent shape in the time-varying coefficients, but spread out over a larger sample. We choose $T = 15 \times 377 = 5665$, and let the kernel bandwidth for estimation increase by $H = \sqrt{15} \times 100$. Our findings suggest that estimated empirical coverage rates get very close to the nominal size as one would expect in large samples.

# 4 The effects of monetary policy on financial variables in the UK

In the following, we apply the methodology to study the time varying effect of monetary policy on financial variables in the United Kingdom (UK).

## 4.1 Data and identification strategy

Our analysis is based on VAR models that include a series of key interest rates and financial variables available on a daily frequency. Specifically, we include the 3 Month Overnight Index Swap rate (OIS), the 2- and 10 year government bond (Gilt) yields, the sterling exchange rate index (ERI), the FTSE all share index, non-financial corporate bond yields and spreads, as well as the 5 year and the 5-year-5-year forward inflation expectation rate based on inflation linked swaps (ILS).[5] The latter measure (average) inflation compensation demanded by market participants for the next five years, and the five years that follow. We include interest rates and spreads in levels, while stock prices and the exchange rate are included in 100 times log levels. For both VAR models, we use data from October 2004 until March 2023, which sum up to more than 4800 observations.

To identify the dynamic effects of monetary policy with our VARs, we follow a high frequency identification approach (Gertler and Karadi, 2015; Altavilla et al., 2019; Swanson, 2021). Specifically, we use surprises in the 2 year intra-day Gilt yield around a narrow windows of Bank of England policy announcements made available in Braun et al. (2023). The idea behind using theses as external instruments is that variation in a narrow window of policy announcements can be credibly thought to be exogenous. Systematic reactions of the central bank should be mostly priced-in just before the announcement, and the small intra-day window ensures that most of the variation is not substantially driven by other macroeconomic news. We rely on surprises based on MPC decision announcements and the publication of Monetary Policy Reports (previously Inflation Reports).[6] Our choice for the two year rate is motivated by the fact that we aim to identify an *average* monetary policy shock. In the light of a changing policy mix that occurs over the sample, surprises in the two year rate are likely to reflect a mix

[5]The ERI is provided by the Bank of England (BOE) and measures the overall change in the trade-weighted exchange value of sterling. Non-financial corporate bond yields and (option-implied) spreads are from ICE BofA tracking the performance of sterling denominated, non-financial corporate debt with at least one year maturity. Finally, the 5y and 5y5y forward rate are based on inflation swaps and obtained from Bloomberg. The latter is approximated by 2 times the 10-year rate minus the 5-year rate.

[6]For details on the construction of the high-frequency surprise see Braun et al. (2023).

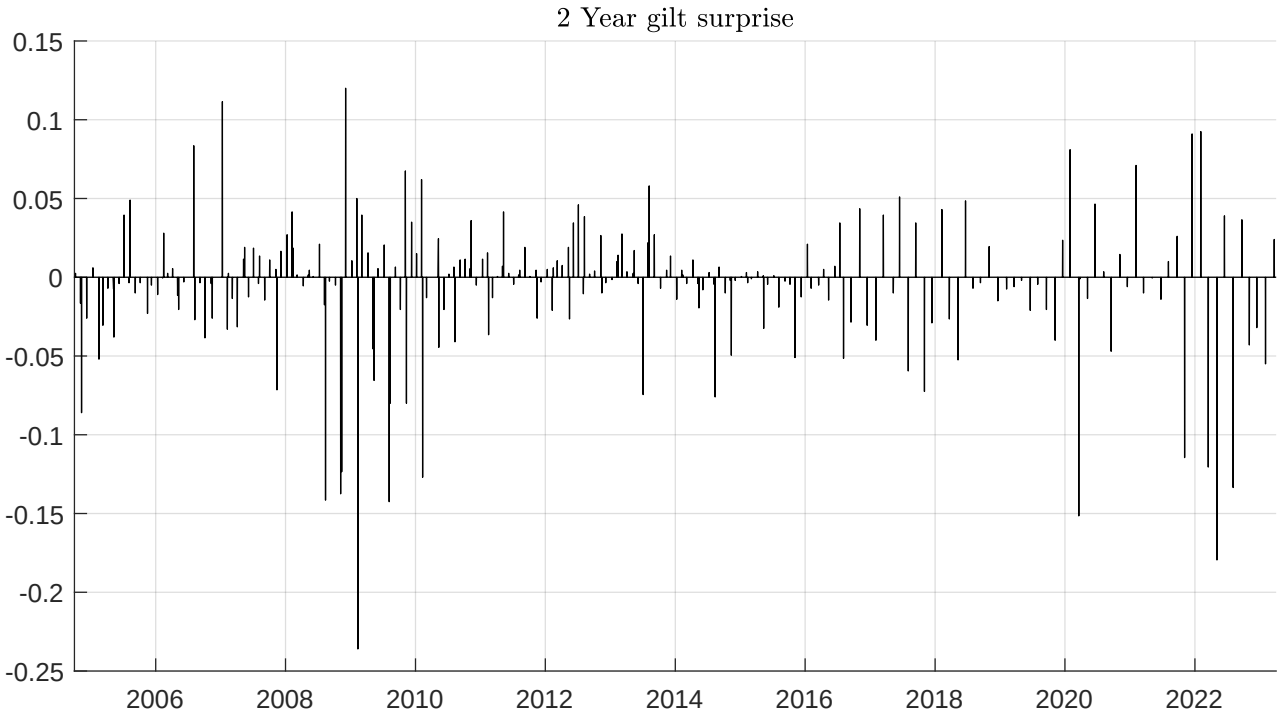


Figure 6: 2 year interest rate surprise series used as external instrument for average monetary policy conducted by the Bank of England (right).

of conventional an unconventional policies.

To allow for sufficient dynamics, we include a total of $p = 25$ lags, which corresponds to 5 weeks of information in a 5-day business week. As for the bandwidth, we set $H = 866$, which we obtain by minimizing squared out-of-sample (conditional) forecast errors up to $h = 25$ between May 2012 and March 2020 (see section 2.5). Note that since the instrument is only non-zero on monetary policy event days, we restrict the model evaluation criteria to include only conditional forecasts made at times that align with those dates (92 days in total).[7]

Figure 6 provides a plot of the external instrument series. There are a total of 234 surprises over the sample period, with an average of 4.5bp variation within those. There is considerable fluctuation throughout the sample, including the period when the BOE Bank Rate was at the zero lower bound. This indicates that as intended, the IV likely captures variation from both conventional and unconventional monetary policy.

[7]Specifically, one-step ahead conditional forecast $E_t[y_{t+1}|\tilde{y}_1, \ldots, \tilde{y}_t, z_{t+1}]$ by the IV augmented VAR are aligned with monetary policy event days such that $z_{t+1}$ contains the monetary policy surprise.

## 4.2 Testing for invertibility and local instrument strength

Before reporting empirical results, we run (local) tests of invertibility to assess if the monetary policy shock can be backed out up to scale within the daily VAR model. To this end, we follow Plagborg-Møller and Wolf (2022) and run Granger causality tests for the null hypothesis that $z_t$ does not predict $y_t$ in the internal instrument VAR model. If the high frequency surprise adds additional predictive power to the model, it can be interpreted as evidence that a VAR monetary policy shock is unlikely to be recoverable as a function of current and past data of that model.

Table 1: Granger causality test results computed at different points of time for the null hypothesis that $z_t$ does not predict $y_t$ in a VAR for $\tilde{y}_t = [z_t, y_t']'$.

| date $t$ | 10/04 | 05/07 | 11/09 | 06/12 | 12/14 | 06/17 | 01/20 | 07/22 | 03/23 |
|---|---|---|---|---|---|---|---|---|---|
| F Statistic | 1.31 | 1.65 | 1.77 | 1.8 | 1.59 | 1.79 | 2.26 | 2.3 | 2.28 |
| p-val | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Table 1 shows F statistics with corresponding $p$ values for the Granger causality tests, based on evolving parameter estimates obtained throughout the sample. The null of no Granger causality is rejected at all times, implying additional predictive ability of the instrument. This can be interpreted as strong statistical evidence against invertibility of the monetary policy shock.

Within our empirical application, this means that we are unable to recover the scale of the monetary policy shock and compute absolute impulse response functions $\lambda_{h,i,t}$. Going forward, unless otherwise noted, we focus on relative IRFs ($\bar{\lambda}_{h,i,t}$) based on the non-invertibility robust internal instrument VAR. Hence, we assume the stronger assumption $\mathrm{E}[z_t \varepsilon_{1t}] = \alpha_t = \alpha$ which is required to interpret $\bar{\lambda}_{h,i,t}$ as IRFs to a monetary policy shock of comparable magnitude over time.

Recall that the computation of $\bar{\lambda}_{h,i,t} = e_{1+i}' C_k \left(\tilde{A}_t\right) \tilde{P}_{\bullet 1,t} / (e_2' \tilde{P}_{\bullet 1,t_b})$ requires normalizing at time $t_b$. As discussed in section 2.4, a weak instrument problem can be mitigated by setting $t_b$ to a time where $z_t$ is strongly correlated with the first VAR prediction error. To aid this choice,

Table 2 provides Wald test statistics for the null hypothesis that $e_2'\tilde{P}_{\bullet 1,t} = 0$. Clearly, our results suggest that the instrument is strong throughout the sample. We set it to June 2017 where we obtain the highest test statistic in our sample, and pre-multiply the IRFs by $-0.0225$, standardizing the shock to yield a roughly 2.5bp point cut in June 2017. This shock size exactly matches that of a unit standard deviation shock obtained under invertibility ($\lambda_{h,i,t}$).

Table 2: Wald test statistic for the null hypothesis that $e_2'\tilde{P}_{\bullet 1,t} = 0$.

| date $t$ | 10/04 | 05/07 | 11/09 | 06/12 | 12/14 | 06/17 | 01/20 | 07/22 | 03/23 |
|---|---|---|---|---|---|---|---|---|---|
| Wald Statistic | 78.55 | 80.42 | 79.79 | 92.03 | 105.28 | 111.07 | 91.44 | 61.98 | 56.59 |

## 4.3 Main Empirical Results

In the following, we discuss the main empirical results: estimates of time varying impulse response functions $\bar{\lambda}_{h,i,t}$ to an expansionary policy shock provided in Figures 7 and 8. For ease of exposition, we choose to plot IRFs at three fixed horizons, that is on impact, after 2 weeks and after 4 weeks. For comparison, we also add a line in red corresponding to estimates of absolute IRFs ($\lambda_{h,i,t}$) despite the statistical evidence against shock invertibility.

Starting with IRFs of interest rates (Figure 7), the first three columns give the time-varying response of the 3 month-, 2 year- and 10 year rate on impact, after 2 weeks and after 4 weeks. Unsurprisingly, we document that in the aftermath of the financial crisis, short-term interest rate started to become relatively unresponsive over the period from 2011 to 2021, and pick up again with the recent rate hike cycle. These time-lines align well with the zero lower bound which was binding short term interest rates from 2009-2021. A similar pattern can be documented for the 2 year rate, although the differences are less significant at 2- and 4 weeks horizons. Surprisingly, we find no evidence that the impact transmission via long term rates has changed over time, despite the onset of quantitative easing as explicit policy toolkit in 2009. However, our estimates suggest that overall persistence of monetary policy has decreased for all interest rates. Towards the end of the sample, we find no evidence that the effect of monetary policy persists at a four weeks horizon. Much alike the 10 year interest rate, transmission of

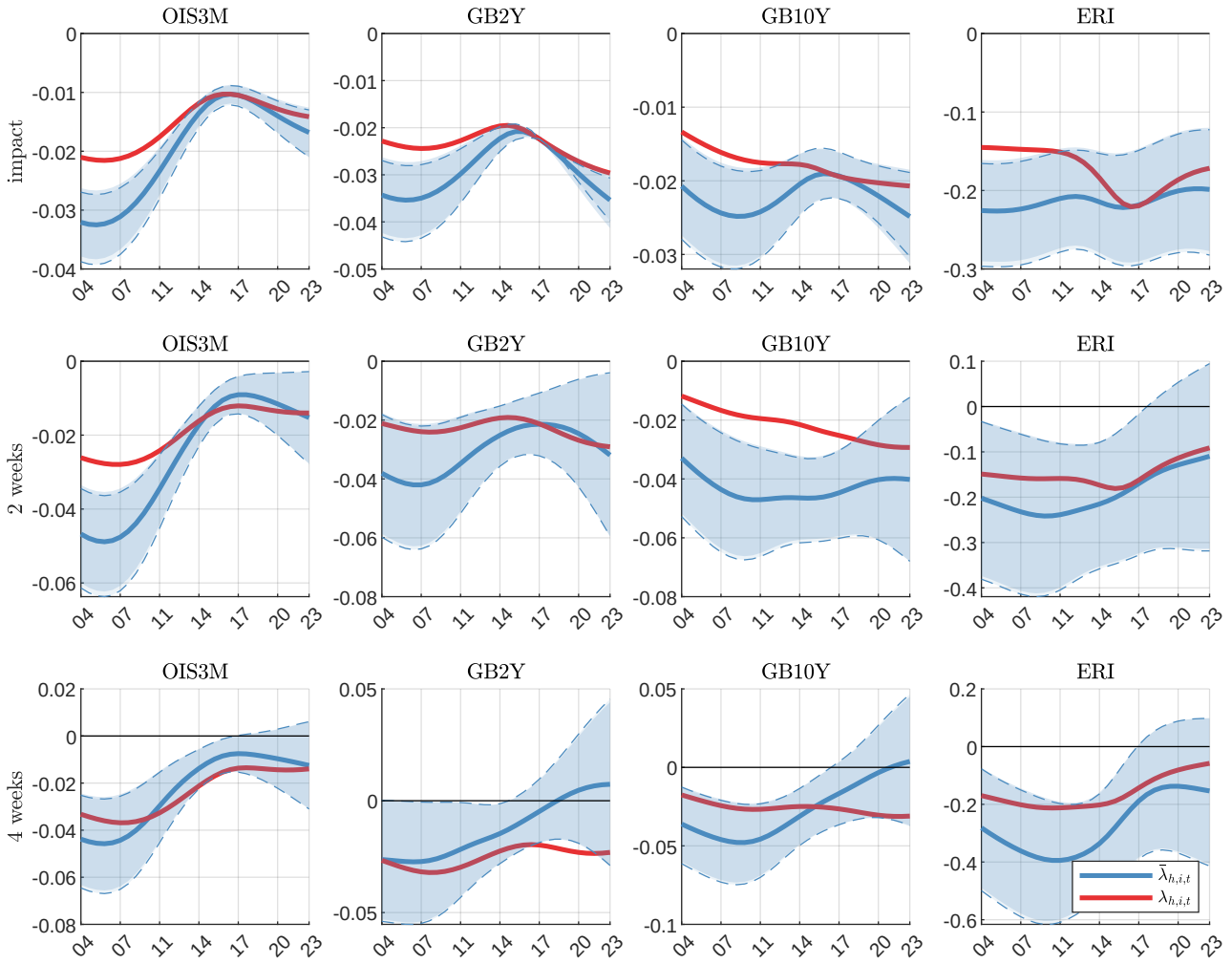


Figure 7: Impulse response functions of an expansionary monetary policy shocks normalized to decrease the 2Y OIS rate by 2.5bp in June 2017. Shaded areas give 90% confidence intervals by the Delta Method, while dashed lines give 90% confidence sets by the AR method.

average policy on the trade weighted exchange rate index (fourth column) is found to be stable over time in the short-run. However, point estimates suggest that the effects have become less persistent over time, since magnitudes at a four week horizon have halved in recent years to what they were estimated to be in the earlier sample. Finally, we note that estimates obtained under the invertibility assumption (red) yield qualitatively similar patterns for time-variation. IRFs estimates for equity prices and corporate financing conditions are given in columns 1-3 of Figure 8. Broadly, we document a similar pattern of time-variation for all three variables: the effect of average policy on prices, yields and spreads have become consistently stronger during the sample. Note that expansionary monetary policy is widely thought of increasing equity and bond prices, given that with an improving economic outlook investors seek yield and are willing

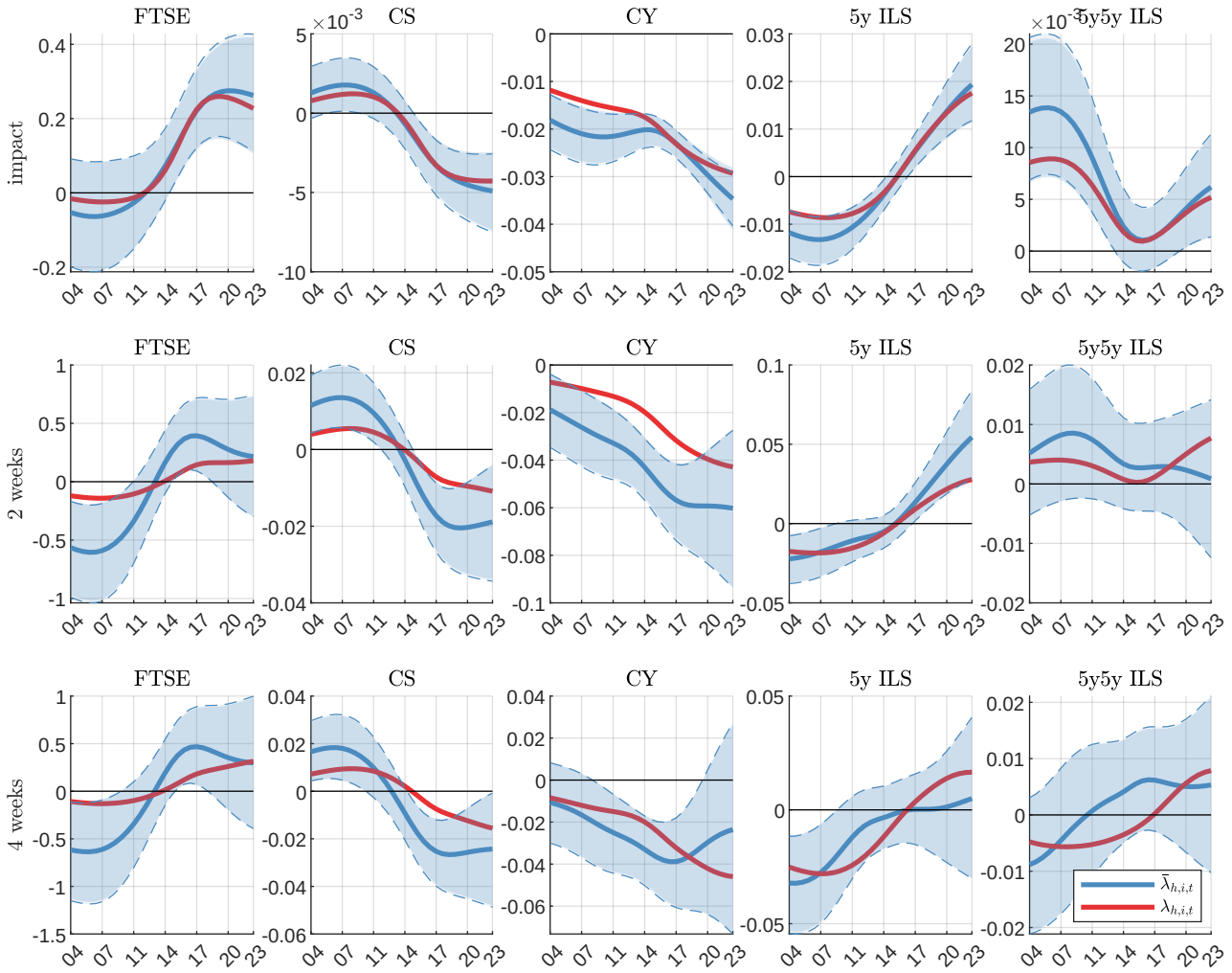


Figure 8: Impulse response functions of an expansionary monetary policy shocks normalized to decrease the 2Y OIS rate by 2.5bp in June 2017. Shaded areas give 90% confidence intervals by the Delta Method, while dashed lines give 90% confidence sets by the AR method.

to bear higher risks in terms of spreads. However, up until about 2012, we find no evidence that negative monetary policy surprises have lead to an increase in the price of risky assets. One way to interpret these findings is the presence of strong information effects in early parts of the sample, a channel through which monetary policy announcements may lead to market participants revising their expectations about fundamentals (Campbell et al., 2012; Nakamura and Steinsson, 2018; Smolyansky and Suarez, 2021). If a negative policy surprise is the result of markets learning from the BOE about worsening economic conditions, this may lower demand for riskier assets. Our estimates would suggest that those effects are less strong in the second part of the sample, where both stock prices and credit spread show more conventional signs.[8]

[8]See also Hoesch et al. (2023) for extensive evidence of a decline in the importance of information effects

Column four gives estimates of TVP IRFs for inflation expectations over the next five years ($5y$ ILS). We find these to be consistent with the presence of information effects in the early part of the sample. Here, negative monetary policy surprises are associated with a decline in inflation expectations, which could reflect markets learning about worsening fundamentals. Towards the end of the sample, however, a negative surprise is associated with a strong positive reaction reflecting increased inflation expectations which aligns well with conventional economic theory. Here, we find that the responses are magnified at a 2 weeks horizon, but there is little evidence that they persist after 4 weeks.

The last column gives the response of long run inflation compensations as measured by the 5-year-5-year forward rate ($5y5y$ ILS). On impact, we document significant time-variation. In response to monetary policy easing, a relatively strong increase is documented around 2008, and again after 2019. However, no significant response is obtained during the period between. Our estimates that long-run inflation expectations seem to react very differently from the medium run measure. Strong impact responses are documented before 2010, and again after 2019, but there is no evidence that these effects persist. Hence, we find no indications of information effects dominating at any part of the sample, suggesting that these predominantly affect short-term expectations.

Similar to IRF estimates of interest rates and exchange rates, absolute IRFs ($\lambda_{h,i,t}$) are very close to those obtained by the internal instrument VAR. Point estimates provided in red are within 90% confidence intervals of $\bar{\lambda}_{h,i,t}$.

In Appendix E.1, we re-estimate the model using a slightly altered instrument that strips out such information effects. Specifically, following Kerssenfischer (2019) and Jarociński and Karadi (2020), a sign-restricted bivariate model is used to decompose the high-frequency policy surprise into pure policy- and information shocks, exploiting distinct co-variability between the 2-year rate and stock market (high frequency) surprises. However, estimates of IRFs are very close to those obtained under the baseline specification, suggesting that conditioning on information

in the United States.

effects in a simple way is not enough to capture the strong time-variation that we document for financial variables.

Appendix E.2 provides a second robustness check of our main empirical results. Here, instead of relying on the 2-year surprise as external instrument, we re-estimate IRFs using an alternative identification strategy that allows for the presence of contaminating shocks. Following Rigobon and Sack (2004), Nakamura and Steinsson (2018) and Wright (2012), we make an assumption that on central bank announcement days, the variance of the (average) monetary policy shock is considerably larger than on a set of control dates which we choose to be the day prior to the announcement day. Assuming that the second moment of contaminating shocks remains broadly stable between the two set of days, an artificial instrument can be created based on reduced-form prediction errors that is relevant and exogenous. Resulting estimates suggest that the documented time-varying patterns for IRFs are robust to such an alternative identification strategy.

## 4.4 The role of different policies in accounting for time-variation

It is possible that the time-variation we document for the effects of monetary policy on financial variables can be explained by an evolving policy toolkit, that is the shift from conventional to unconventional policies. If this is the case, we would expect stable estimates once we condition on instruments that capture variation in the different policy tools. Commonly discussed dimensions thereof are standard rate cuts, anticipated rate cuts and asset purchasing programmes or Quantitative Easing. To shed some light on this question, we proceed using more granular structural factors as instruments that can be related explicitly to instances of conventional monetary policy (target factor), anticipated policy (path factor) and quantitative easing (QE factor). Such measures are widely available in the literature, see e.g. Gürkaynak et al. (2005) and Swanson (2021) for the US, Altavilla et al. (2019) for the Euro Area and Braun et al. (2023) for the UK.

These factors are typically obtained in three steps. First, a panel of high frequency monetary

policy surprises are gathered for the entire term structure of risk free rates up to 10 years maturity. Then, principal component analysis is conducted to extract three orthogonal factors that summarize the bulk of variation in high frequency surprises. The final step involves imposing identifying restrictions to back out a unique rotation of the factors that allow for structural interpretation. Here, only the target factor is allowed to affect the short end of the term-structure. Furthermore, the path and QE factors are disentangled assuming that QE policies were not important prior to the introduction of the first BOE asset purchase program in March 2009. For a detailed description of the construction, we refer to Braun et al. (2023). To gain some intuition on how the factors relate to our benchmark IV, Figure 9 provides a decomposition of the two-year surprise into contributions by the orthogonal target-, path-, and QE factor. Clearly, with an evolving policy mix, the factors have been of varying importance in explaining our benchmark IV throughout the sample. For example, the target factor explains substantial variation of the 2-year surprise at the beginning of the sample, as well as at the end. However, almost no variation is explained by the target factor between 2009 and 2020, which aligns well with the period of the zero-lower bound.

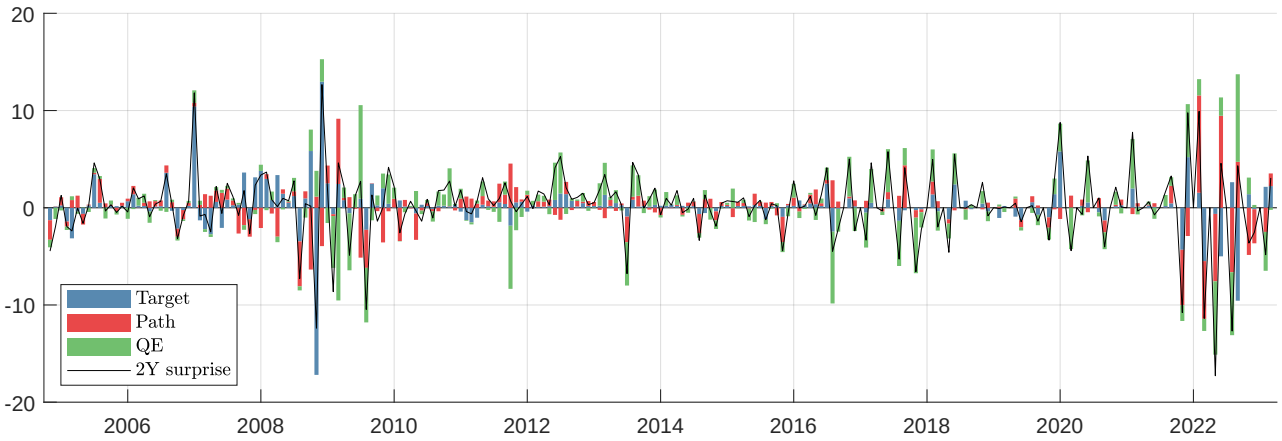


Figure 9: Decomposition of the 2-year Gilt surprise into variation explained by the Target, Path and QE factors.

Equipped with the structural factors, we re-estimate the daily VAR models replacing the 2-year surprise with one factor at a time. Recomputing the Wald statistics of Table 2 for each instrument separately suggests that we can benefit from shifting the date of standardization

$t_b$. Hence, we re-standardize the IRFs to decrease the 2 year rate in April 2014 for the Target factor, in January 2020 for the Path factor and in September 2018 for the QE factor. For comparison, we align the size of the shock to match that of our baseline result at those times.

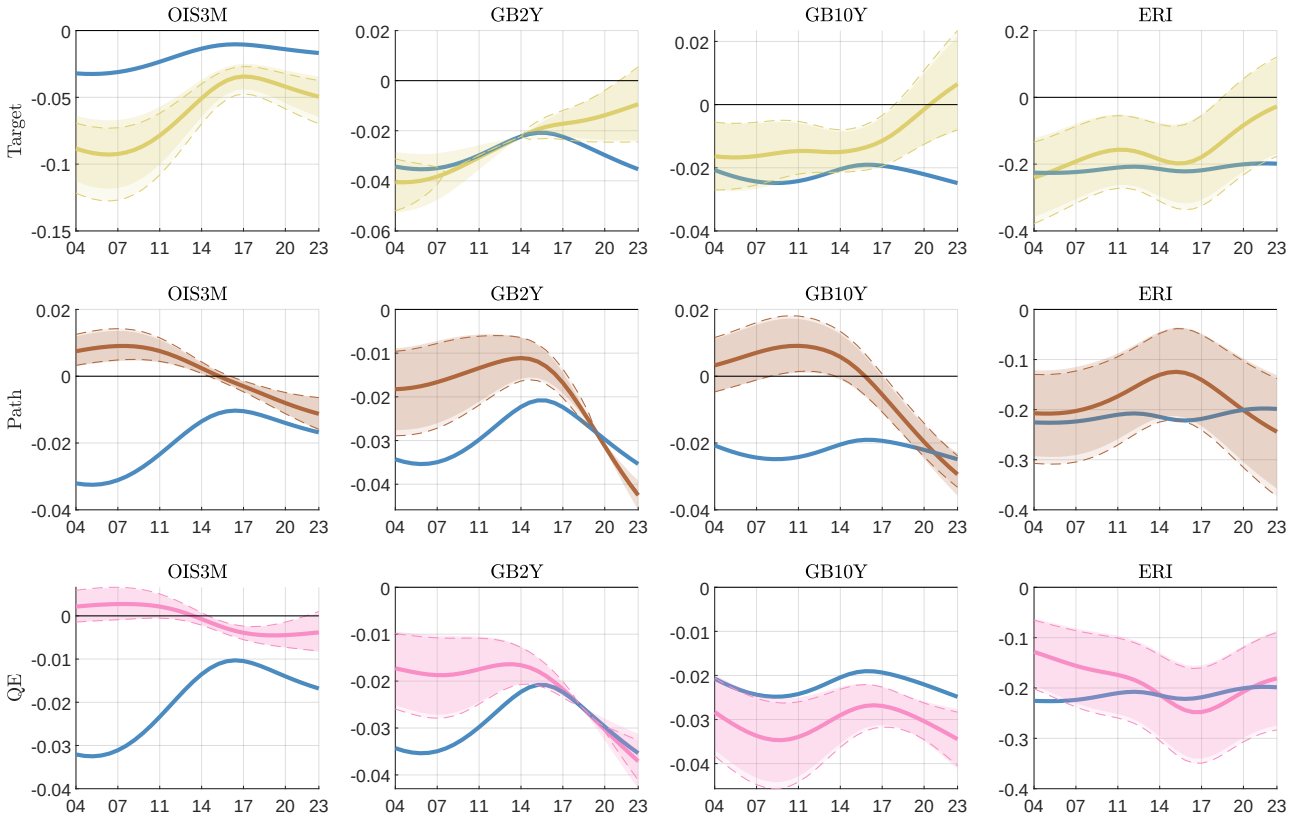


Figure 10: Impact impulse response functions of a expansionary monetary policy shocks normalized to decrease the 2 year rate in April 2014 (Target factor), January 2020 (Path factor) and September 2018 (QE factor). The shock size is chosen to match that of the baseline VAR identified by the 2 year surprise (blue line). Shaded areas give 90% confidence intervals by the Delta Method, while dashed lines give 90% confidence sets by the AR method.

Figure 10 provides estimates of (relative) impact impulse response function of interest rates and the exchange rate index. Here, each row in the panel corresponds to impact IRFs identified using either the Target, the Path or the QE factor as instrumental variable. For comparison, the blue line gives the time-varying IRF obtained when the 2 year raw surprise is used as external instrument, as we study in section 4.3.

We document that generally, conditioning on the monetary policy sub-dimensions doesn't yield time-invariant estimates. This suggests that the evolving policy mix is unlikely to explain the time-variation we document for IRFs in average monetary policy shocks.

Looking at the effects of conventional rate surprises (first row), we find that transmission to interest rates and the exchange rates have declined over time. In more recent years, no significant effect can be documented for the 2-year, 10-year and exchange rate index.

Anticipated policy captured by the path factor (second row) has become more effective transmitting to interest rates, particularly towards the end of the sample. The effect on the Exchange Rate Index, however, is estimated to be broadly time-invariant.

Finally, monetary policy shocks identified by the QE factor (third row) is estimated to have a somewhat stronger impact on the 2 year rate over time. However, no significant time-variation can be documented for the 10-year rate and the Exchange Rate Index. Overall, Figure 10 suggests that a (constant) shock in anticipated and QE policies are estimated to display a stronger transmission to interest rates, while the opposite is the case for standard rate hikes.

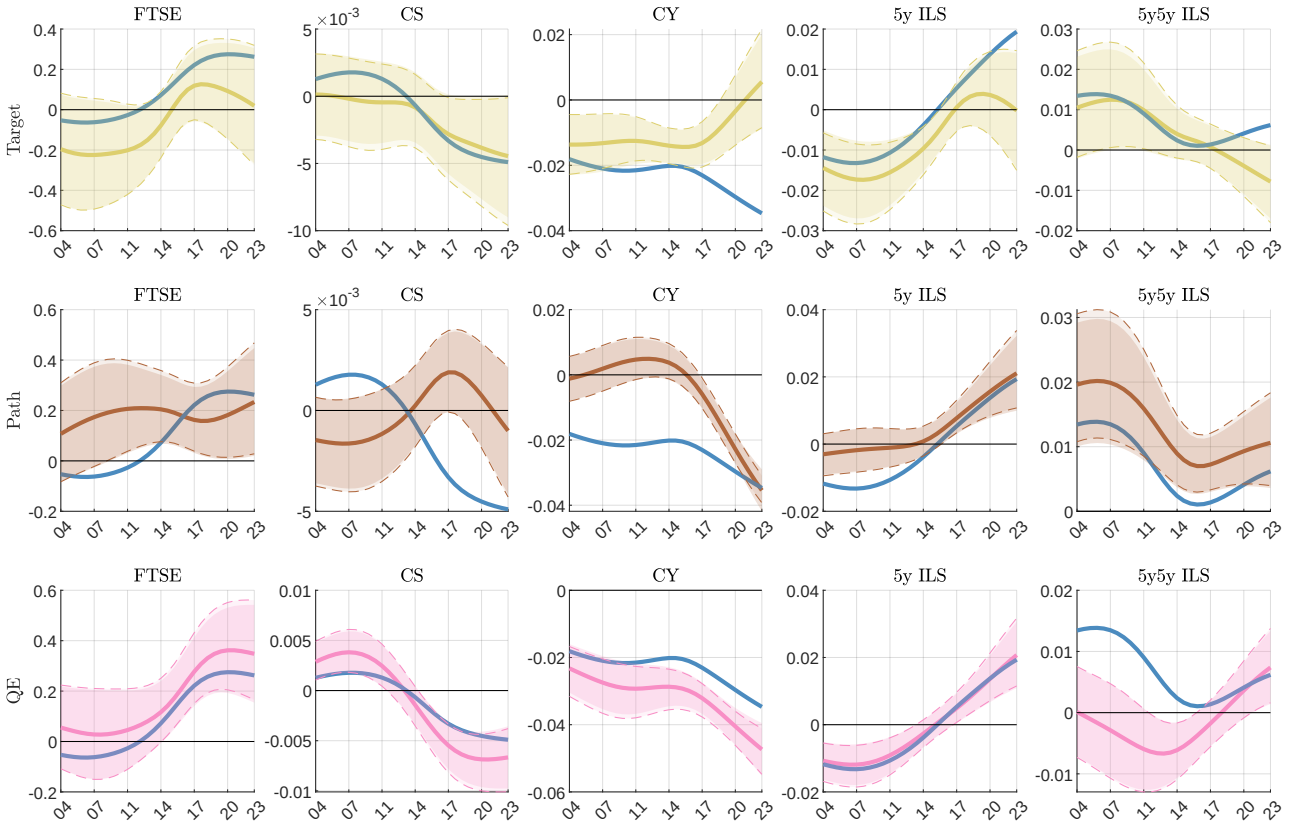


Figure 11: Impact impulse response functions of a expansionary monetary policy shocks normalized to decrease the 2 year rate in February 2006 (Target factor), September 2020 (Path factor) and June 2019 (QE factor). The shock size is chosen to match that of the baseline VAR identified by the 2 year surprise (blue line). Shaded areas give 90% confidence intervals by the Delta Method, while dashed lines give 90% confidence sets by the AR method.

Figure 11 shows time-varying impact IRFs of asset prices and market inflation compensations. Point estimates from a VAR identified by the Target factor suggests that information effects were present during the early part of the sample, possibly explaining why stock prices and 5-year inflation expectations are estimated to decrease in response to rate cuts. Similar to our baseline results, IRFs of inflation compensation for the following 5 years ($5y5y$) seem less prone to those effects and display a conventional sign. These trends seem to reverse towards the end of the sample where no significant impact is documented for any of the variables. In line with a more muted transmission to medium- and long run rates, the response of corporate yields is estimated to get smaller over time, while spreads remain broadly constant at very low levels.

The second row shows impact IRFs of a VAR identified by the Path factor. Somewhat surprisingly, there is no evidence of information effects in this model, as the effect of stock prices is estimated to be positive and constant over time. In line with stronger effects on 2-year and 10-year government bond yields (Figure 10), the effect on corporate yields increases towards the end of the sample while spreads remain insignificant throughout. Effects of medium run inflation expectations are estimated to get stronger over time, while the effect on long run inflation compensation halves over the sample period.

Fairly significant time-varying impact effects can be documented for the model identified by the QE factor (last row). The effect of stock prices, corporate bond yields and spreads, as well as inflation expectations is estimated to get significantly stronger over time, particularly after 2014. Unlike for interest rates and the ERI, time-varying patterns seem to be highly correlated with our benchmark estimates discussed in section 4.3). Policies and communications that affect the longer-term interest rates (e.g. QE) seem to be important contributors to the time-variation we observe in risky asset prices and inflation expectations.

Summing up, the time-variation that we observe in stock prices and 5-year inflation expectations are very similar across policy dimensions. Patterns for corporate financing conditions seems most aligned with those obtained using the QE factor as instrument. Finally, the in-

stability in the transmission to long-run inflation expectations is similar to estimates obtained under the Path factor. Overall, we find that further conditioning on more granular policy dimensions cannot explain the time-variation documented for the overall effect of monetary policy on financial variables.

# 5 The time-varying effects of oil supply news on US industrial production

In the following, we use the methodology to revisit the time-varying effects of an exogenous oil supply news shock on US industrial production. Our analysis builds on the work of Känzig (2021) who studies the effects of exogenous changes in oil price expectations caused by OPEC communications, an intergovernmental organization of major oil-producing nations. To capture changes in price expectations orthogonal to the global business cycle, Känzig (2021) constructs an external instrument based on quotes of WTI oil price futures in a narrow window around OPEC production quota announcements. A constant parameter IV-SVAR estimated over the period from 1974 to 2017 suggest consequences for the US economy that mimic a typical supply shock; activity falls, as measured by US industrial production, while both consumer prices and inflation expectations rise.

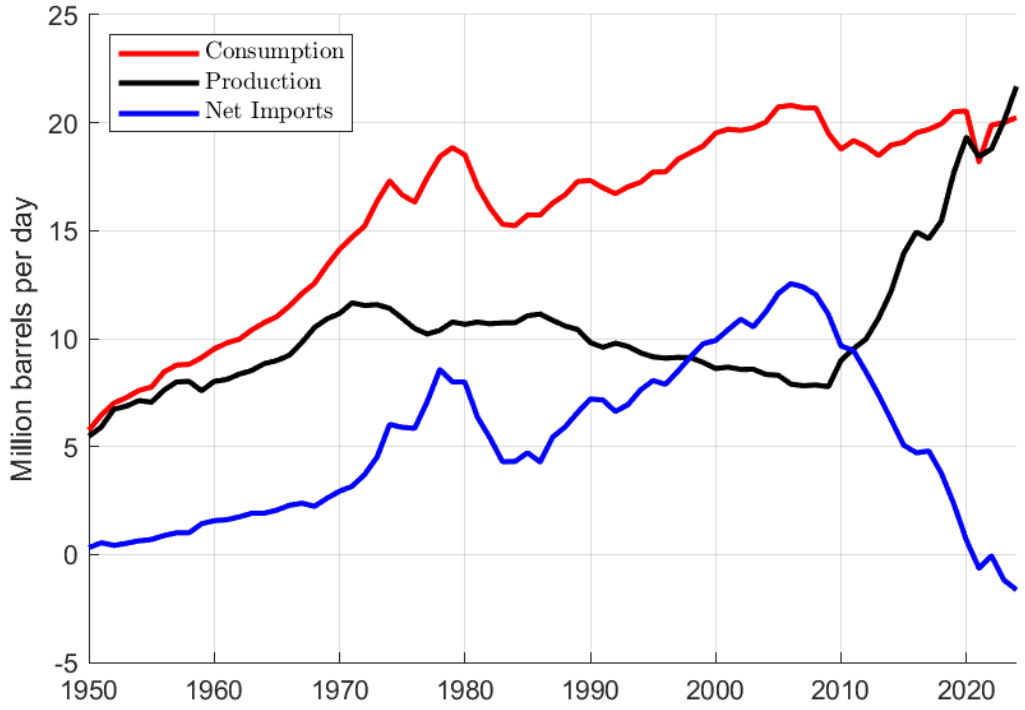


Figure 12: US petroleum consumption, production, and net imports (1950-2023).

There is large evidence in the oil market literature that the relationship between oil prices and US macroeconomic conditions has changed over time, see for example Baumeister and Peersman (2013), Ramey and Vine (2011). Kilian (2009) notes that large part of this instability can be explained by the time-varying importance of supply- and demand shocks. However, even conditioning on supply shocks a some degree of time-variation remains (Baumeister and Peersman, 2013). The variation has been attributed to a variety of potential drivers, including the oil intensity of economic activity, improved monetary policy, or changing importance of certain sectors in the US (Ramey and Vine, 2011). In more recent history, a major driver behind the time-varying relationship has been found to be the shale oil revolution. A combination of hydraulic fracturing and horizontal drilling allowed the US to sharply increase the production of crude oil and natural gas. Over the period 2005-2023, total US petroleum production more than doubled, from an average of 7.9 to 21.7 million barrels per day as shown in Figure 12 (black line).[9] This allowed to United States to transition from a large petroleum net-importer to a

[9]Petroleum production includes field production of crude oil and natural gas, as well as products produced from refining crude oil and from processing natural gas plant liquids.

petroleum net exporter in 2020. Indeed, Bjørnland and Skretting (2024) document extensive evidence that the shale oil revolution aligns with changes in the transmission of oil-market specific shocks to the US economy. Within a Bayesian time-varying factor model, the authors identify a oil-market specific shock by exclusion restrictions, finding that industrial production and investments reacts more positive to oil price increases since the the shale-oil revolution, boosted by activity in oil-intensive regions and industries. Based on our methodology, we find similar results for oil-supply news identified by instrumental variables.

## 5.1 Data and identification strategy

Our analysis is based on a SVAR-IV model with specifications following Känzig (2021). Our model includes a measure of real oil prices, world crude oil production, a proxy of world crude oil stocks, and world industrial production as in Baumeister and Hamilton (2019).[10] We augment the model further by the index of US industrial production, split into manufacturing- and mining output. All variables are included in log levels. To identify shocks to oil supply expectations, we rely on an updated surprise series as external instrument, made available on the homepage of Diego Kanzig. The estimation sample includes data from January 1974 to December 2023. To deal with Covid outliers, we use a series of dummies thereby excluding the signal from the data between February 2020 and June 2022.[11] Following the original paper, we use $p = 13$ lags.

For the hyperparameter $H$ that governs the overall time-variation, the cross validation procedure discussed in section 2.5 suggests an optimal bandwidth of $H = 190$ when applied to the pre-covid period. Given a relatively flat objective function for any bandwidth between 150-250, we choose a slightly lower bandwidth of $H = 150$ to allow for more time-variation in the IRFs, particularly during the shale-oil revolution. A sequence of F-tests provides no clear evidence (at 5% significance level) that the instrument is Granger causing the endogenous variables in

[10] The real price of oil defined as the WTI price deflated by US CPI, and the proxy of world crude oil stocks is included in seasonally adjusted log levels.

[11] This is similar to setting the weighting function to zero for that time period [need to double check this]

the model (see Table 3), with exception of the very beginning and end of the sample. For that reason, we proceed assuming shock invertibility, and study impulse response functions to a shock of constant size across time, that is of unit standard deviation ($\lambda_{k,i,t}$).

Table 3: Granger causality test results computed at different points of time for the null hypothesis that $z_t$ does not predict $y_t$ in a VAR for $\tilde{y}_t = [z_t, y_t']'$.

| date $t$ | July 77 | May 86 | Feb 95 | Dec 03 | Sep 12 | Jun 21 |
|---|---|---|---|---|---|---|
| F Statistic | 1.35 | 1.08 | 1.01 | 1.08 | 1.27 | 1.67 |
| p-val | 0.040 | 0.317 | 0.464 | 0.327 | 0.078 | 0.001 |

## 5.2 Results

Figure D.21 displays estimates of time-varying impulse response functions to an oil supply news shock at various points in time over the sample (blue line) and compares it to the constant parameter estimates (red line). Shaded areas denote 90% confidence intervals.

As expected, the constant parameter results are very similar to that of Känzig (2021). A supply news shock raises the oil price relatively persistently. Crude oil production is declining gradually, while stocks are rising reflecting precaution by market participants. Global economic activity declines, measured by world industrial production, and so does US manufacturing output. US mining output, which includes extraction of oil and gas, increases slightly but with some lag.

There is a striking amount of time-variation in the transmission of the shock that aligns well with the US shale-oil revolution. An oil supply news shock of constant size is estimated to have just 2/3 towards of the the price effect towards the end of the sample, compared to estimates from 1977-2003. Furthermore, the price effects is less persistent. Despite the more muted price signals, US mining output is estimated to increase by larger amounts towards the end of the sample, and react more quickly. Such a quicker reaction of US mining output aligns well with the micro-evidence of shale-oil producers, who are often found to have a larger price elasticity of supply (Aastveit et al., 2022).

Since 2003, world crude oil production is not found to decline significantly in reaction to the

oil-supply news shock. This may reflect, in part, that the increase by US producers offsets cuts in the OPEC production quotas. In line with increasing world oil production, the IRF of the world industrial production (IP) index which includes energy output is estimated to increase over time. A similar pattern arises for US manufacturing output, although the evidence is less clear cut.

In Appendix D, we provide further industry-level results to learn about the key driving forces of the observed time-variation documented for aggregate US manufacturing output. First, the weight of downstream industries to natural gas and crude oil extraction has gained over the last two decades, such as that of Petroleum and Coal Products Manufacturing which includes refinery output, as well as Chemical Manufacturing and Plastic and Rubber production which often rely on crude oil and other hydrocarbons. As expected, these industries have shown a strong pattern of time-variation, responding more positive to oil price surges than in the past. Other important industries that show strong time-variation are some of the energy-intensive durable goods producing that are upstream, such as Primary Metals and Fabricated Metal products. Also the response of machinery sector, as well as the transportation equipment sector has increased over time, from strongly negative to insignificant.

Our findings suggest that OPEC oil-supply news shocks transmit differently in recent times than a constant parameter model would suggest. Price effects are less pronounced and less persistent, while there is no evidence that crude oil production declines. Both world- and US industrial production output respond fairly positive, and hence these shocks no longer transmit as aggregate cost-push shocks.

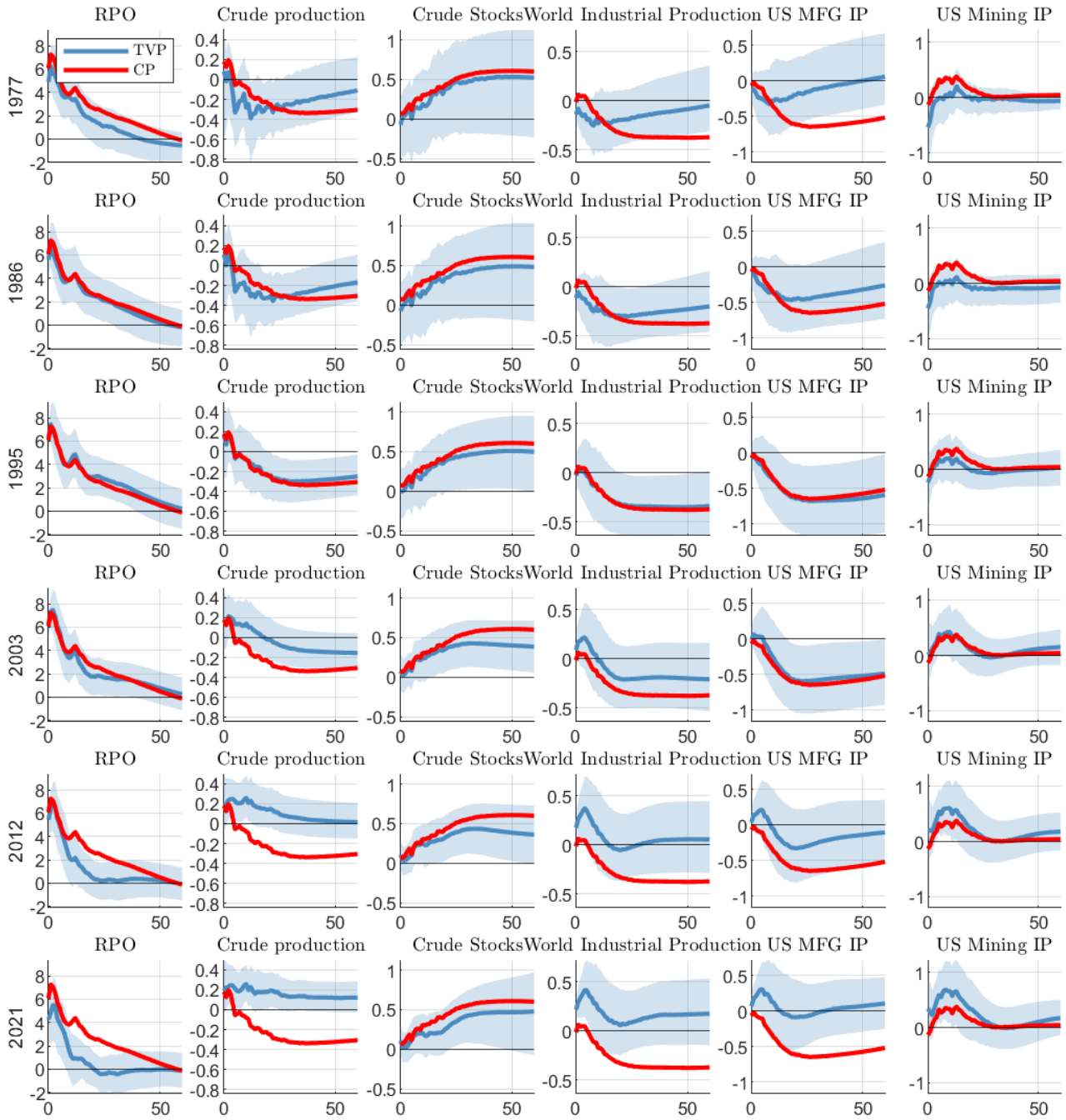


Figure 13: Time-varying impulse response functions to an oil-supply shock of unit variance.

# 6 Conclusion

In this paper, we develop kernel based estimators for time varying impulse response functions of structural VAR models identified by external instruments. Compared to prominent Bayesian approaches, our frequentist estimators are particularly simple to implement, computationally efficient and require no choice for the law of motion and corresponding priors. The amount

of time-variation in a given dataset can be set in an automatic fashion, e.g. by optimizing out-of-sample model fit. Importantly, inference can be reliably conducted even if identification is only weak.

# References


Aastveit, K. A., H. C. Bjørnland, and T. S. Gundersen (2022): "The price responsiveness of shale producers: Evidence from micro data," *Available at SSRN 4273926*.

Altavilla, C., L. Brugnolini, R. S. Gürkaynak, R. Motto, and G. Ragusa (2019): "Measuring euro area monetary policy," *Journal of Monetary Economics*, 108, 162–179.

Arias, J. E., J. F. Rubio-Ramírez, and D. F. Waggoner (2021): "Inference in Bayesian Proxy-SVARs," *Journal of Econometrics*, 225, 88–106.

Baumeister, C. and J. D. Hamilton (2019): "Structural interpretation of vector autoregressions with incomplete identification: Revisiting the role of oil supply and demand shocks," *American Economic Review*, 109, 1873–1910.

Baumeister, C. and G. Peersman (2013): "Time-varying effects of oil supply shocks on the US economy," *American Economic Journal: Macroeconomics*, 5, 1–28.

Bjørnland, H. C. and J. Skretting (2024): "The shale oil boom and the US economy: Spillovers and time-varying effects," *Journal of Applied Econometrics*.

Braun, R., S. Miranda-Agrippino, and T. Saha (2023): "Measuring Monetary Policy in the UK: the UK Monetary Policy Event-Study Database," .

Caldara, D. and E. Herbst (2019): "Monetary policy, real activity, and credit spreads: Evidence from Bayesian Proxy SVARs," *American Economic Journal: Macroeconomics*, 11, 157–92.

Campbell, J. R., C. L. Evans, J. D. Fisher, A. Justiniano, C. W. Calomiris, and M. Woodford (2012): "Macroeconomic effects of federal reserve forward guidance [with comments and discussion]," *Brookings papers on economic activity*, 1–80.

Cochrane, J. H. and M. Piazzesi (2002): "The fed and interest rates-a high-frequency identification," *American economic review*, 92, 90–95.

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

Dahlhaus, R. (1997): "Fitting time series models to nonstationary processes," *The annals of Statistics*, 25, 1–37.

Fieller, E. C. (1944): "A fundamental formula in the statistics of biological assay, and some applications," *Quart. J. Pharm*, 17, 117–123.

Forni, M., L. Gambetti, G. Ricco, et al. (2023): "External Instrument SVAR Analysis for Noninvertible Shocks," Tech. rep.

Fry, R. and A. Pagan (2011): "Sign restrictions in structural vector autoregressions: a critical review," *Journal of Economic Literature*, 49, 938–960.

Gertler, M. and P. Karadi (2015): "Monetary policy surprises, credit costs, and economic activity," *American Economic Journal-Macroeconomics*, 7, 44–76.

Giacomini, R., T. Kitagawa, and M. Read (2022): "Robust Bayesian inference in proxy SVARs," *Journal of Econometrics*, 228, 107–126.

Giraitis, L., G. Kapetanios, and M. Marcellino (2021): "Time-varying instrumental variable estimation," *Journal of Econometrics*, 224, 394–415.

Giraitis, L., G. Kapetanios, and T. Yates (2014): "Inference on stochastic time-varying coefficient models," *Journal of Econometrics*, 179, 46–65.

——— (2018): "Inference on multivariate heteroscedastic time varying random coefficient models," *Journal of Time Series Analysis*, 39, 129–149.

Gürkaynak, R. S., B. Sack, and E. Swanson (2005): "The sensitivity of long-term interest rates to economic news: Evidence and implications for macroeconomic models," *American economic review*, 95, 425–436.

Hoesch, L., B. Rossi, and T. Sekhposyan (2023): "Has the Information Channel of Monetary Policy Disappeared? Revisiting the Empirical Evidence," *American Economic Journal: Macroeconomics*, 15, 355–87.

Jarocinski, M. (2021): "Estimating Fed's Unconventional Policy Shocks," *ECB Working Paper 2585*.

Jarociński, M. and P. Karadi (2020): "Deconstructing monetary policy surprises—the role of information shocks," *American Economic Journal: Macroeconomics*, 12, 1–43.

Känzig, D. R. (2021): "The macroeconomic effects of oil supply news: Evidence from OPEC announcements," *American Economic Review*, 111, 1092–1125.

Kerssenfischer, M. (2019): "The puzzling effects of monetary policy in VARs: Invalid identification or missing information?" *Journal of Applied Econometrics*, 34, 18–25.

Kilian, L. (2008): "Exogenous oil supply shocks: How big are they and how much do they matter for the U.S. economy?" *The Review of Economics and Statistics*, 90, 216–240.

——— (2009): "Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market," *American Economic Review*, 99, 1053–1069.

Koop, G., M. H. Pesaran, and S. M. Potter (1996): "Impulse response analysis in nonlinear multivariate models," *Journal of econometrics*, 74, 119–147.

Kuttner, K. N. (2001): "Monetary policy surprises and interest rates: Evidence from the Fed funds futures market," *Journal of monetary economics*, 47, 523–544.

Lewis, D. J. (2022): "Robust inference in models identified via heteroskedasticity," *Review of Economics and Statistics*, 104, 510–524.

Lütkepohl, H. (1993): "Testing for causation between two variables in higher dimensional VAR models," in *Studies in Applied Econometrics*, ed. by H. Schneeweiß and K. F. Zimmermann, Springer-Verlag, Heidelberg, 75–91.

Mertens, K. and M. O. Ravn (2013): "The dynamic effects of personal and corporate income tax changes in the United States," *American Economic Review*, 103, 1212–1247.

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

Montiel-Olea, J. L., J. H. Stock, and M. W. Watson (2021): "Inference in structural vector autoregressions identified with an external instrument," *Journal of Econometrics*, 225, 74–87.

Nakamura, E. and J. Steinsson (2018): "High-frequency identification of monetary non-neutrality: the information effect," *The Quarterly Journal of Economics*, 133, 1283–1330.

Paul, P. (2020): "The time-varying effect of monetary policy on asset prices," *Review of Economics and Statistics*, 102, 690–704.

Plagborg-Møller, M. and C. K. Wolf (2021): "Local projections and VARs estimate the same impulse responses," *Econometrica*, 89, 955–980.

——— (2022): "Instrumental variable identification of dynamic variance decompositions," *Journal of Political Economy*, 130, 2164–2202.

Primiceri, G. E. (2005): "Time varying structural vector autoregressions and monetary policy," *Review of Economic Studies*, 72, 821–852.

Ramey, V. A. and D. J. Vine (2011): "Oil, automobiles, and the US economy: How much have things really changed?" *NBER macroeconomics annual*, 25, 333–368.

Rigobon, R. and B. Sack (2004): "The impact of monetary policy on asset prices," *Journal of monetary economics*, 51, 1553–1575.

Smolyansky, M. and G. Suarez (2021): "Monetary policy and the corporate bond market: How important is the Fed information effect?" .

Staiger, D. and J. H. Stock (1997): "Instrumental Variables Regression with Weak Instruments," *Econometrica*, 65, 557–586.

Stock, J. (2008): "What Is New in Econometrics: Time Series," Tech. rep., Lecture 7. In: Short Course Lectures, NBER Summer Institute.

Stock, J. H. and M. W. Watson (1996): "Evidence on structural instability in macroeconomic time series relations," *Journal of Business & Economic Statistics*, 14, 11–30.

——— (2012): "Disentangling the channels of the 2007-09 recession," *Brookings Papers on Economic Activity*, 43, 81–156.

——— (2016): "Dynamic factor models, factor-augmented vector autoregressions, and structural vector autoregressions in macroeconomics," in *Handbook of macroeconomics*, Elsevier, vol. 2, 415–525.

——— (2018): "Identification and estimation of dynamic causal effects in macroeconomics using external instruments," *The Economic Journal*, 128, 917–948.

Swanson, E. T. (2021): "Measuring the effects of Federal Reserve forward guidance and asset purchases on financial markets," *Journal of Monetary Economics*, 118, 32–53.

Waggoner, D. F. and T. Zha (2003): "A Gibbs sampler for structural vector autoregressions," *Journal of Economic Dynamics and Control*, 28, 349–366.

WRIGHT, J. H. (2012): "What does monetary policy do to long-term interest rates at the zero lower bound?" *The Economic Journal*, 122, 447–466.

# Appendix A Proofs

## A.1 Proof of Theorem 1

Theorem 1 states that under Assumption 2-4 and $H = o(T^{\frac{1}{2}})$, it holds that:

$$\sqrt{H}\begin{pmatrix} \hat{\beta}_t - \beta_t \\ \hat{\Gamma}_t - \Gamma_t \\ \text{vech}(\hat{\Sigma}_t) - \sigma_t \end{pmatrix} \xrightarrow{d} \mathcal{N}(0, V_{\theta_t}),$$

for $V_{\theta_t} = S_t \Pi_{ww,t} S_t'$ and

$$S_t = \begin{pmatrix} I_n \otimes \Pi_{x,t}^{-1} & 0 & 0 \\ -\left(I_n \otimes \Pi_{xz,t}\Pi_{x,t}^{-1}\right) & I & 0 \\ 0 & 0 & S_\sigma \end{pmatrix},$$

for $\Pi_{x,t} = \text{plim}_{T\to\infty} \frac{1}{H} \sum_{j=1}^{T} w_{j,t} x_j x_j'$, $\Pi_{xz,t} = \text{plim}_{T\to\infty} \frac{1}{H} \sum_{j=1}^{T} w_{j,t} z_j x_j$,

$\Pi_{ww,t} = \text{plim}_{T\to\infty} \frac{1}{H} \sum_{j=1}^{T} w_{j,t}^2 \xi_j \xi_j'$, $\xi_j = [\text{vec}(x_j u_j)', (z_j u_j - \Gamma)', \text{vec}\left(u_j' u_j - \Sigma_t\right)']'$, and $S_\sigma$ such that $\text{vech}(\Sigma_t) = S_\sigma \,\text{vec}(\Sigma_t)$.

For $x_t' = [y_{t-1}', y_{t-2}', \ldots, y_{t-p}', 1]'$ a $1 \times k$ vector, the model reads

$$\underbrace{y_t'}_{1\times n} = \underbrace{x_t'}_{1\times k}\underbrace{\Theta_t}_{k\times n} + u_t'$$

$$\underbrace{y_t}_{n\times 1} = \underbrace{(I_n \otimes x_t')}_{n\times nk}\underbrace{\beta_t}_{nk\times 1} + \underbrace{u_t}_{n\times 1}$$

$$y_t = \tilde{x}_t \beta_t + u_t$$

where $\tilde{x}_t = (I_n \otimes x_t')$ and $\beta_t = \text{vec}(\Theta_t)$. Let $z_t$ be a $m \times 1$ random vector that is *correlated* with $u_t$. We wish to consider estimating $\underbrace{\Gamma_t}_{n\times m} = E\left(u_t z_t'\right)$ allowing for this quantity to vary over time. To do so we wish to derive the asymptotic distribution of $\frac{1}{\sqrt{H}} \sum_j w_{t,j}\left(H\right)\left(\hat{u}_{t,j} z_j' - E\left(u_j z_j'\right)\right)$ where $w_{t,j}\left(H_2\right) = \frac{H\tilde{w}_{t,j}(H)}{\sum_j \tilde{w}_{t,j}(H)}$, $\hat{u}_j = y_j - \tilde{x}_j \hat{\beta}_t$ and

$$\hat{\beta}_t = \left[I_n \otimes \sum_{j=1}^{T} w_{t,j}\left(H_1\right) x_j x_j'\right]^{-1}\left[\sum_{j=1}^{T} w_{t,j}\left(H_1\right) \text{vec}(x_j y_j')\right]$$

where

$$w_{t,j}(H) = K(|t-j|/H), \tag{18}$$

where $H \to \infty$, $H = o(T)$. $K(x)$, $x \in (0,\, a)$ is a non-negative continuous function with finite or infinite support, such that for some $C > 0$ and $\nu > 3$,

$$K(x) \leq C(1+x^{\nu})^{-1}, \quad |(d/dx)K(x)| \leq C(1+x^{\nu})^{-1}, \quad x \in (0,\, a). \tag{19}$$

First, consider $\sqrt{T}\left(\hat{\beta}_t - \beta_t\right)$:

$$\begin{aligned}
\sqrt{H}\left(\hat{\beta}_t - \beta_t\right) &= \underbrace{\left[I_n \otimes \left(\frac{1}{H}\sum_{j=1}^{T} w_{t,j}(H)\, x_j x_j'\right)^{-1}\right]}_{S_{t,xx}(H)} \frac{1}{\sqrt{H}}\sum_{j=1}^{T} w_{t,j}(H)\,\mathrm{vec}(x_j u_j') \\
&= S_{t,xx}(H)\frac{1}{\sqrt{H}}\sum_{j=1}^{T} w_{t,j}(H)\,\mathrm{vec}(x_j u_j') \\
&= S_{t,xx}(H)\frac{1}{\sqrt{H}}\sum_{j=1}^{T} w_{t,j}(H)\,(I_n \otimes x_j) u_j
\end{aligned}$$

Next, consider $\hat{\Gamma}_t = \frac{1}{H}\sum_{j=1}^{T} w_{t,j}(H)\,\hat{u}_{t,j} z_j'$ and $\hat{\gamma}_t = \mathrm{vec}(\hat{\Gamma}_t) = \frac{1}{H}\sum_{j=1}^{T} w_{t,j}(H)\,(z_j \otimes I_n)\,\hat{u}_{t,j}$. Denote by $\gamma_t = E_t[\mathrm{vec}(u_t z_j')]$ and use that $\hat{u}_t = u_t - \tilde{x}_t\left(\hat{\beta}_t - \beta_t\right)$ :

$$\begin{aligned}
\sqrt{H}\left(\hat{\gamma}_t - \gamma_t\right) &= \frac{1}{\sqrt{H}}\left(\sum_{j=1}^{T} w_{t,j}(H)\,\hat{u}_{t,j} z_j' - \Gamma_t\right) \\
&= \frac{1}{\sqrt{H}}\left(\sum_{j=1}^{T} w_{t,j}(H)\,(z_j \otimes I_n)\,u_j - \gamma_t - \frac{1}{H}\sum_{j=1}^{T} w_{t,j}(H)\,(z_j \otimes I_n)\,\tilde{x}_j H\left(\hat{\beta}_t - \beta_t\right)\right) \\
&= \frac{1}{\sqrt{H}}\left(\sum_{j=1}^{T} w_{t,j}(H)\,(z_j \otimes I_n)\,u_j - \gamma_t\right) - \underbrace{\left(\frac{1}{H}\sum_{j=1}^{T} w_{t,j}(H)\,(z_j \otimes I_n)\,\tilde{x}_j\right)}_{S_{t,zx}(H)} \sqrt{H}\left(\hat{\beta}_t - \beta_t\right).
\end{aligned}$$

Define $S_t(H) = \begin{bmatrix} S_{t,xx}(H) & 0 & 0 \\ -S_{t,zx}(H)S_{t,xx}(H) & I & 0 \\ 0 & 0 & S_\sigma \end{bmatrix}$, then it is:

$$\sqrt{H}\begin{pmatrix} \hat{\beta}_t - \beta_t \\ \hat{\gamma}_t - \gamma_t \\ \hat{\sigma}_t - \sigma_t \end{pmatrix} = \underbrace{\begin{bmatrix} S_{t,xx}(H) & 0 & 0 \\ -S_{t,zx}(H)S_{t,xx}(H) & I & 0 \\ 0 & 0 & S_\sigma \end{bmatrix}}_{S_t} \frac{1}{\sqrt{H}} \sum_{j=1}^{T} w_{t,j}(H) \underbrace{\begin{pmatrix} \text{vec}(x_j u_j') \\ \text{vec}(u_j z_j' - \Gamma_t) \\ \text{vec}(u_j u_j - \Sigma_t) \end{pmatrix}}_{\xi_j}$$

and therefore the asymptotic covariance is given by $V_{\theta_t} = S_t \Pi_{ww,t} S_t'$ for $\frac{1}{\sqrt{H}} \sum_{j=1}^{T} w_{t,j}\xi_j \rightarrow \mathcal{N}(0, \Pi_{ww,t})$.

The results of the Theorem follow directly from Theorem 2.2 of Giraitis et al. (2018) (GKY18) once we account for the presence of the exogenous variable, $z_t$ (Extension 1 (E1)) and the introduction of a lag order greater than 1 (Extension 2 (E2)). The only other difference between the analysis of GKY18 and ours is that GKY18 allow for stochastic parameter processes. We choose to restrict ourselves to deterministic sequences for the parameter processes, to simplify the presentation of our asymptotic results.

We consider each extension in turn, starting with E1. There are two matters relating to proving E1. The first relates to extending Theorem 2.1 of GKY18 to this case (Result E11, (RE11)), and the second is to establish asymptotic normality as in (2.15) of GKY18 (Result E12, (RE12)). RE11 follows immediately by 2 and (6.2)-(6.3) of GKY18.

RE12 relates to showing normality of term $T_{n,t;1}$ (the first term of $T_{n,t}$) in page 41 of the online appendix of GKY18. Normality follows immediately by Lemma 6.2 (ii) of GKY18 using Assumption 3.

Next, consider E2. The result here follows immediately by considering the companion form given by

$$\tilde{y}_t = \tilde{A}_t \tilde{y}_{t-1} + \nu_t, \tag{20}$$

where $\tilde{y}_t = (y_t', y_{t-1}', ..., y_{t-p+1}')'$, $\tilde{A}_t = \begin{pmatrix} A_{1t} & A_{1t} & ... & A_{pt} \\ I & 0 & ... & 0 \\ 0 & ... & ... & ... \\ ... & ... & I & 0 \end{pmatrix}$, $\nu_t = \left((B_t\varepsilon_t)', 0, ..., 0\right)'$ and applying Theorem 2.2 of GKY18.

The only result that needs to be proven is the asymptotic independence of $\hat{\beta}_t$ and $\hat{\sigma}_t$. We revisit the proof of Theorem 2.2 of GKY18. The asymptotically relevant terms of $\sqrt{H}\left(\hat{\beta}_t - \beta_t\right)$ and $\sqrt{H}\left(\hat{\sigma}_t - \sigma_t\right)$ are given by $T_{n,t;1}$ and $q_{n,t}$ which are both defined in page 41 of the online appendix of GKY18. The expectation of their cross product involves the third moments of $\varepsilon_t$ which are zero by the symmetry assumption of Theorem 2 proving the result. The proof for independence between $\hat{\gamma}_t$ and $\hat{\sigma}_t$ can be established in an equivalent way.

## A.2 Proof of Corollary 2

We discuss the covariance term of the two differently dated estimators in the statement of the Corollary. To do this we need to extend slightly the work of GKY18. To do so we will revert to the notation of the proof of their Lemma 6.2. Recall $\xi_{tj} := K_{2,t}^{-1/2} b'\varepsilon_j y_{j-1}' V_{\psi,t_0}^{-1/2} a$ where $K_t = \sum_{j=1}^{T} w_{tj}$, $K_{2,t} = \sum_{j=1}^{T} w_{tj}^2$ $V_{\psi,t_0}$ is defined in (2.17) of GKY18 and $b$, and $a$ are vectors of constants. Then, following the proof of Lemma 6.2, following (6.28) of GKY18, it suffices to determine the probability limit of $\sum_{|t-j|<h} w_{t-1j} w_{tj} E[\xi_{tj}\xi_{t-1j}|\mathcal{F}_{j-1}]$. We note that

$$E[\xi_{tj}\xi_{t-1j}|\mathcal{F}_{j-1}] = K_{2,t}^{-1/2} K_{2,t-1}^{-1/2} E(b'\varepsilon_j)^2 a' V_{\psi,t_0}^{-1/2} y_{j-1} y_{j-1}' V_{\psi,t_0}^{-1/2} a$$

where $E(b'\varepsilon_j)^2 = ||b||^2$. Setting $\tilde{V}_{yyc,t} := K_{(1),t}^{-1} \sum_{|t-j|<h} w_{t-1j} w_{tj} y_{j-1} y_{j-1}'$, for $K_{(q),t} = \sum_{j=1}^{n} w_{t-qj} w_{tj}$, we obtain

$$j_{tn} := \tilde{K}_{(1)t} ||b||^2 a' V_{\psi,t_0}^{-1/2} \tilde{V}_{yyc,t} V_{\psi,t_0}^{-1/2} \boldsymbol{a} = \tilde{K}_{(1)t} ||b||^2 + r_{tn},$$

where $\tilde{K}_{(q)t} = K_{2,t}^{-1/2} K_{2,t-q}^{-1/2} K_{(q),t}$, and $r_{cn} = \tilde{K}_{(1)t} ||b||^2 a' V_{\psi,t_0}^{-1/2} (\tilde{V}_{yyc,t} - V_{\psi,t_0}) V_{\psi,t_0}^{-1/2}$. It remains to show that $r_{tn} \to_p 0$ which involves checking that

$$||\tilde{V}_{yyc,t} - V_{\psi,t_0}||_{sp} = o_p(1).$$

We need to consider $||K_{(1),t}^{-1} \sum_{j=1}^{n} w_{t-1j} w_{tj} \boldsymbol{y}_{j-1} \boldsymbol{y}_{j-1}' - V_{\psi,t}||_{sp}$ which is $o_p(1)$ by Lemma 6.1(i) of GKY18. This of course easily generalises to $\sum_{|t-j|<h} w_{t-qj} w_{tj} E[\xi_{tj} \xi_{t-qj} | \mathcal{F}_{j-1}]$ for all finite $q$.

### A.3 Proof of Theorem 2

All the results of this Theorem follow directly from the proof of Theorem 1.

## Appendix B Inference for structural impulse response functions

This part of the Appendix gives detailed formulas in order to compute closed form Delta Method and Anderson Rubin confidence sets for the (time-varying) IV-SVAR estimator and the internal IV-VAR estimator.

### B.1 Absolute Impulse Response Functions (IV-SVAR estimator)

In this paper, we use the IV-SVAR estimator to recover absolute impulse response functions $\lambda_{h,i,t}$, that is the $i$th element of the $n \times 1$ vector $\lambda_{k,t}$. The corresponding function is given by:

$$\hat{\lambda}_{k,t} = C_k(\hat{A}_t) \hat{\Gamma}_t / \sqrt{\hat{\Gamma}_t' \hat{\Sigma}_t^{-1} \hat{\Gamma}_t}$$

Building on the reduced form results given in theorem 1, we get

$$\sqrt{H} \begin{pmatrix} \hat{\beta}_t - \beta_t \\ \hat{\Gamma}_t - \Gamma_t \\ \text{vech}(\hat{\Sigma}_t) - \sigma_t \end{pmatrix} \xrightarrow{d} \mathcal{N}(0, V_{\theta_t}).$$

Starting with the Delta Method, as described in section 2.4, we have $\sqrt{H}\left(\hat{\lambda}_{k,t} - \lambda_{k,t}\right) \xrightarrow{d} \mathcal{N}(0, \Omega_{k,t})$, where $\Omega_{k,t} = J_k(\beta_t, \Gamma_t, \sigma_t) V_{\theta_t} J_k(\beta_t, \Gamma_t, \sigma_t)'$ for $\beta_t = \text{vec}(A_t)$, and

$$J_k(\beta_t, \Gamma_t, \sigma_t) = \left[\frac{\partial \lambda_{k,t}}{\partial \beta_t} : \frac{\partial \lambda_{k,t}}{\partial \Gamma_t} : \frac{\partial \lambda_{k,t}}{\partial \sigma_t}\right]$$

is the $n \times (n^2p + n + n(n+1)/2)$ dimensional gradient. The corresponding derivatives are stated in the following. First, note that $C_k(A_t) = J_s \mathbf{A}_t^k J_s'$ where $J_s = [I_n, 0, \ldots, 0]$ and

$$\mathbf{A}_t = \begin{pmatrix} A_{1t} & A_{2t} & \ldots & A_{p-1,t} & A_{pt} \\ I_n & 0 & \ldots & 0 & 0 \\ 0 & I_n & & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & 0 \\ 0 & 0 & \ldots & I_n & 0 \end{pmatrix}.$$

Hence, it is $\frac{\partial \lambda_{k,t}}{\partial \beta_t} = 0$ for $k = 0$ while for $k > 1$:

$$\frac{\partial \lambda_{k,t}}{\partial \beta_t'} = \left((\Gamma_t/\alpha_t)' \otimes I_n\right) G_k,$$

where $\alpha_t = \sqrt{\Gamma_t' \Sigma_t^{-1} \Gamma_t}$ and $G_k = \frac{\partial \, \text{vec}(C_k(A_t))}{\partial \beta_t'} = \sum_{m=0}^{k-1} \left[J(\mathbf{A}_t')^{k-1-m}\right] \otimes C_m(A_t)$ (Lütkepohl, 1993). Next, define $\frac{\partial [\Gamma_t, \alpha_t]}{\partial [\Gamma_t', \sigma_t']'} = \left[\frac{\partial \Gamma_t}{\partial [\Gamma_t', \sigma_t']'} : \frac{\partial \alpha_t}{\partial [\Gamma_t', \sigma_t']'}\right]$ where it holds that:

$$\begin{aligned} \frac{\partial \Gamma_t}{\partial [\Gamma_t', \sigma_t']'} &= [I_n : 0], \\ \frac{\partial \alpha_t}{\partial [\Gamma_t', \sigma_t']'} &= \frac{1}{2}\left(\Gamma_t' \Sigma_t^{-1} \Gamma_t\right)^{-1/2} \left[2\Gamma_t' \Sigma_t^{-1}, -\left(\Gamma_t \Sigma_t^{-1} \otimes \Gamma_t' \Sigma_t^{-1}\right) D\right], \end{aligned}$$

for $D$ is the duplication matrix such that $\text{vec}(\Sigma_t) = D \, \text{vech}(\Sigma_t)$. Also, it holds that:

$$\frac{\partial \lambda_{k,t}}{\partial [\Gamma_t', \alpha_t']'} = C_k(A_t) \left[I_n/\alpha_t : \Gamma_t/\alpha_t^2\right].$$

Combining both results via the Chain rule yields the missing parts of $J_k()$:

$$\left[\frac{\partial \lambda_{k,t}}{\partial \Gamma_t} : \frac{\partial \lambda_{k,t}}{\partial \sigma_t}\right] = \frac{\partial \lambda_{k,t}}{\partial [\Gamma_t', \alpha_t']'} \times \frac{\partial [\Gamma_t, \alpha_t]}{\partial [\Gamma_t', \sigma_t']'}.$$

With respect to the AR confidence set, the first step is to obtain the asymptotic distribution

of the $(n+1) \times 1$ vector:

$$L_{k,t} = \begin{pmatrix} C_k(A_t)\Gamma_t \\ \sqrt{\Gamma_t' \Sigma_t^{-1} \Gamma_t} \end{pmatrix}$$

for which it holds that $\lambda_{i,t} = (e_i' L_{k,t})/(e_{n+1}' L_{k,t})$. Via the Delta Method we get: $\sqrt{H}\left(\hat{L}_{k,t} - L_{k,t}\right) \xrightarrow{d} \mathcal{N}(0, \Omega_{k,t}^L)$ for $\Omega_{k,t}^L = J_k^{(2)}\left(\beta_t, \Gamma_t, \sigma_t\right) V_{\theta_t} J_k^{(2)}\left(\beta_t, \Gamma_t, \sigma_t\right)'$ where

$$J_k^{(2)}\left(\beta_t, \Gamma_t, \sigma_t\right) = \left[\frac{\partial L_{k,t}}{\partial \beta_t} : \frac{\partial L_{k,t}}{\partial \Gamma_t} : \frac{\partial L_{k,t}}{\partial \sigma_t}\right].$$

Similar to above, it holds that $\frac{\partial L_{k,t}}{\partial \beta_t} = 0$ for $k = 0$ while for $k > 1$:

$$\frac{\partial L_{k,t}}{\partial \beta_t'} = \begin{pmatrix} \left(\Gamma_t' \otimes I_n\right) G_k \\ 0 \end{pmatrix}.$$

Finally, the last step is:

$$\left[\frac{\partial L_{k,t}}{\partial \Gamma_t} : \frac{\partial L_{k,t}}{\partial \sigma_t}\right] = \frac{\partial L_{k,t}}{\partial \left[\Gamma_t', \alpha_t'\right]'} \times \frac{\partial \left[\Gamma_t, \alpha_t\right]}{\partial \left[\Gamma_t', \sigma_t'\right]'},$$

where $\frac{\partial [\Gamma_t, \alpha_t]}{\partial [\Gamma_t', \sigma_t']'}$ is as defined above and

$$\frac{\partial L_{k,t}}{\partial \left[\Gamma_t', \alpha_t'\right]'} = \begin{pmatrix} C_k(A_t) & 0 \\ 0 & 1 \end{pmatrix}.$$

Next, consider the linear test $e_i' \hat{L}_{k,t} - \lambda_0 e_{n+1}' \hat{L}_{k,t} = 0$ with the corresponding Wald test statistic $q(\lambda_0) = \frac{H(e_i' \hat{L}_{k,t} - \lambda_0 e_{n+1}' \hat{L}_{k,t})^2}{\hat{\omega}_{ii} - 2\lambda_0 \hat{\omega}_{i,n+1} + \lambda_0^2 \hat{\omega}_{n+1,n+1}}$ where $\hat{\omega}_{ij}$ is the $ij$th element of of $\hat{\Omega}_{k,t}^L$. The AR confidence set of coverage $1-a$ is then given by inverting the test statistic, yielding $\text{CS}^{AR}\{\lambda_{k,i,t} | q(\lambda_{k,i,t}) \leq \chi_{1,1-a}^2\}$. The inversion can be solved in closed form following, e.g. footnote 14 in Montiel-Olea et al. (2021).

## B.2 Relative Impulse Response Functions (internal IV VAR estimator)

For relative impulse response functions, the corresponding function of reduced form parameters is given by:

$$\hat{\bar{\lambda}}_{k,t} = C_k\left(\hat{\tilde{A}}_t\right)\hat{\tilde{P}}_{\bullet 1,t}/(e_2'\hat{\tilde{P}}_{\bullet 1,t_b}),$$

where $\hat{\bar{\lambda}}_{k,i,t} = e_{1+i}'\hat{\bar{\lambda}}_{k,t}$. Furthermore, $\hat{\tilde{A}}_t$, $\hat{\tilde{P}}_{\bullet 1,t} = e_1'\text{chol}(\hat{\tilde{\Sigma}}_t)$ and $\hat{\tilde{P}}_{\bullet 1,t_b} = e_1'\text{chol}(\hat{\tilde{\Sigma}}_{t_b})$ are based on kernel estimates of the TVP internal instrument VAR.

Starting from the reduced form results of Theorem 2 and Corollary 3, we have:

$$\sqrt{H}\left(\hat{\tilde{\beta}}_t - \tilde{\beta}_{t_b}\right) \xrightarrow{d} \mathcal{N}\left(0, \underbrace{\tilde{\Sigma}_t \otimes \left(\tilde{\Pi}_{x,t}\right)^{-1}\tilde{\Pi}_{ww,t}\left(\tilde{\Pi}_{x,t}\right)^{-1}}_{V_1}\right),$$

$$\sqrt{H}\left(\hat{\tilde{\sigma}}_{t,t_b} - \tilde{\sigma}_{t,t_b}\right) \xrightarrow{d} \mathcal{N}\left(0, \underbrace{L_{2(n+1)}\Pi_{uu,uu,t,t_b}L_{2(n+1)}' - \tilde{\sigma}_{t,t_b}\tilde{\sigma}_{t,t_b}'}_{V_2}\right),$$

for $\tilde{\sigma}_{t,t_b} = \text{vech}(\tilde{\Sigma}_{t,t_b})$. To obtain $\bar{\Omega}_{k,t}$ in $\sqrt{H}\left(\hat{\bar{\lambda}}_{k,t} - \bar{\lambda}_{k,t}\right) \xrightarrow{d} \mathcal{N}\left(0, \bar{\Omega}_{k,t}\right)$, an application of the Delta method yields $\Omega_{k,t} = \bar{J}_k\left(\tilde{\beta}_t, \tilde{\sigma}_{t,t_b}\right)\text{diag}(V_1, V_2)\bar{J}_k\left(\tilde{\beta}_t, \tilde{\sigma}_{t,t_b}\right)'$ for

$$\bar{J}_k\left(\tilde{\beta}_t, \tilde{\sigma}_{t,t_b}\right) = \left[\frac{\partial\bar{\lambda}_{k,t}}{\partial\tilde{\beta}_t'} : \frac{\partial\bar{\lambda}_{k,t}}{\partial\tilde{\sigma}_{t,t_b}'}\right].$$

The first part of $\bar{J}_K()$ is given by $\frac{\partial\bar{\lambda}_{k,t}}{\partial\tilde{\beta}_t'} = 0$ for $k = 0$ while for $k > 1$:

$$\frac{\partial\bar{\lambda}_{k,t}}{\partial\tilde{\beta}_t'} = \left(\left(\tilde{P}_{\bullet 1,t}/\left(e_2'\tilde{P}_{\bullet 1,t_b}\right)\right)' \otimes I_n\right)G_k,$$

for $G_k = \frac{\partial\,\text{vec}(C_k(\tilde{A}_t))}{\partial\tilde{\beta}_t'} = \sum_{m=0}^{k-1}\left[J(\mathbf{A}_t')^{k-1-m}\right] \otimes C_m\left(\tilde{A}_t\right)$. To obtain $\frac{\partial\bar{\lambda}_{k,t}}{\partial\tilde{\sigma}_{t,t_b}'}$ we make use of the Chain rule. First, consider the Gradient $\frac{\partial[\tilde{P}_{\bullet 1,t}', e_2'\tilde{P}_{\bullet 1,t_b}]'}{\partial\tilde{\sigma}_{t,t_b}}$. To this end, let $S_{\sigma_t}$ be selection matrix such that $\tilde{\sigma}_t = S_{\sigma_t}\tilde{\sigma}_{t,t_b}$ and $S_{\sigma_t}$ a selection matrix such that $\tilde{\sigma}_{t_b} = S_{\sigma_{t_b}}\tilde{\sigma}_{t,t_b}$. Define $S_P$ a matrix of 0 and 1's such that $[\tilde{P}_{\bullet 1,t}', e_2'\tilde{P}_{\bullet 1,t_b}]' = S_P[\text{vech}(\tilde{P}_t)', \text{vech}(\tilde{P}_{t_b})]'$. Furthermore, let $L_{n+1}$ be the elimination matrix such that $\text{vech}(\tilde{\Sigma}_t) = L_{n+1}\,\text{vec}(\tilde{\Sigma}_t)$, and $K_{mn}$ be the $mn \times mn$ commutation

matrix such that $\text{vec}(A') = \text{vec}(A)$ for $A$ any $m \times n$ matrix. Then:

$$\frac{\partial[\tilde{P}'_{\bullet 1,t}, e_2'\tilde{P}_{\bullet 1,t_b}]'}{\partial \tilde{\sigma}_{t,t_b}} = S_P \begin{bmatrix} \left(L_{n+1}\left(I_{(n+1)^2} + K_{n+1,n+1}\right)\left(\tilde{P}_t \otimes I_{n+1}\right)L'_{n+1}\right)^{-1} S_{\sigma_t} \\ \left(L_{n+1}\left(I_{(n+1)^2} + K_{n+1,n+1}\right)\left(\tilde{P}_{t_b} \otimes I_{n+1}\right)L'_{n+1}\right)^{-1} S_{\sigma_{t_b}} \end{bmatrix}.$$

Finally:

$$\frac{\partial \bar{\lambda}_{k,t}}{\partial[\tilde{P}'_{\bullet 1,t}, e_2'\tilde{P}_{\bullet 1,t_b}]} = C_k\left(\tilde{A}_t\right)\left[I_{n+1}/\left(e_2'\tilde{P}_{\bullet 1,t_b}\right) : \tilde{P}_{\bullet 1,t}/\left(e_2'\tilde{P}_{\bullet 1,t_b}\right)^2\right],$$

and hence the second part of $\bar{J}_K()$ is given by:

$$\frac{\partial \bar{\lambda}_{k,t}}{\partial \tilde{\sigma}'_{t,t_b}} = \frac{\partial \bar{\lambda}_{k,t}}{\partial[\tilde{P}'_{\bullet 1,t}, e_2'\tilde{P}_{\bullet 1,t_b}]} \times \frac{\partial[\tilde{P}'_{\bullet 1,t}, e_2'\tilde{P}_{\bullet 1,t_b}]'}{\partial \tilde{\sigma}'_{t,t_b}}.$$

With respect to the AR confidence set, the first step is to obtain the asymptotic distribution of the $(n+2) \times 1$ vector:

$$\bar{L}_{k,t} = \begin{pmatrix} C_k(A_t)\tilde{P}_{\bullet 1,t} \\ \\ e_2'\tilde{P}_{\bullet 1,t_b} \end{pmatrix}$$

for which it holds that $\bar{\lambda}_{i,t} = (e'_{1+i}L_{k,t})/(e'_{n+2}L_{k,t})$. Via the Delta Method we get: $\sqrt{H}\left(\hat{\bar{L}}_{k,t} - \bar{L}_{k,t}\right) \xrightarrow{d} \mathcal{N}(0, \Omega^{\bar{L}}_{k,t})$ for $\Omega^{\bar{L}}_{k,t} = \bar{J}^{(2)}_k\left(\tilde{\beta}_t, \tilde{\sigma}_{t,t_b}\right)\text{diag}(V_1, V_2)J^{(2)}_k\left(\tilde{\beta}_t, \tilde{\sigma}_{t,t_b}\right)'$ where

$$\bar{J}^{(2)}_k\left(\tilde{\beta}_t, \tilde{\sigma}_{t,t_b}\right) = \left[\frac{\partial \bar{L}_{k,t}}{\partial \tilde{\beta}_t} : \frac{\partial \bar{L}_{k,t}}{\partial \tilde{\sigma}_{t,t_b}}\right].$$

Similar to above, it holds that $\frac{\partial \bar{L}_{k,t}}{\partial \beta_t} = 0$ for $k = 0$ while for $k > 1$:

$$\frac{\partial \bar{L}_{k,t}}{\partial \beta'_t} = \begin{pmatrix} \left(\tilde{P}'_{\bullet 1,t} \otimes I_n\right)G_k \\ \\ 0 \end{pmatrix}.$$

Finally:

$$\frac{\partial \bar{L}_{k,t}}{\partial \tilde{\sigma}_{t,t_b}} = \frac{\partial \bar{L}_{k,t}}{\partial[\tilde{P}'_{\bullet 1,t}, e_2'\tilde{P}_{\bullet 1,t_b}]} \times \frac{\partial[\tilde{P}'_{\bullet 1,t}, e_2'\tilde{P}_{\bullet 1,t_b}]'}{\partial \tilde{\sigma}'_{t,t_b}},$$

where $\frac{\partial[\tilde{P}'_{\bullet 1,t}, e_2'\tilde{P}_{\bullet 1,t_b}]'}{\partial \tilde{\sigma}'_{t,t_b}}$ is as above and

$$\frac{\partial \bar{L}_{k,t}}{\partial[\tilde{P}'_{\bullet 1,t}, e_2'\tilde{P}_{\bullet 1,t_b}]} = \begin{pmatrix} C_k(A_t) & 0 \\ \\ 0 & 1 \end{pmatrix}.$$

# Appendix C Supplementary Monte Carlo Results

This part of the Appendix illustrates the performance of the kernel based confidence sets in large samples. Specifically, we increase sample size and kernel bandwidth by a factor of 15 and $\sqrt{15}$ respectively, interpolating coefficients from the same data-generating process described in section 3. Figure C.14 shows the true underlying impulse response functions, which display the same dynamics just over a larger time frame.

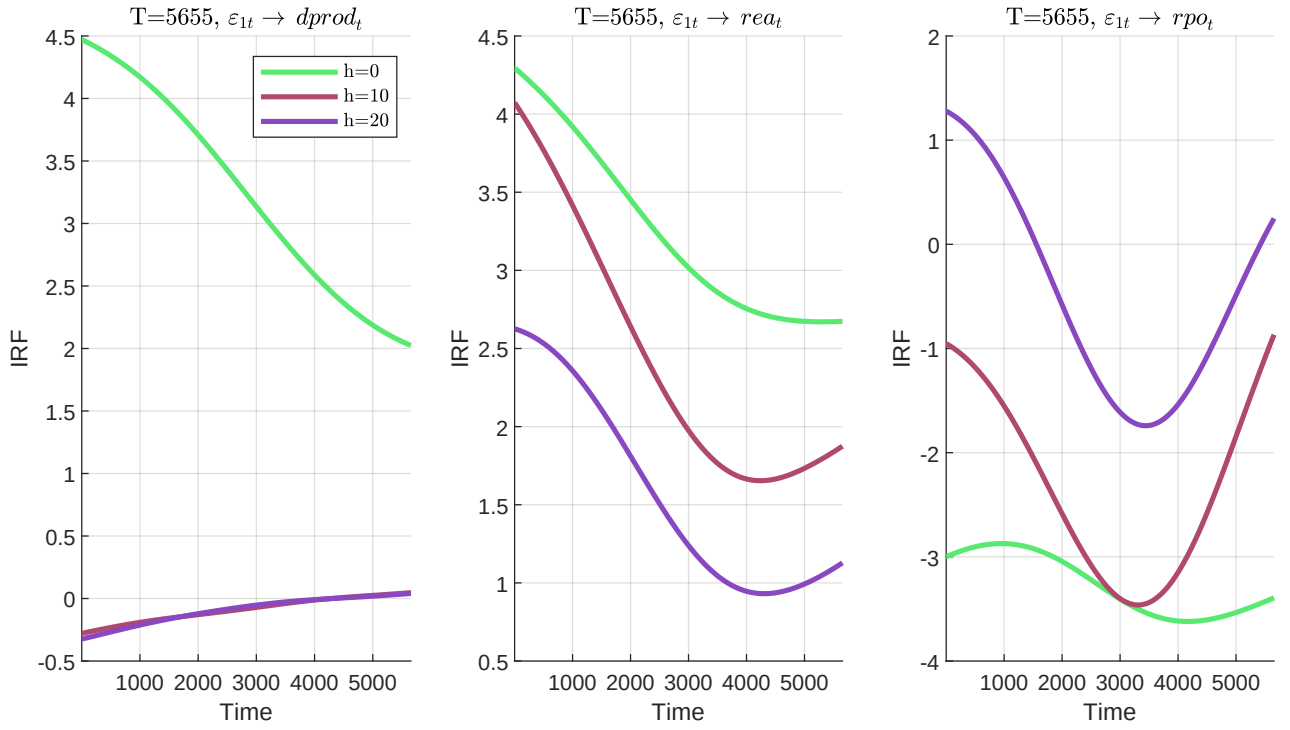


Figure C.14: True impulse response functions $\lambda^{\sigma}_{h,i,t}$.

As reported in Figure C.15, estimated empirical coverage gets very close to the nominal 95% confidence level in the DGP that considers a strong instrument. For the parameter constellation that mimics a weak instrument (Figure C.16), estimated coverage between the DM and AR confidence sets slightly diverges, with some undercoverage noticeable by the Delta Method. This suggests that under time-varying coefficients it remains particularly important to consider robust confidence sets, even in relative large samples.

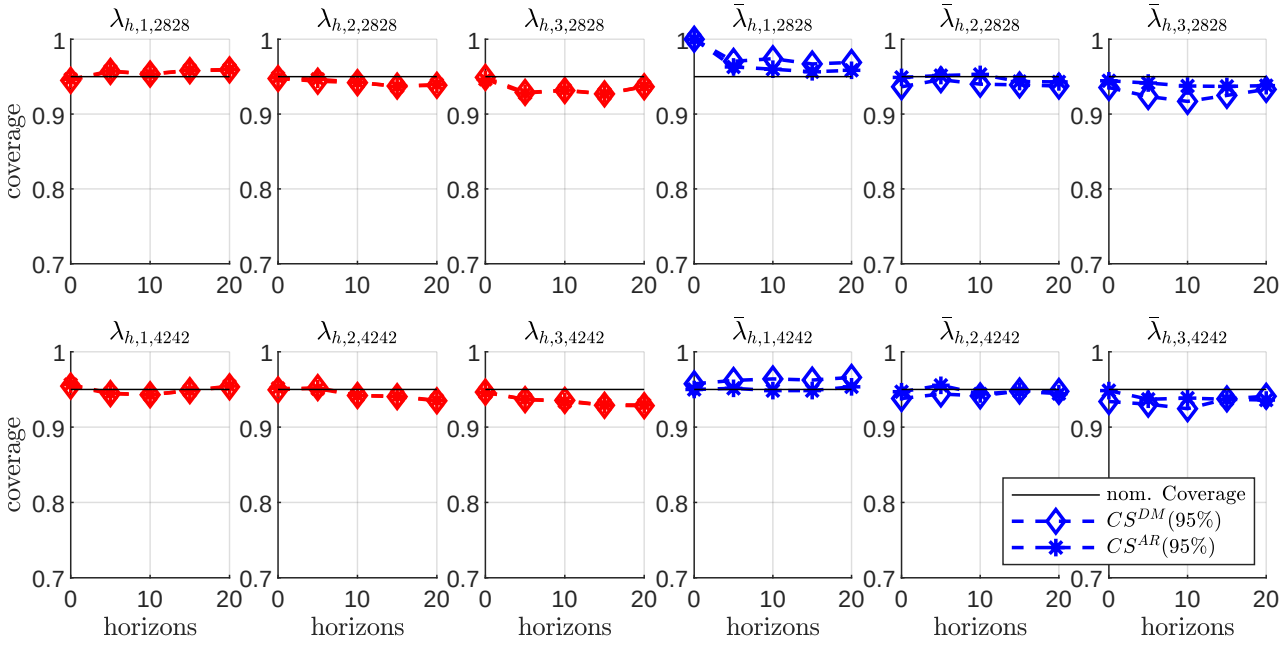


Figure C.15: Estimated empirical coverage at 95% confidence level obtained for $\lambda_{h,i,t}$ (red) and $\bar{\lambda}_{h,i,t}$ (blue) at $t = 1/4T = 2828$ (first row) and $t = 3/4T = 4242$ (second row). $\theta^{strong} = \{\phi_z = 0.86, \sigma_z = 0.06\}$. Confidence Sets (CS) based on the Delta Method (DM) are highlighted by diamonds, while Anderson Rubin confidence sets (AR) by stars.

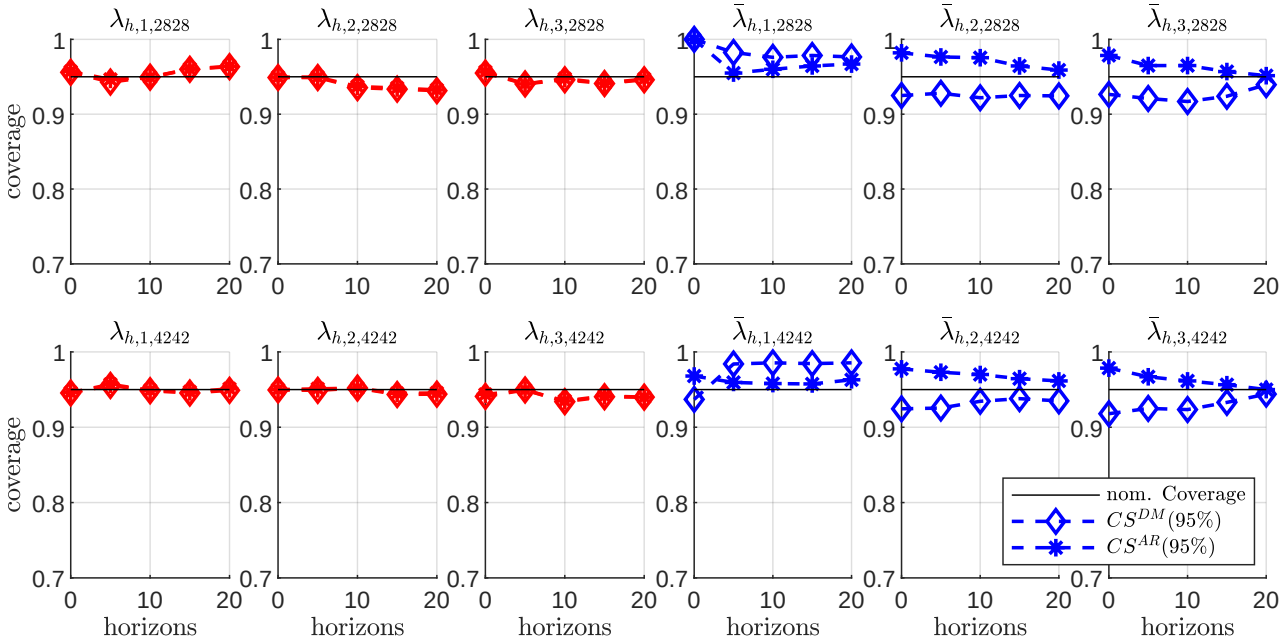


Figure C.16: Estimated empirical coverage at 95% confidence level obtained for $\lambda_{h,i,t}$ (red) and $\bar{\lambda}_{h,i,t}$ (blue) at $t = 1/4T = 2828$ (first row) and $t = 3/4T = 4242$ (second row). $\theta^{weak} = \{\phi_z = 0.48, \sigma_z = 0.71\}$. Confidence Sets (CS) based on the Delta Method (DM) are highlighted by diamonds, while Anderson Rubin confidence sets (AR) by stars.

# Appendix D Supplementary Empirical results: oil supply news shocks

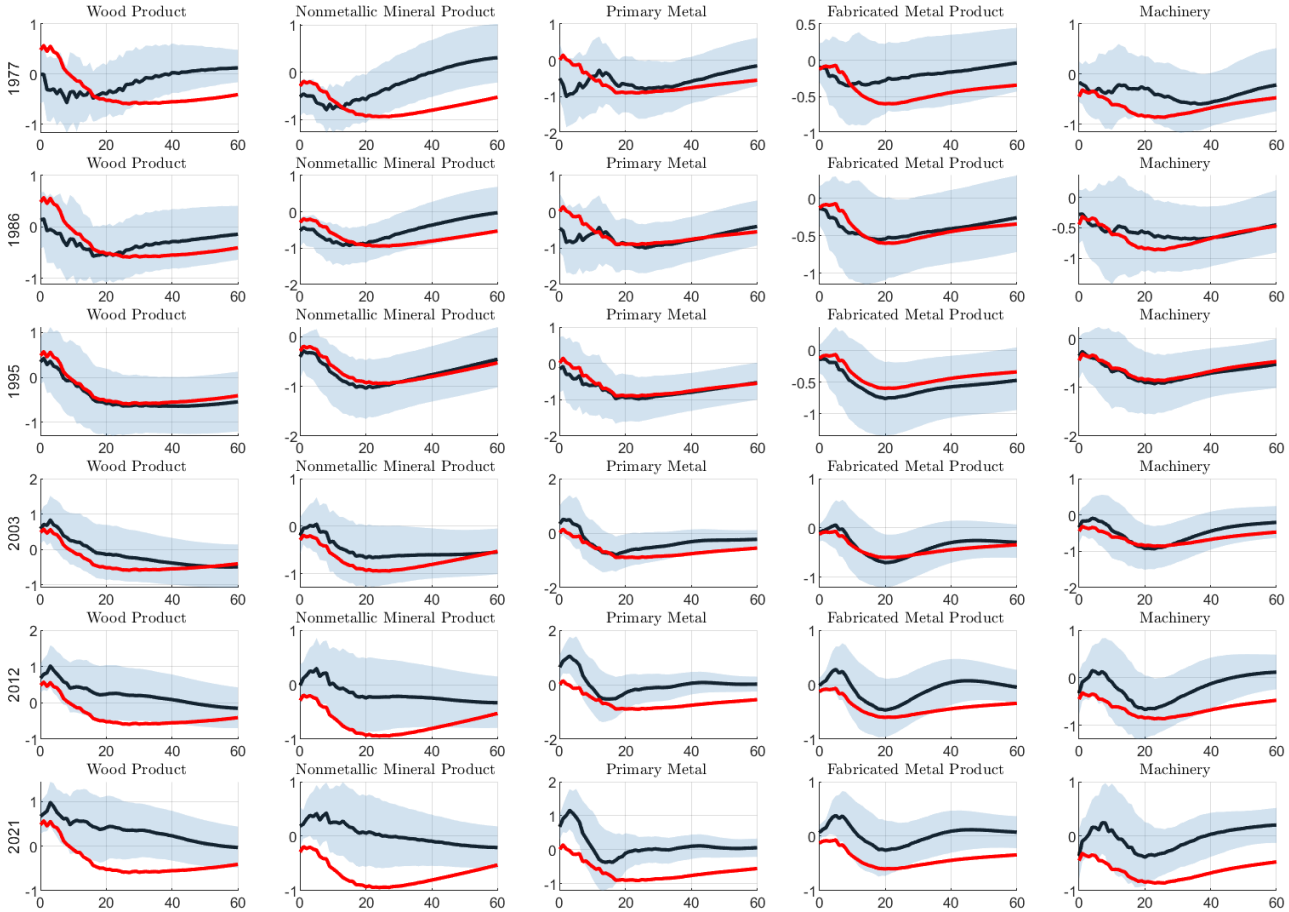


Figure D.17: Time-varying impulse response functions to an oil-supply shock of unit variance.

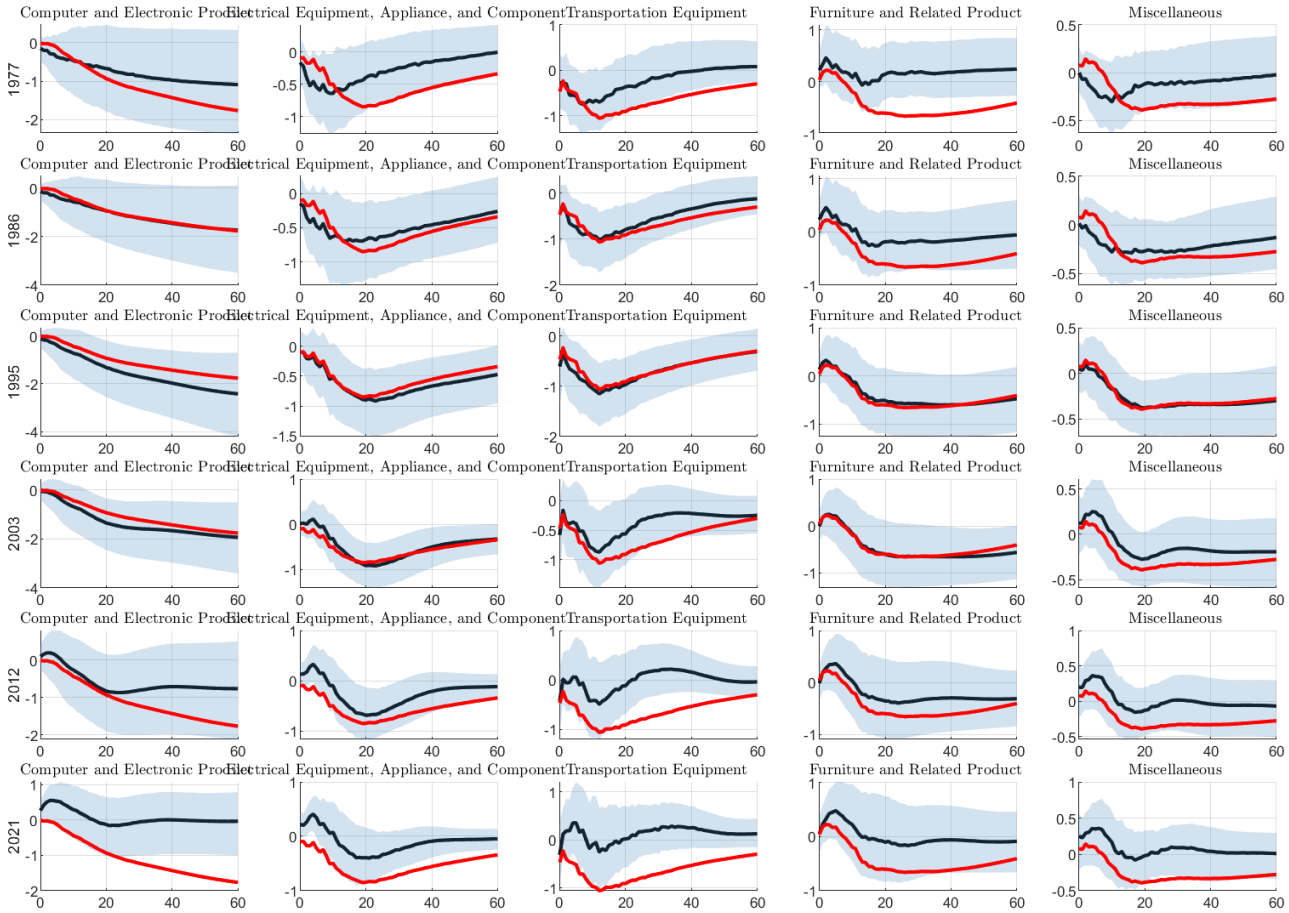


Figure D.18: Time-varying impulse response functions to an oil-supply shock of unit variance.

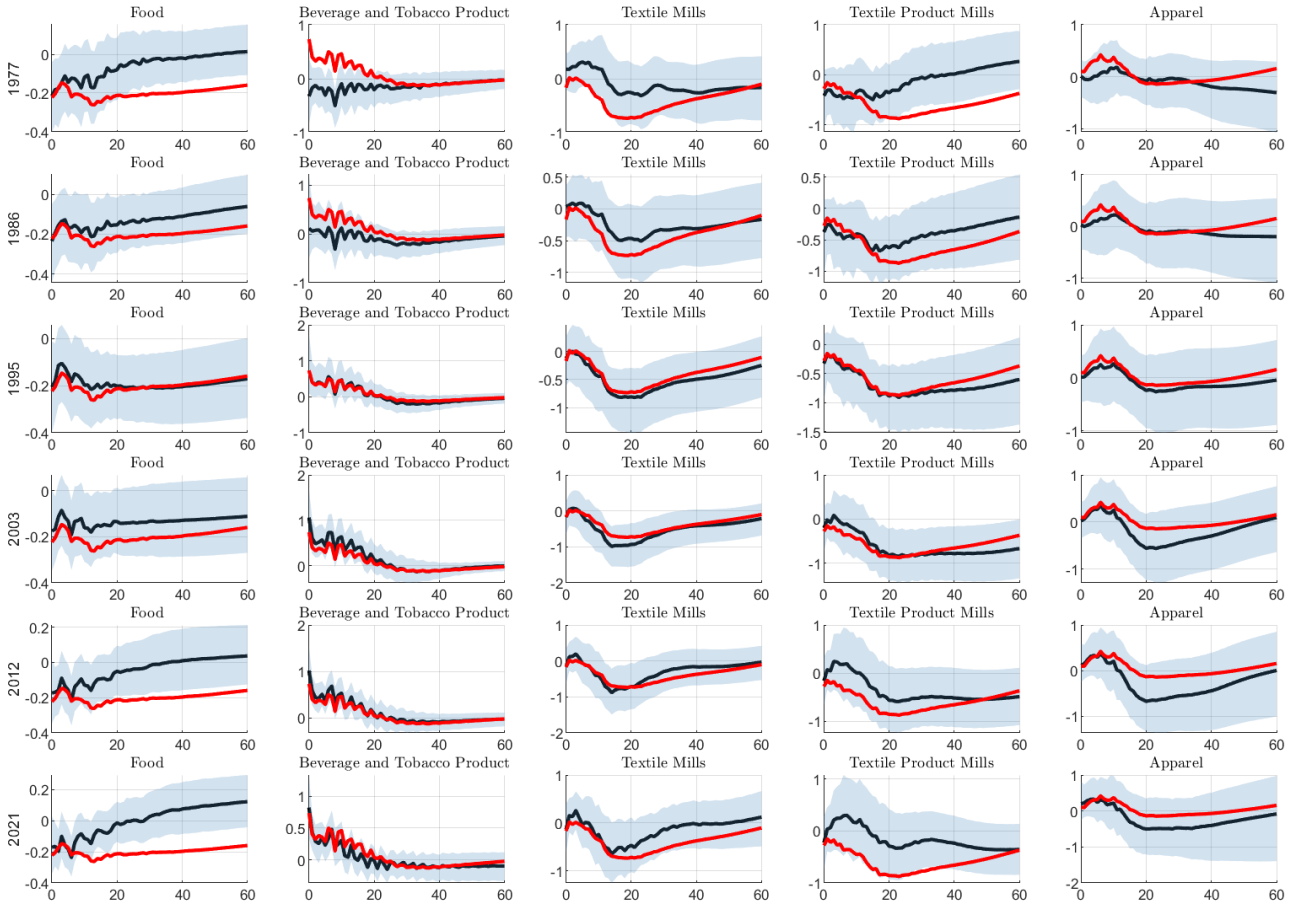


Figure D.19: Time-varying impulse response functions to an oil-supply shock of unit variance.

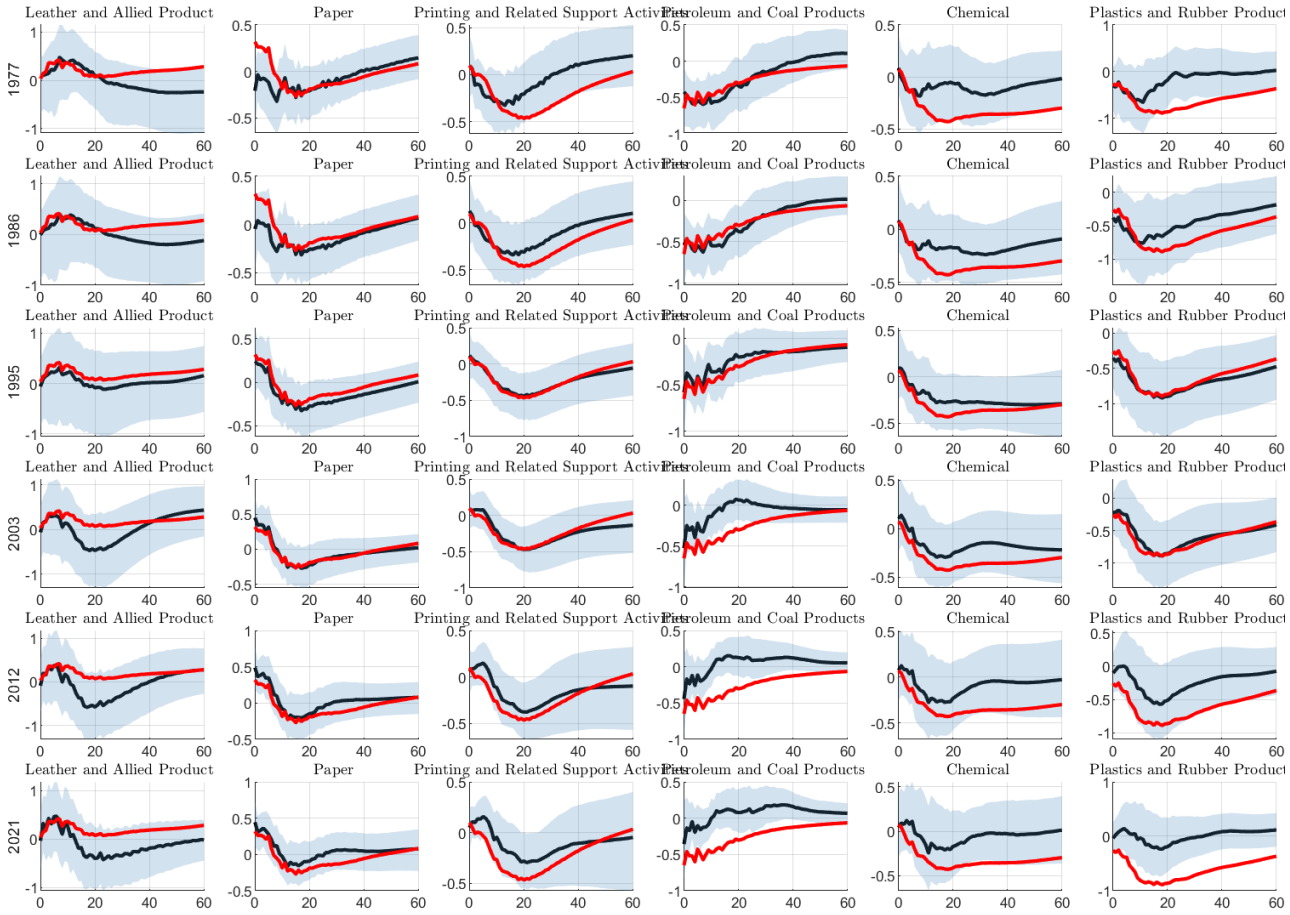


Figure D.20: Time-varying impulse response functions to an oil-supply shock of unit variance.

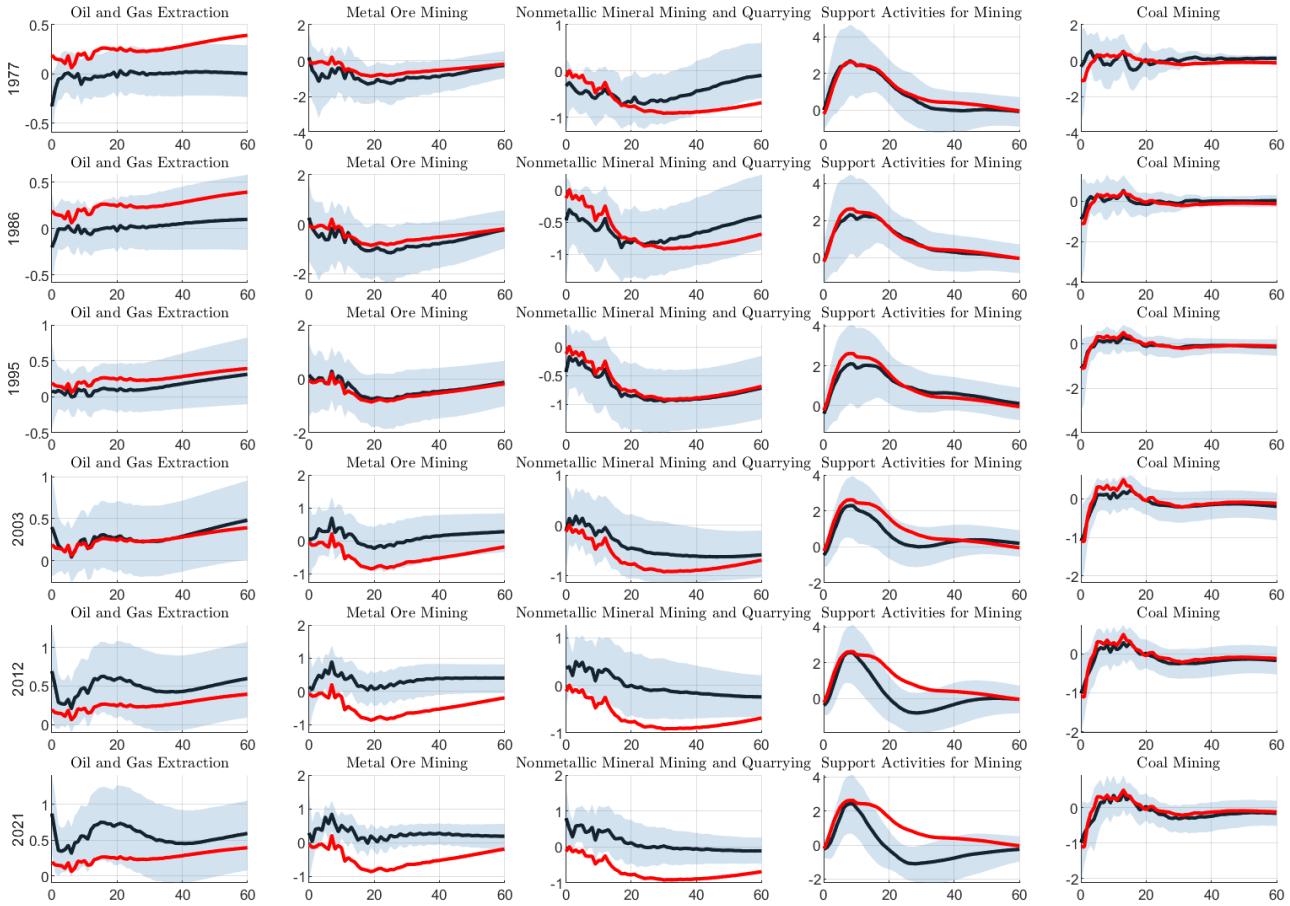


Figure D.21: Time-varying impulse response functions to an oil-supply shock of unit variance.

# Appendix E Supplementary Empirical results: UK monetary policy shocks

## E.1 The role of information effects

In this part of the appendix, we re-estimate the baseline results using a variation of the external instrument that teases out information effect. To this end, we follow Jarocinski (2021) and Kerssenfischer (2019) in exploiting the contemporaneous high frequency surprise of stock prices to construct an instrument that is robust to such information frictions. Specifically, we separate policy from information shocks via a simple sign-restricted model:

$$\begin{pmatrix} \Delta \text{GBP2Y}_t \\ \Delta \text{FTSE}_t \end{pmatrix} = \begin{pmatrix} + & + \\ - & + \end{pmatrix} \begin{pmatrix} \varepsilon_t^p \\ \varepsilon_t^i \end{pmatrix}$$

where $\Delta\text{GBP2Y}_t$ and $\Delta\text{FTSE}_t$ are the high frequency surprises in the two year Government bond rate and the FTSE all share index respectively. Effectively, the model assumes that genuine monetary policy shocks induce a negative co-movement between interest rates and stock prices, while information shocks are assumed to push them into the same direction. Following Kerssenfischer (2019), we use the median target model of Fry and Pagan (2011) to define the informational robust instrument as $z_t^{policy} = \hat{\varepsilon}_t^p$. The resulting estimates are made available in Figure E.22.

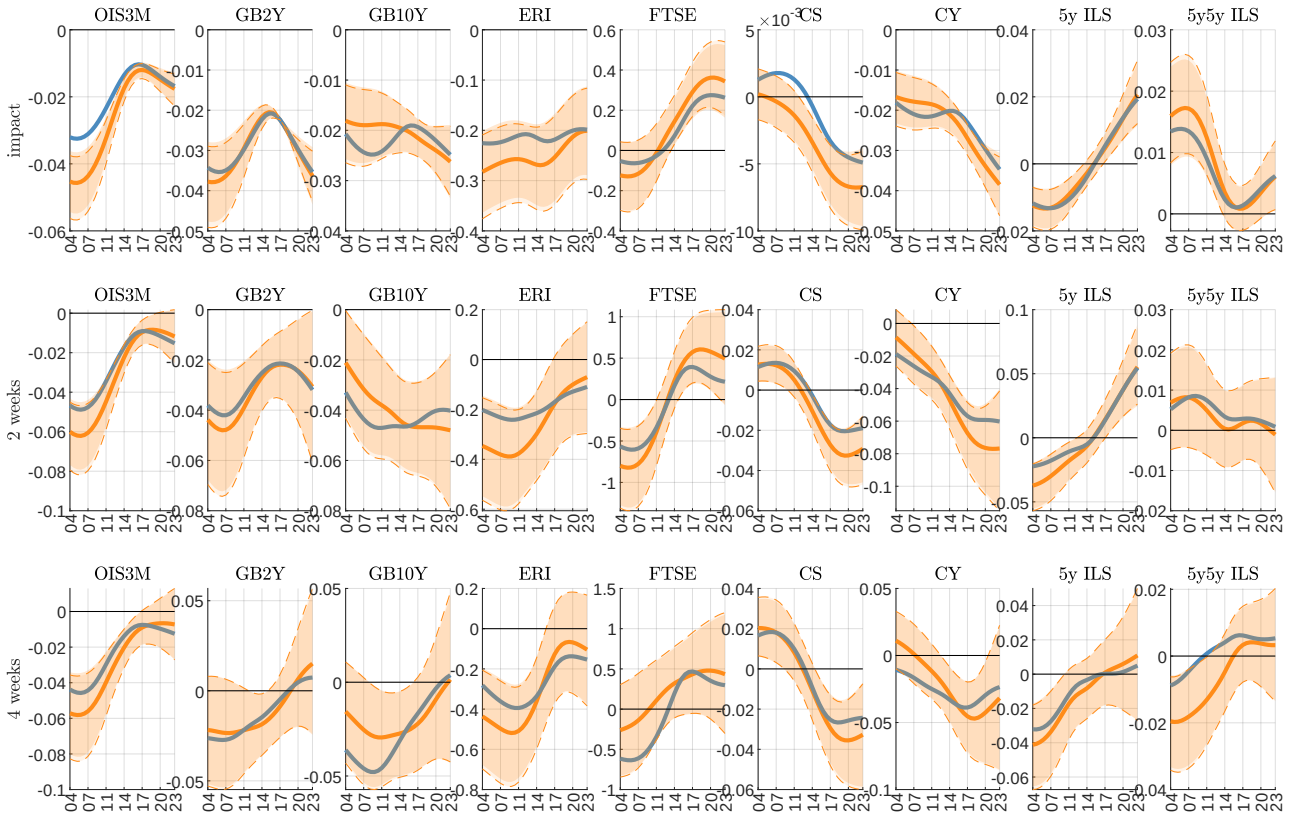


Figure E.22: Impulse response functions of an expansionary monetary policy shocks normalized to decrease the 2Y OIS rate by 2.5bp in June 2017. Here, the external instrument is net of information effects as in Kerssenfischer (2019). Shaded areas give 90% confidence intervals by the Delta Method, while dashed lines give 90% confidence sets by the AR method. The blue line gives the IRF obtained under the baseline IV.

## E.2 Identification via heteroskedasticity

Our main empirical results are based on instruments constructed using intra-daily data. Here, a small 30 minute window around the policy announcement is used to measure the unexpected news in policy, minimizing the risk of contamination by other shocks. However, it is not

completely unreasonable to think that part of the variation in $z_t$ is driven by economic news unrelated to monetary policy, that is $z_t = \varepsilon_{1,t} + \sum_{j \neq 1} \varepsilon_{j,t} + \eta_t$, where $\varepsilon_{1,t}$ is the monetary policy shock and the $\varepsilon_{j,t}$ are other economic news. To assess the robustness of our results, we re-estimate the time-varying impulse response functions using an alternative identification strategy that allows for the presence of contaminating shocks. Following Rigobon and Sack (2004), Nakamura and Steinsson (2018) and Wright (2012), we make an assumption that on central bank announcement days, the variance of the monetary policy shock is considerably larger than on a set of control dates. At the same time, the difference in variance of other structural shocks is assumed to remain broadly comparable across those set of days. Note that this doesn't rule out that the conditional variance of other structural shocks varies over time, as long as the mean variance remains comparable on announcement and control days.

Formally, let us denote the monetary policy announcement days by $\mathcal{T}_1 = \{t_{a_1}, \ldots, t_{a_M}\}$, of which there are $T_1 < T$ in our sample, and denote control dates by $\mathcal{T}_2 = \{t_{a_2}, \ldots, t_{a_J}\}$ where $T_2 \leq (T - T_1)$. Define standardized residuals $\tilde{u}_t = \Sigma_t^{\frac{1}{2}} u_t$, where $\Sigma_t^{\frac{1}{2}} (\Sigma_t^{\frac{1}{2}})' = \Sigma_t$. The covariance matrix of those standardized reduced form errors is given by $E(\tilde{u}_t \tilde{u}_t') = \Sigma_1, t \in \mathcal{T}_1$ at the announcement dates, and by $E(\tilde{u}_t \tilde{u}_t') = \Sigma_2, t \in \mathcal{T}_2$ at control dates. We assume that the variance of monetary policy shock at announcement dates, $\mathrm{E}(\varepsilon_{1t}^2) = \sigma_1^2$, is much larger than on the control dates $\mathrm{E}(\varepsilon_{1t}^2) = \sigma_2^2$, that is $\sigma_1^2 > \sigma_2^2$. The variance of the other shocks remains constant, and for simplicity, is standardized to unity. Denote by $b_i$ the $i$th column of the structural impact matrix $B$, then it holds that $\Sigma_1 = \sigma_1^2 b_1 b_1' + \sum_{i=2}^n b_i b_i'$ and $\Sigma_2 = \sigma_2^2 b_1 b_1' + \sum_{i=2}^n b_i b_i'$. Hence, the difference in the covariance matrices is given by $\Sigma_1 - \Sigma_2 = (\sigma_1^2 - \sigma_2^2) b_1 b_1'$, identifying the structural parameters of interest.

To implement the identification strategy in our IV-SVAR, we make use of the fact that identification can be thought of as an instrumental variables regression (Rigobon and Sack, 2004). Specifically, let $\hat{u}_t^{2y}$ be daily standardized forecast errors in the 2 year interest rate. Then, the

contemporaneous impact matrix can be consistently estimated using the following instrument:

$$z_t = \left( \mathbf{1}(t \in \mathcal{T}_1) \frac{T}{T_1} - \mathbf{1}(t \in \mathcal{T}_2) \frac{T}{T_2} \right) \hat{u}_t^{2y},$$

see also Lewis (2022).

Under the assumptions above, one can show that the instrument is relevant and exogenous, specifically $\mathrm{E}[z_t \varepsilon_{1t}] = b_{11}^2(\sigma_1^2 - \sigma_2^2) = \alpha$ and $\mathrm{E}[z_t \varepsilon_{jt}] = 0, j \neq 1$. Note that this translates easily to the time-varying coefficient case, but requires shock invertibility to get comparable shocks sizes given that $\mathrm{E}_t[z_t \varepsilon_{1t}] = b_{11,t}^2(\sigma_{1t}^2 - \sigma_{2t}^2) = \alpha_t$ is generally time-varying.

Figure E.23 provides estimates under identification by heteroskedasticity, using the day prior to the announcement days as the control days. Estimates are remarkably close to those obtained under an IV-SVAR (assuming invertibility), suggesting that results are fairy robust to the alternative identification scheme.

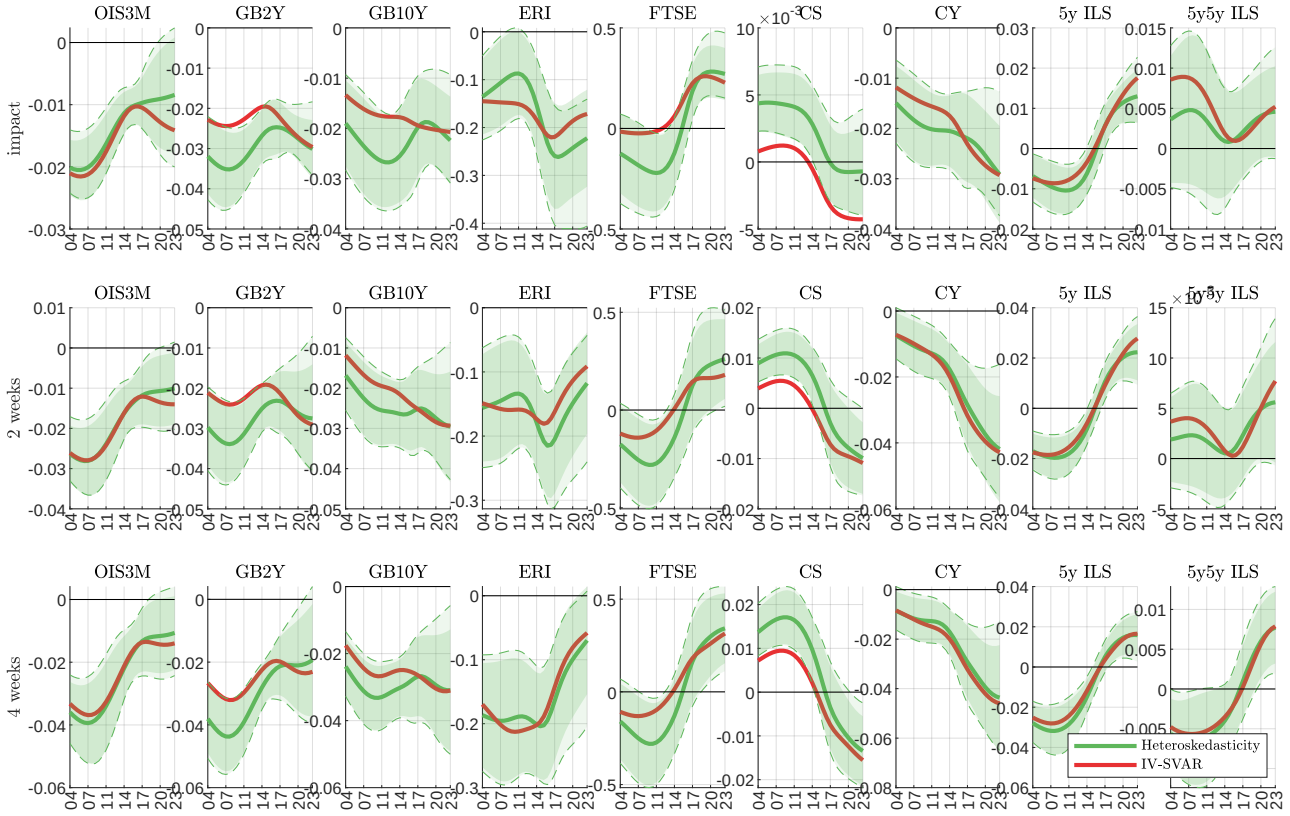


Figure E.23: Impulse response functions of an expansionary monetary policy shock of unit standard deviation obtained under identification by heteroskedasticity. Shaded areas give 90% confidence intervals by the Delta Method, while dashed lines give 90% confidence sets by the AR method.