Finance and Economics Discussion Series

Federal Reserve Board, Washington, D.C.
ISSN 1936-2854 (Print)
ISSN 2767-3898 (Online)

Nonparametric Time Varying IV-SVARs: Estimation and Inference

Robin Braun, George Kapetanios, Massimiliano Marcellino

2025-004

Please cite this paper as:
Braun, Robin, George Kapetanios, and Massimiliano Marcellino (2025). "Nonparametric Time Varying IV-SVARs: Estimation and Inference," Finance and Economics Discussion Series 2025-004. Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/FEDS.2025.004.

Nonparametric Time Varying IV-SVARs: Estimation and Inference *

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

December 9, 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 oil supply news shocks on US industrial production.

JEL classification: C14, C32, C53, C55
Keywords: Time-varying parameters, Nonparametric estimation, Structural VAR, External instruments, Weak instruments, Oil supply news 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, Mark Watson and Yiru Wang 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 to study dynamic causal effects in empirical macroeconomics, including those of monetary policy (Gertler and Karadi, 2015; Caldara and Herbst, 2019; Jarociński and Karadi, 2020), oil price shocks (Känzig, 2021) or technology shocks (Miranda-Agrippino et al., 2024) among many. 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, while Inoue et al. (2024a,b) leverage the path estimator of Müller and Petalas (2010) for a similar purpose.

In this paper we contribute to the literature of VARs identified by external instruments, developing estimators for IV-SVARs with slowly changing parameters aimed at capturing instabilities salient in macroeconomic relationships (Stock and Watson, 1996). While we take no stance on what causes the parameter changes, often discussed factors include institutional modifications, technological developments, economic trends such as globalization, or an evolving policy toolkit.

Our paper complements and extends previous work on time-varying IV-SVARs in various ways. First, instead of assuming a Gaussian process 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 and easily handles very persistent time-series.

Second, we provide results for two estimators that cater to different needs of the researchers, 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 possible to construct time-varying IRFs that remain comparable over time in response to shocks of constant scale. Shock invertibility can be tested. If rejected, it is still possible to rely on the internal instrument estimator, which allows to estimate relative IRFs consistently in the absence of invertibility. However, given that the shock scale remains unknown in the internal instrument VAR, one cannot set a comparable shock size across time without further assumptions on the relationship between the shock of interest and the external instrument (see Paul (2020)).

Our main object of interest are IRFs. 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 effec-

tive 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 include a Monte Carlo exercise. Here, we calibrate a data-generating processes based on time-varying estimates of the global oil market VAR by Kilian (2009). We are able to obtain satisfactory empirical coverage if the evolution of parameters is sufficiently smooth, particularly for the weak-IV robust confidence sets. 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 illustrate the methodology revisiting estimates of the transmission of oil supply news shocks on US industrial production (IP). Building on Känzig (2021), we study a time-varying IV-SVAR that includes monthly macroeconomic variables for the global crude-oil market, as well as US mining- and manufacturing IP. For identification, the model relies on an external instrument that leverages futures price movements around OPEC production quota announcements. A constant parameter model suggests that these shocks transmit as a cost-push shock to the US economy, where manufacturing production declines with increasing real oil prices. However, our methodology reveals strong time-variation in the estimated impulse response functions, which seems to align 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 the oil-market specific shock still transmits as cost-push shock to the US. 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 instrumental variables instead of exclusion restrictions, and use the proposed kernel based methods instead of Bayesian techniques.

Related Literature

Our paper builds on the seminal work of Cogley and Sargent (2005) and Primiceri (2005) who introduced TPV into VARs by letting coefficients evolve according to a random walk in a Bayesian setting. Paul (2020) extends their framework to achieve identification of IRFs by external instruments, including the instrument as a regressor. Inoue et al. (2024a,b) also estimate time-varying IRFs identified by IV, but rely on the frequentist path estimator of Müller and Petalas (2010). While this allows for less prior dependence, the underlying implementation still relies on a Gaussian random walk assumption to obtain point estimates and standard errors.

Our methodological approach is distinct from these papers along various dimensions. First, we rely on an entirely frequentist, nonparametric approach that leverages kernel estimators. This approach has several practical advantages. First, it avoids a parametric choice for the law of motion underlying the SVAR parameters and instead relies on nonparametric smoothness conditions. Second, it is computationally simple and can handle very large datasets as well as persistent time series. In such circumstances, existing methods may struggle as they require to loop trough each observation and potentially need to deal with non-stationary draws implying explosive IRFs. Third, similar to a constant parameter VAR, our framework provides very simple formulations for standard errors. We leverage those to conduct robust inference valid under weak instruments (see also Inoue et al. (2024a)). Finally, unlike previous papers, we cover both the standard IV-SVAR model as well as the internal IV estimator (Plagborg-Møller and Wolf, 2021).[1] The theory behind the 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

[1]The estimators by Paul (2020) and Müller and Petalas (2010) are compared with our proposal in Appendix E, where we find that with simulated data they perform similarly, while with actual data our method seems better capable of handling very persistent time series.

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.

We note that we are not the first to leverage kernel-based estimators to introduce TVP into VARs, see e.g. Kapetanios et al. (2019) and Hipp (2020). However, unlike these papers we focus on external instrument identification. Finally, we would also like to relate our paper to Amir-Ahmadi et al. (2023) who also allow for a time-varying relationship between the instrument and the structural shock of interest. Unlike their paper, however, we allow the IRFs to be time-varying.

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 transmission of oil supply news shocks on US industrial production, and section 5 recapitulates. Proofs and additional empirical results are gathered in the online supplementary material. This also includes a comparison of our methodology to a Bayesian- and a path estimator.

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 the shock size across time, and inference of reduced form quantities via kernel based methods. Finally, we show how the results can be leveraged to compute confidence sets of IRFs.

2.1 Identification of VAR impulse response functions via external instruments

Consider the n-variate SVAR(p) model given by:

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

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

where $y_t = (y_{1t}, \ldots, y_{nt})'$ is a $n \times 1$ vector of endogenous time series, ν is a $n \times 1$ vector of intercepts, $A_i, i = 1, \ldots, p$ are $n \times n$ matrices of autoregressive coefficients and the error terms u_t and ε_t are, for simplicity, assumed to be i.i.d. white noise with covariance matrix Σ_t and I_n respectively. Equation (1) describes the reduced form VAR dynamics of y_t as a function of lagged realizations and a vector of $n \times 1$ error terms u_t with full covariance matrix Σ. Equation (2) relates the prediction errors u_t to $n \times 1$ structural shocks ε_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 ε_t. For the moment, we assume that the model is stable, which implies that the SVAR(p) has a MA(∞) 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(∞) 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 elements in the MA(∞) 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 ith 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 ε_{1t} on variable i, k periods ahead.

It is important to note that, without further assumptions, 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 $\Sigma_u = BB' = BQQ'B = \tilde{B}\tilde{B}'$. In 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. $E(z_t\varepsilon_{1,t}) = \alpha \neq 0$,
2. $E(z_t\varepsilon_{j,t}) = 0$ *for* $j \neq 1$.

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

$$\Gamma = \mathrm{E}\left(u_t z_t\right) = B \begin{pmatrix} \mathrm{E}[\varepsilon_{1,t} z_t] \\ \mathrm{E}[\ \varepsilon_{2:n,t} z_t] \end{pmatrix} = \left[\begin{array}{c|c} b_1 & b_2 \end{array}\right] \begin{pmatrix} \alpha \\ 0 \end{pmatrix} = \alpha b_1,$$

where $\varepsilon_{2:n,t} = [\varepsilon_{2t}, \ldots, \varepsilon_{nt}]'$. In words, the correlation between the external instrument and reduced form prediction error is proportional to the first column of the impact matrix b_1.

In order to obtain interpretable magnitudes, there are two popular approaches to normalize IRFs in IV-SVARs. The first is known as the unit shock standardization (Stock and Watson, 2016) or *relative* IRFs. Here, the shock variance is re-normalized to yield IRFs that increase the first variable by unity on impact (say $\tilde{b}_{11} = 1$). 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:

$$\tilde{\lambda}_{k,i} = e_i' C_k(A)\Gamma / e_1'\Gamma. \tag{4}$$

The second approach is to normalize the standard deviation of the shock to unit variance ($\text{Var}(\varepsilon_{1,t}) = 1$), yielding *absolute* IRFs. Here, one is required to incorporate additional information of the reduced form covariance matrix Σ to recover α and hence b_1. 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 *absolute* 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: $\tilde{\lambda}_{k,i}$ does not 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 $\tilde{\lambda}_{k,i}$, even if invertibility does not hold. Specifically, for $\tilde{y}_t = [z_t, y_t']'$, the resulting *internal instrument* VAR model reads:

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

and robust relative IRFs are obtained by $\tilde{\lambda}_{k,i} = e_{1+i}' C_k\left(\tilde{A}\right)\tilde{\Gamma}/(e_2'\tilde{\Gamma}_t)$ for $\tilde{\Gamma} = e_1'\text{chol}(\tilde{\Sigma})$. On the other hand, without invertibility it is no longer possible to identify absolute IRFs ($\lambda_{k,i}$).[2] 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

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

impulse response functions that remain comparable over time.

2.2 Introducing time-varying coefficients

Introducing time-varying coefficients into the SVAR reads:

$$y_t = A_{1t}y_{t-1} + A_{2t}y_{t-2} + \ldots + A_{pt}y_{t-p} + B_t\varepsilon_t, \qquad \varepsilon_t \sim (0, I_n) \tag{7}$$

where $\mathrm{E}_t(\varepsilon_t) = 0$, $\mathrm{E}_t(\varepsilon_t\varepsilon_t') = I_n$ and $\mathrm{E}_t(u_t u_t') = \Sigma_t = B_t B_t'$. Also, let $\Gamma_t = \mathrm{E}_t(z_t u_t)$ and update the IV assumption to the time-varying case:

Assumption 2 (External Instrument)**.** *Let z_t be an instrument for the first shock. The stochastic process $\{(\varepsilon_t, z_t)\}_{t=1}^{\infty}$ satisfies:*

1. $E_t(z_t\varepsilon_{1,t}) = \alpha_t \neq 0$,
2. $E_t(z_t\varepsilon_{j,t}) = 0$ *for* $j \neq 1$.

At this point, one approach would be to impose a specific parametric 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) or quasi likelihood methods as Müller and Petalas (2010). Instead, in this paper we follow a nonparametric 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 3. *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:*

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

Assumption 3 states that the model parameters are bounded and that changes to those parameters are restricted to be small. The rate is assumed 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, though allowing for smooth stochastic processes is also feasible. The theory in all existing work (such as, e.g., Giraitis et al. (2018)) has been developed with a single rate of change for all parameter processes. As a result, we align with this setting. Extending the analysis to different rates for different parameters is possible but is outside the scope of the paper. Assumption 3 enables consistency and rate results as presented in Theorem 1-3. The results are similar in nature to those presented in, e.g., Giraitis et al. (2018). The assumption requires that the deviation in parameters is shrinking with T, so the idea behind consistency is close to an infill asymptotic setup. Under Assumption 3, Giraitis et al. (2018) show that the MA(∞) representation can be expressed 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 3, the MA(∞) 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 2, time-varying IRFs to a shock of size one standard deviation are given by:[3]

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

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:

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

effectively measuring IRFs to a re-normalized target shock $\tilde{\varepsilon}_{1,t}$ with variance $b_{11,t}^2$.

[3]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(∞) representation.

In a time-varying parameter VAR, the choice of IRF normalization is not just a matter of preference. To unpack the differences, consider a toy model of supply and demand for quantities and prices $y_t = [q_t, p_t]'$:

$$\begin{aligned} &\text{supply:} && q_t = \alpha_t p_t + \sigma_{1t}\varepsilon_{1t} \\ &\text{demand:} && q_t = \beta_t p_t + \sigma_{2t}\varepsilon_{2t} \end{aligned} \qquad \begin{pmatrix} \varepsilon_t^s \\ \varepsilon_t^d \end{pmatrix} \sim (0, I_2),$$

where α_t and β_t are supply- and demand elasticities, respectively, while $[\sigma_{1t}, \sigma_{2t}]$ are the volatilities of the structural shocks. In this case, $A_t = \begin{pmatrix} 1 & -\alpha_t \\ 1 & -\beta_t \end{pmatrix}$, and

$$B_t = A_t^{-1}\text{diag}([\sigma_{1t}, \sigma_{2t}]') = \begin{pmatrix} -\sigma_{1t}\beta_t/(\alpha_t - \beta_t) & \sigma_{2t}\alpha_t/(\alpha_t - \beta_t) \\ -\sigma_{1t}1/(\alpha_t - \beta_t) & \sigma_{2t}/(\alpha_t - \beta_t) \end{pmatrix}.$$

Assume the availability of an instrument for the supply shocks. Then, the absolute IRFs $\lambda_{k,i,t}$ leverage estimates of Γ_t and Σ_t to identify b_{1t}, the first column of B_t. This IRF is a function of both, time-varying elasticities and volatilities facilitating interpretation by holding constant the scale of the shock to unity throughout time. For example, in the case of a monetary policy shock, this means that estimates of $\lambda_{k,i,t}$ will not only reflect variation in the elasticities but also shock volatility, which summarizes how successful a central bank is in steering interest rates. This may be time-varying for many reasons, e.g. the design of new policies or temporary policy constraints such as the zero lower bound. Relative IRFs, on the other hand, yield $\Gamma_t/e_1'\Gamma_t = \begin{pmatrix} 1 & 1/\beta_t \end{pmatrix}$ in our toy-model, and hence only depend on the structural elasticities. In this case, the IRF corresponds to a shock with a time-varying scale $\tilde{\varepsilon}_{1t} \sim (0, \sigma_{1t}^2 (\beta_t/(\alpha_t - \beta_t))^2)$. Depending on the application at hand, this might still be a very useful quantity to study. It is also worth noting that unlike absolute IRFs, relative IRF does not depend on the shock volatilities. Hence, smoothness as stated in Assumption 3 is not necessary for σ_t, and our inference procedures for relative IRFs could be adjusted to allow for other forms of heteroskedasticity in

structural shocks, e.g. as in Gonçalves and Kilian (2004), but we do not consider this in the current paper.

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 relative IRFs to a fix shock size 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 as in Assumption 1, 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 a constant shock size over time:

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

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 remains comparable across time. Note that this is generally not the case if α_t itself was subject to time-variation.

Summing up our discussion, once time-variation is introduced into the model, a trade-off arises. Under invertibility, an IV-SVAR allows to identify the scale of the shock throughout the sample and study absolute IRFs of a constant shock size ($\lambda_{k,i,t}$). Whenever invertibility does not hold, unit shock IRFs based on a internal instrument VAR may be a useful alternative. While relative IRFs require less assumption on the smoothness of shock volatilities, they requires stronger assumptions on α_t to obtain (relative) IRFs that remain comparable across time ($\tilde{\lambda}_{k,i,t,t_b}$). 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, proceeding with the IV-SVAR estimator may be preferable as it requires minimal assumptions on α_t. However, if shock invertibility is rejected,

informative results under the stronger assumption $\alpha_t = \alpha$ may still be obtained based on the internal instrument VAR. 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 $\tilde{\lambda}_{k,i,t,t_b}$ we proceed in two steps. We start deriving the joint asymptotic distribution of kernel-based estimators of the reduced form parameters A_t, Γ_t, Σ_t and Σ_{t_b}, both for the IV-SVAR and internal instrument VAR. 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.

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:

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

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

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

where $\hat{u}_j = y_j - (I_n \otimes x_j')\hat{\beta}_t$ and $w_{t,j}(H) = K(|t-j|/H)$ 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 on the error term and kernel:

Assumption 4. *$\varepsilon_t = (\varepsilon_{1t}, \cdots, \varepsilon_{nt})'$ is an iid process such that $E[\varepsilon_{i1}^4] < \infty$. z_t is a stationary, α-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 5. *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

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

Concerning Assumption 4, two remarks are in order. First, as in previous work on time-varying regressions using kernel estimators, we assume the errors to be iid processes. Contemporaneous work by Giraitis et al. (2024a) shows that this can be relaxed to allow for martingale difference processes. However, a rigorous treatment is technically demanding and we do not think that, for the purposes of our analysis, much would be gained. Second, Assumption 4 precludes the presence of serial correlation in the error term, and the associated need for potential heteroskedasticity and autocorrelation robust (HAR) corrections. We note three things in this context. First, it is reasonable to expect that the presence of enough lags in the VAR model will soak up serial correlation. Second, one can test for remaining residual correlation, although recent work by Giraitis et al. (2024b) notes several problems with such tests, typically associated with over-rejection. Finally, we are not aware of any rigorous work on HAR procedures for time varying models. Such work would want to account for time variation in the correction and, while this is a very interesting topic of research, we consider it beyond the scope of the current paper.

Concerning Assumption 5, one suitable choice that we adapt is the 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 3-5 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 $\Pi_{x,t} = plim_{T\to\infty} \frac{1}{H}\sum_{j=1}^{T} w_{j,t} x_j x_j'$, $\Pi_{xz,t} = plim_{T\to\infty} \frac{1}{H}\sum_{j=1}^{T} w_{j,t} z_j x_j$, $\Pi_{ww,t} = plim_{T\to\infty} \frac{1}{H}\sum_{j=1}^{T} w_{j,t}^2 \xi_j \xi_j'$, $\xi_j = [\text{vec}(x_j u_j)', (z_j u_j - \Gamma)', \text{vec}\left(u_j' u_j - \Sigma_t\right)']'$, *and* S_σ *such that* $\text{vech}(\Sigma_t) = S_\sigma \text{vec}(\Sigma_t)$.

It is worth mentioning that the bandwidth H assumed in Theorem 1 is strictly related to the amount of time-variation permitted in Assumption 3. If a different rate is assumed in Assumption 3, say $T^{-\gamma}$ rather than T^{-1}, a different H should be used (the more time variation, the smaller the optimal H). While our theoretical results imply certain conditions for the bandwidth ($H = o(T^{\frac{1}{2}})$), it is reasonable to allow for larger bandwidths in practice, for smaller sample sizes that are common in fields such as macroeconomics. From a practical point of view, since γ 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 nonparametric 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 $\tilde{\beta}_t = \text{vec}\left(\left[\tilde{A}_{1t}, \ldots, \tilde{A}_{pt}\right]\right)$ and $\tilde{P}_t = \text{chol}(\tilde{\Sigma}_t)$ 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 $\tilde{\lambda}_{k,i,t,t_b} = e_{1+i}' C_k\left(\tilde{A}_t\right) \tilde{P}_{\bullet 1,t} / (e_2' \tilde{P}_{\bullet 1,t_b})$, 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 3-5 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 jointly asymptotically normal and asymptotically independent of each other. Their respective distributions are given by*

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

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

Under an additional normality assumption for the errors, the asymptotic variance of $\hat{\tilde{\sigma}}_t$ further reduces to $2D_{n+1}^{+}\left(\tilde{\Sigma}_t \otimes \tilde{\Pi}_{uu,t}\right) D_{n+1}^{+'}$, 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 $\tilde{\lambda}_{k,i,t,t_b}$ 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 $\tilde{\lambda}_{k,i,t,t_b}$, 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 3-5 hold and $H = o(T^{\frac{1}{2}})$. Define $w_t(H) = [w_{t,1}(H), \ldots, w_{t,T}(H)]'$*

and $\tilde{\sigma}_{t,t_b} = \text{vech}(\tilde{\Sigma}_{t,t_b})$. *Let* $\Pi_{uu,uu,t,t_b} = plim_{T\to\infty} \frac{1}{H} \sum_{j=1}^{T} \text{vec}(\tilde{\xi}_{wj}\tilde{\xi}'_{wj}) \, \text{vec}(\tilde{\xi}_{wj}\tilde{\xi}'_{wj})'$ *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 $2D^{+}_{2(n+1)} \left(\tilde{\Sigma}_{t,t_b} \otimes \tilde{\Pi}_{uu,t,t_b} \right) D^{+\prime}_{2(n+1)}$, 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:

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

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 α under the null hypothesis and hence remains valid even under asymptotically weak instruments, that is if $\alpha \to 0$.

Starting with the Delta Method, its 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_{θ_t} denotes the joint distribution

of the IV-SVAR reduced form parameters which we give in Theorem 1. Furthermore, $J_k\left(\beta_t, \Gamma_t, \sigma_t\right)$ denotes the derivative of $\lambda_{k,t}$ with respect to the reduced form parameters. Similarly, for relative IRFs it is

$$\sqrt{H}\left(\hat{\tilde{\lambda}}_{k,t,t_b} - \tilde{\lambda}_{k,t,t_b}\right) \xrightarrow{d} \mathcal{N}\left(0, \tilde{\Omega}_{k,t,t_b}\right),$$

where $\tilde{\Omega}_{k,t,t_b} = \tilde{J}_k\left(\tilde{\beta}_t, \tilde{\sigma}_{t,t_b}\right) \tilde{V}_{\theta_{t,t_b}} \tilde{J}_k\left(\tilde{\beta}_t, \tilde{\sigma}_{t,t_b}\right)'$ for $\tilde{V}_{\theta_{t,t_b}}$ being elements of the joint asymptotic covariance matrix stated in Theorem 2 and Corollary 3, and $\tilde{J}_k()$ is the gradient of $\tilde{\Omega}_{k,t,t_b}$ 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. This may be particularly true in TVP models where the effective sample size is fairly small. Hence, we also cover weak identification robust confidence sets. Consider the $(n+1) \times 1$ vectors L and the $(n+2) \times 1$ vector $\tilde{L}_{k,t}$:

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

for which it holds that $\lambda_{k,i,t} = (e_i' L_{k,t})/(e_{n+1}' L_{k,t})$ and $\tilde{\lambda}_{k,t,t_b} = (e_{1+i}' \tilde{L}_{k,t,t_b})/(e_{n+2}' \tilde{L}_{k,t,t_b})$. 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 $\sqrt{H}\left(\hat{\tilde{L}}_{k,t,t_b} - \tilde{L}_{k,t,t_b}\right) \xrightarrow{d} \mathcal{N}(0, \Omega_{k,t,t_b}^{\tilde{L}})$ 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 ijth element of $\hat{\Omega}_{k,t}^L$. 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, in the strong instrument case, Montiel-Olea et al. (2021) prove that the AR confidence set converges to Delta Method implied confidence intervals. For those reasons, we generally recommend to use the weak-IV robust confidence intervals.

Third, we note that for relative IRFs $\tilde{\lambda}_{k,t,t_b}$, both the Delta method and AR CS depend on the choice of t_b. Montiel-Olea et al. (2021) show that the $100\%(1-a)$ AR confidence set is finite only if the Wald test statistic for $e_2'\tilde{P}_{\bullet 1,t}$ is above its corresponding critical value ($\chi^2_{1,1-a}$). For this reason, we recommend setting t_b to a point in the sample where the Wald-test statistic is very high, which can be done fairly automatic way.

Finally, we note that this paper covers the simplest case of a single instrument identifying one target shock. However, the reduced form results hold for a more general case of k instruments. Hence, it is possible to extend the results in this subsection to the case of r target shocks and r instruments, which requires additional restrictions, e.g. as employed in Mertens and Montiel Olea (2018). We point the interested reader to the supplementary appendix of Montiel-Olea et al. (2021), discussing how robust confidence sets can be constructed in this case.

2.5 Bandwidth selection

To control for the amount of time-variation in any given application, the user is required to set a bandwidth H prior to estimation. 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 conditional forecasts, we propose to evaluate the models out-of-sample predictive performance comparing data realizations y_{t+h} to conditional forecasts $\hat{y}_{t+h}(H) = E_t[y_{t+h}|\tilde{y}_1, \ldots, \tilde{y}_t, z_{t+1}, \ldots, z_{t+h}]$. 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).[4] 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 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.

[4] 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 $w_{t,j}(H) = 0, j > t$.

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 the TVP VAR kernel estimators to 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:

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

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, 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 $\tilde{\lambda}_{k,i}$ in the fixed parameter case. However, as we will document, these parameter constellations do not necessarily translate to strong and weak instrument dynamics for absolute IRFs $\lambda_{k,i}$, since it leverages all the information coming from Γ and Σ exploiting the 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.

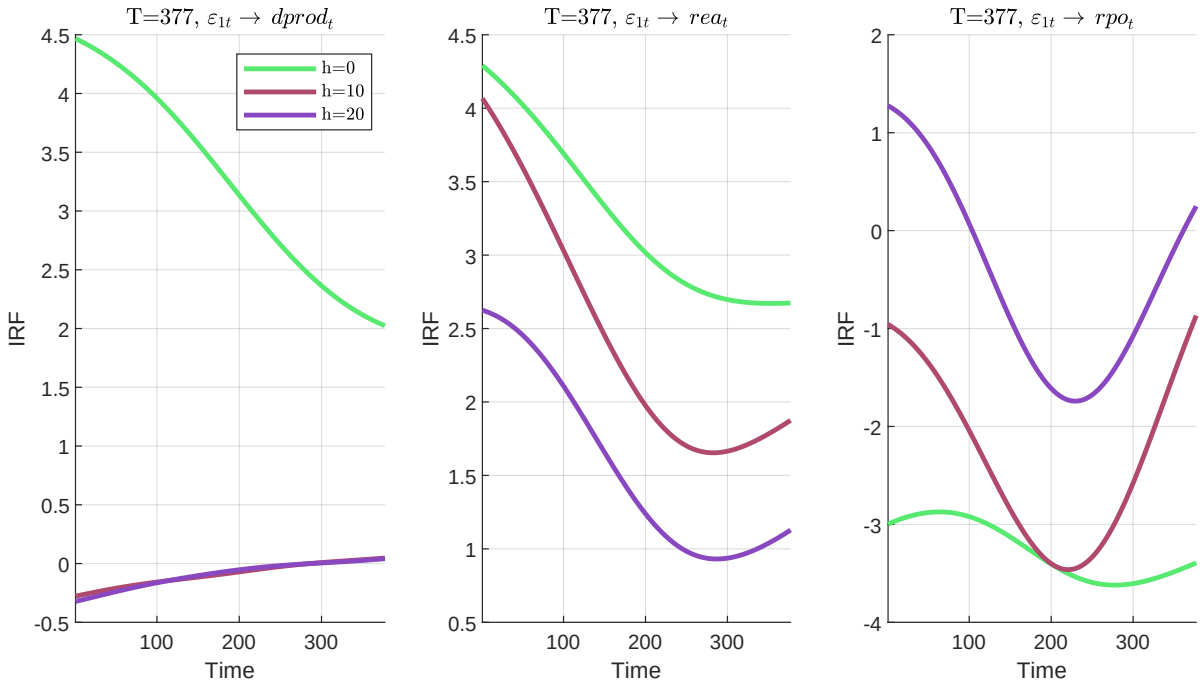

Figure 1: True impulse response functions $\lambda_{h,i,t}$ for horizons $h = 0$, $h = 10$ and $h = 20$.

3.2 Empirical coverage

We proceed simulating a total of 5000 datasets from the DGP, each of sample size $T = 377$. Empirical coverage at 95% confidence level is then computed for TVP kernel estimates of $\lambda_{h,i,t}$ (IV-SVAR) and $\tilde{\lambda}_{h,i,t,t_b}$ (internal instrument VAR). For the latter, IRFs are re-standardized to increase the first variable by one at the fixed time point $t_b = T/2$, and for ease of readability we drop the subscript in the remainder of this section. 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, to keep the discussion simple, we focus on empirical coverage at two points of time, $t = T/2$ and $t = 3/4T$.

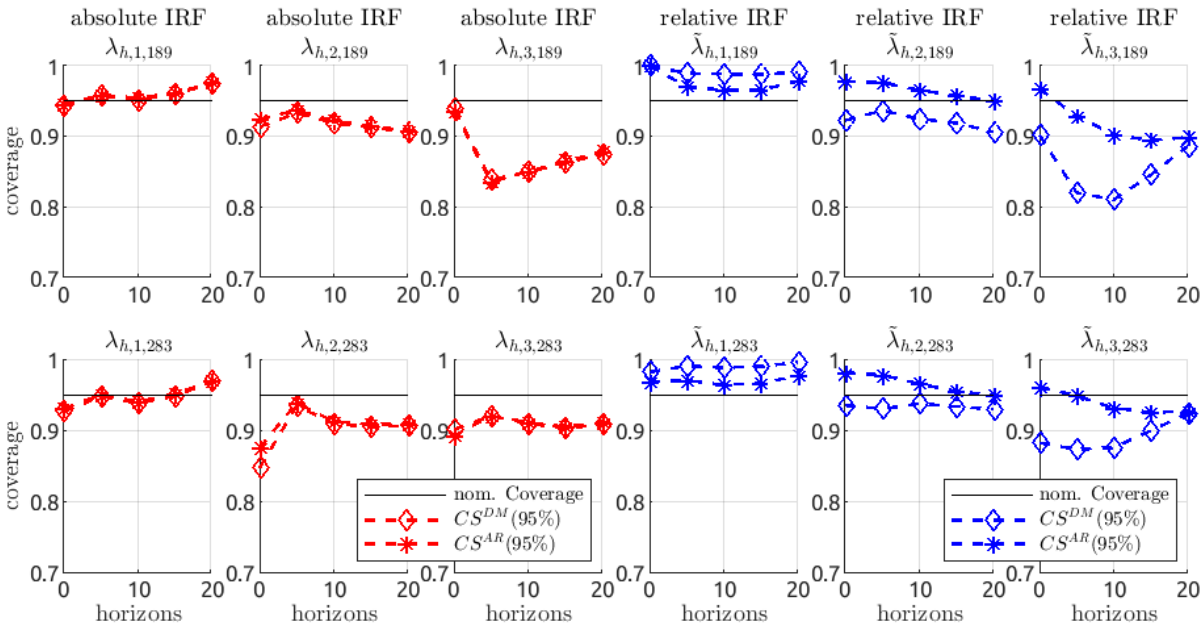

Figure 2: Estimated empirical coverage at 95% confidence level obtained for $\lambda_{h,i,t}$ (red) and $\tilde{\lambda}_{h,i,t}$ (blue) at $t = 1/2T = 189$ (first row) and $t = 3/4T = 283$ (second row). $\theta^{strong} = \{\phi_z = 0.86, \sigma_z = 0.06\}$ and H is known. Confidence Sets (CS) based on the Delta Method (DM) are highlighted by diamonds, while Anderson Rubin confidence sets (AR) by stars.

Figure 2 shows simulation results under the strong instrument parameters setting and known bandwidth. Regarding absolute IRFs $\lambda_{h,i,t}$ (red), we document that the Delta Method (DM) and Anderson Rubin (AR) confidence sets (CS) 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 performance we document is for $\lambda_{h,3,t}$ at $t = T/2 = 189$ for horizons $h = 5$ and $h = 10$, where coverage is about 85%. With respect to estimates of relative IRFs $\tilde{\lambda}_{h,i,t}$ (blue), we document that the DM- and AR CS provide somewhat different empirical coverage. This suggests that the instrument may be weak in the time-varying case. This is not surprising, given that relative to a constant parameter setup, the kernel estimator 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, $\tilde{\lambda}_{h,1,t}$, the DM provides

considerable over-coverage throughout horizons, while the AR confidence sets are close to the nominal level. On the other hand, for $\tilde{\lambda}_{h,2,t}$ and $\tilde{\lambda}_{h,3,t}$ the DM provides coverage that is generally too low, whereas coverage by the weak-IV robust confidence sets performs better and remains at or above 90%.

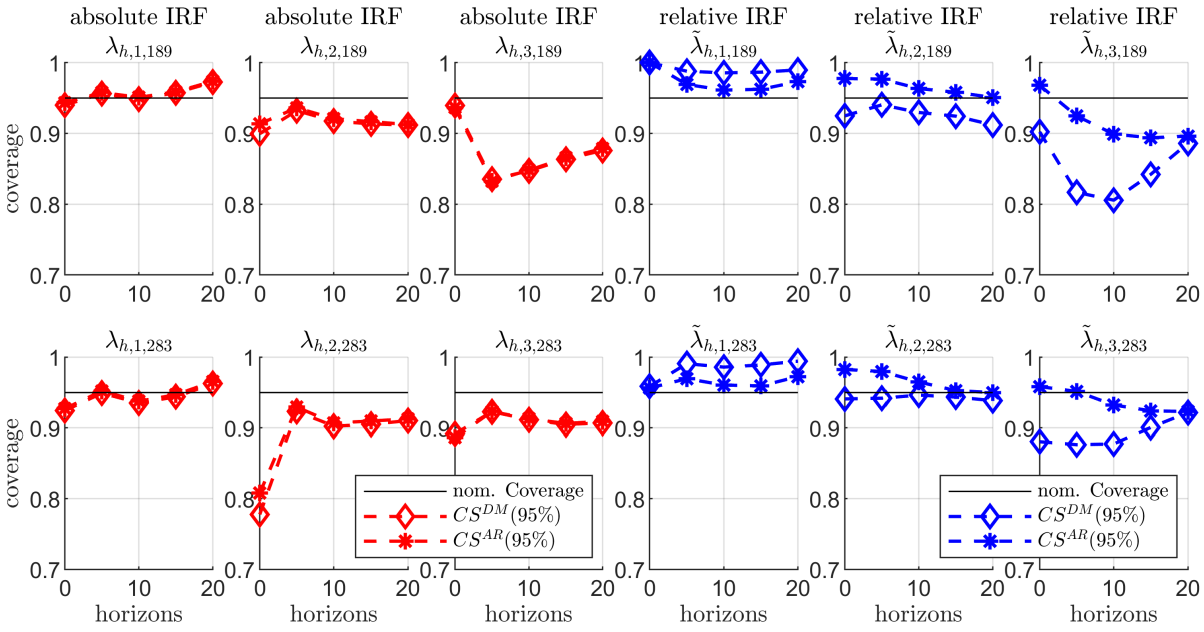

Figure 3: Estimated empirical coverage at 95% confidence level obtained for $\lambda_{h,i,t}$ (red) and $\tilde{\lambda}_{h,i,t}$ (blue) at $t = 1/2T = 189$ (first row) and $t = 3/4T = 283$ (second row). $\theta^{strong} = \{\phi_z = 0.86, \sigma_z = 0.06\}$ and H is estimated. Confidence Sets (CS) based on the Delta Method (DM) are highlighted by diamonds, while Anderson Rubin confidence sets (AR) by stars.

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 $\lambda_{h,2,289}$ where we see some deterioration to levels of about 80%.

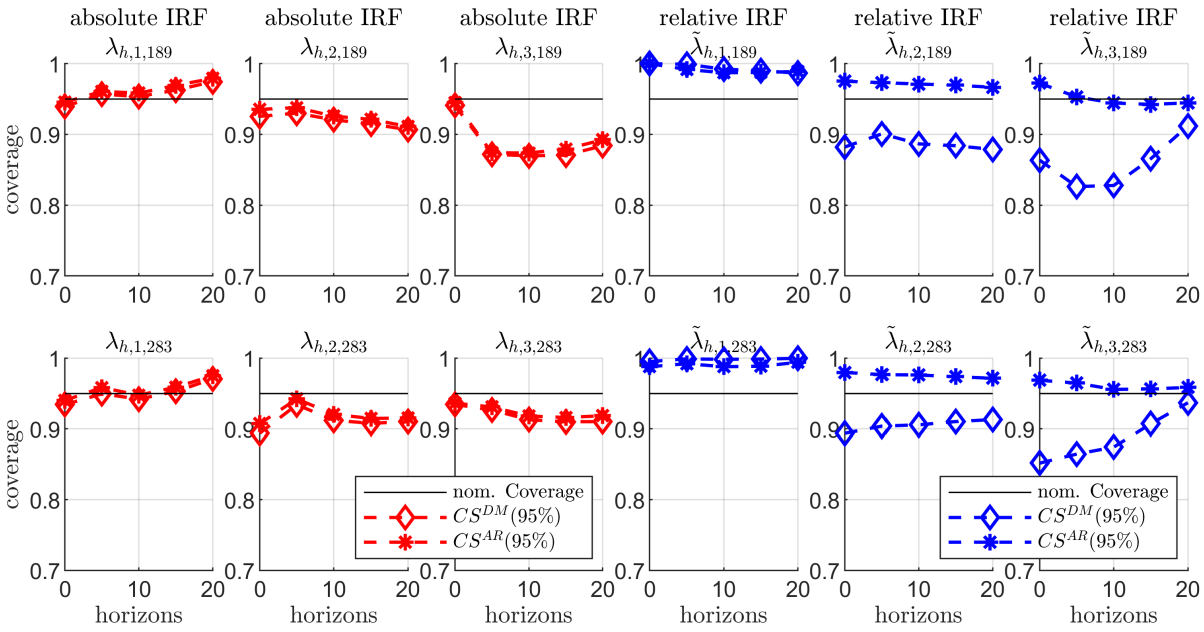

Figure 4: Estimated empirical coverage at 95% confidence level obtained for $\lambda_{h,i,t}$ (red) and $\tilde{\lambda}_{h,i,t}$ (blue) at $t = 1/2T = 189$ (first row) and $t = 3/4T = 283$ (second row). $\theta^{weak} = \{\phi_z = 0.48, \sigma_z = 0.71\}$. Confidence Sets (CS) based on the Delta Method (DM) are highlighted by diamonds, while Anderson Rubin confidence sets (AR) by stars.

For the parameter constellation θ^{weak}, results obtained under the known bandwidth are reported in Figure 4. Starting with $\lambda_{h,i,t}$ (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, it is still the case 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 in the time-varying case.[5]

With respect to $\tilde{\lambda}_{h,i,t}$ (blue), the performance of both confidence sets deteriorates as one would expect when the instruments becomes 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 $\tilde{\lambda}_{h,1,t}$.

[5]Simulation results available upon request find that for absolute IRFs, a much weaker instrument is needed to note a difference between the AR and DM confidence sets, e.g. $\theta = \{\phi_z = 0.48/4, \sigma_z = 0.71\}$.

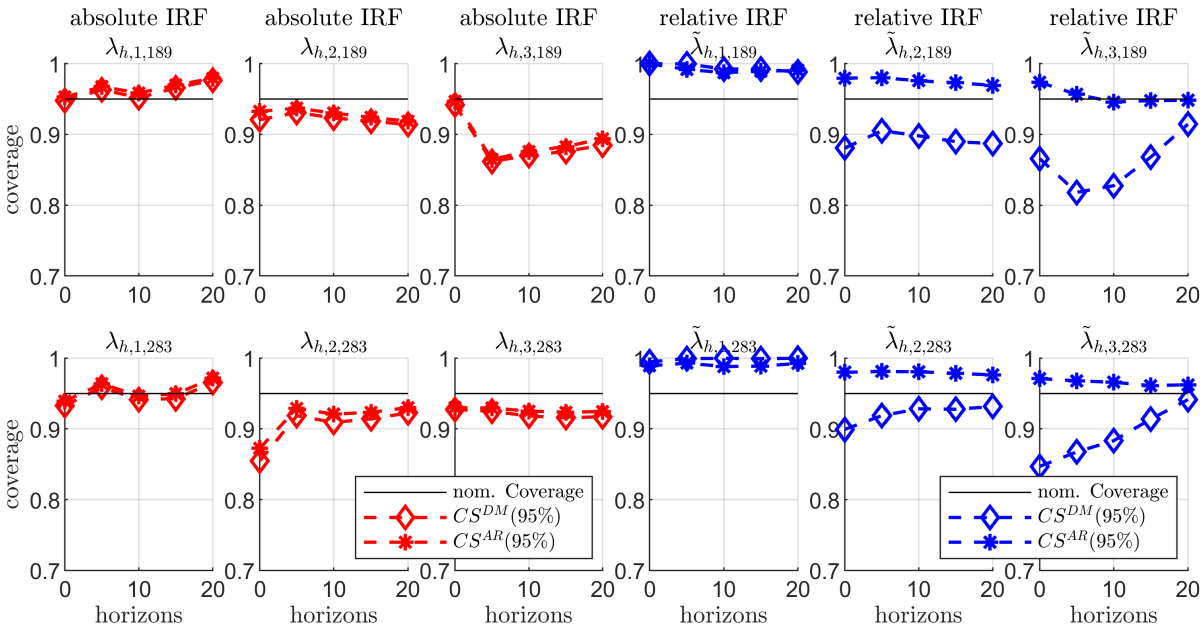

Figure 5: Estimated empirical coverage at 95% confidence level obtained for $\lambda_{h,i,t}$ (red) and $\tilde{\lambda}_{h,i,t}$ (blue) at $t = 1/2T = 189$ (first row) and $t = 3/4T = 283$ (second row). $\theta^{weak} = \{\phi_z = 0.48, \sigma_z = 0.71\}$. and H is estimated. Confidence Sets (CS) based on the Delta Method (DM) are highlighted by diamonds, while Anderson Rubin confidence sets (AR) by stars.

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 of the IRFs (see Figure 5).

In Appendix C, we provide supplementary Monte Carlo results obtained for larger effective sample sizes. Here, we interpolate coefficients linearly to obtain an equivalent shape in the time-varying coefficients, but spread out over a larger sample and hence much smoother. We choose $T = 30 \times 377 = 11310$, and let the kernel bandwidth for estimation increase by $H = \sqrt{30} \times 100 = 547$. Our findings suggest that estimated empirical coverage rates get very close to the nominal size as one would expect in large effective sample sizes.

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

In the following, we illustrate the use of our methodology revisiting the effects of oil supply news shocks 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 oil prices orthogonal to the 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.

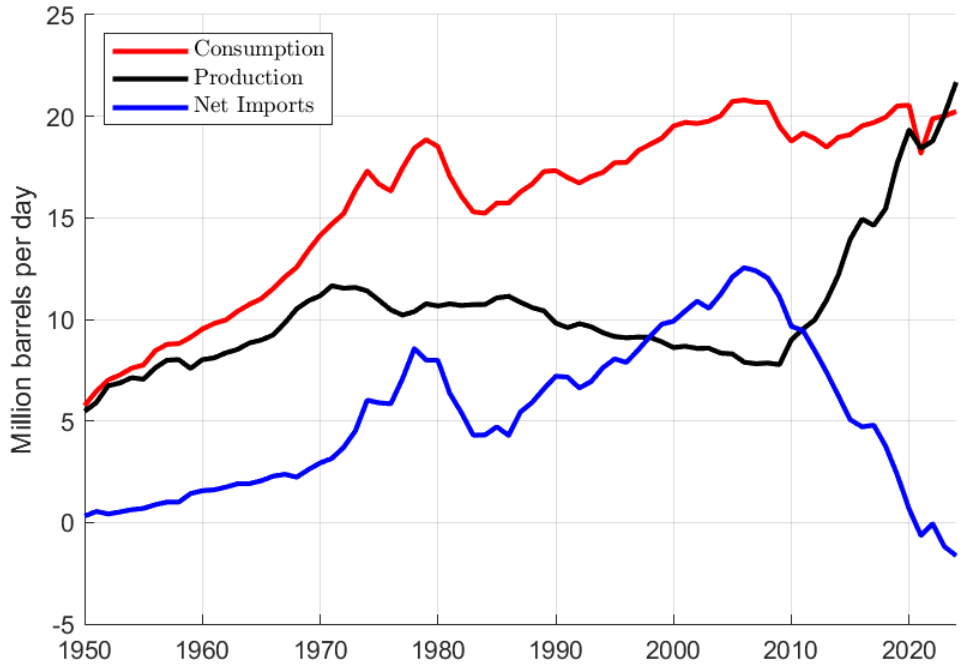

Figure 6: US petroleum consumption, production, and net imports (1950-2023). Source: US Energy Information Administration.

Based on our methodology, we extend the analysis to study instabilities over time. A large body of literature has found that the relationship between oil prices and US macroeco-

nomic conditions has changed, see for example Baumeister and Peersman (2013), Ramey and Vine (2011). Kilian (2009) notes that large part of the instability can be explained by the varying importance of supply- and demand shocks. However, even conditioning on oil-market specific supply shocks a large degree of time-variation remains (Baumeister and Peersman, 2013).

A variety of potential drivers have been put forward to explain the variation, including the time-varying oil intensity of economic activity, improved monetary policy, or changing importance of certain sectors in the economy (Ramey and Vine, 2011). In recent history, a plausible explanation may also be the shale oil revolution. A combination of hydraulic fracturing and horizontal drilling allowed the US to sharply increase its 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 6 (black line).[6] This allowed to United States to transition from a large petroleum net-importer to a petroleum net-exporter in 2020 (blue line, Figure 6).

Indeed, Bjørnland and Skretting (2024) document evidence that the shale oil revolution aligns well with changes in the transmission of oil-market specific shocks to the US economy. Within a Bayesian time-varying factor model, the authors identify an oil-market specific shock by exclusion restrictions, finding that US industrial production and investment reacts more positive to oil price increases since the the shale-oil revolution, boosted by activity in oil-intensive regions and industries. Our empirical findings complement those results, instead relying on kernel estimators and on the identification of oil-supply news shocks by instrumental variables.

[6]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.

4.1 Data and identification strategy

As in Känzig (2021), our VAR model includes a measure of real oil prices, world crude oil production, a proxy of world crude oil stocks, and world industrial production.[7] We augment the model further by two sub-aggregates of the index of US industrial production, that is manufacturing and mining output.[8] 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 mitigate the effect of Covid outliers on our estimates, we use a series of dummies, thereby discounting any signal from the data between February 2020 and June 2022. We confirm that this is equivalent to setting the kernel weighting function to zero over that time period. Following the original paper, we use $p = 13$ lags.

For the hyperparameter H governing overall time-variation, the cross validation procedure discussed in section 2.5 suggests a bandwidth of $H \approx 190$ when applied to the pre-covid period.[9] We note that the objective function is relatively flat between 110 and 250, yielding at most 5% deterioration in the MSE relative to $H = 190$. For this reason, we choose a slightly lower bandwith of $H = 150$ allow for more time-variation in the IRFs, particularly since the shale-oil revolution takes place relatively late in the sample. A sequence of Granger causality tests presented in Table 1 provide no clear evidence at the 5% significance level that the instrument is Granger causing the endogenous variables in the model, with the exception of the very end of the sample. A constant parameter model yields similar conclusions with p-values of 0.8 and 0.85 for the Wald and F-test. Therefore,

[7] The real price of oil is defined as the WTI price deflated by US CPI, and the proxy of world crude oil stocks is included in seasonally adjusted log levels. World industrial production is downloaded from Christiane Baumeisters homepage, see Baumeister and Hamilton (2019).

[8] The US industrial production index measures the combined real output of the manufacturing, mining, and electric and gas utilities industries. The variability in the latter is mostly driven by weather, and hence excluded from our analysis.

[9] Here, the objective function is based on out of sample one-step ahead conditional forecasts computed over the second half of the sample.

we proceed assuming shock invertibility, and study absolute impulse response functions to a shock of unit standard deviation throughout time ($\lambda_{k,i,t}$). To assess the sensitivity of the results to our model choices, Appendix D displays estimates using different bandwidths, and relative IRFs estimated relying on the internal instrument VAR.[10]

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	July 77	May 86	Feb 95	Dec 03	Sep 12	Jun 21
Wald Statistic	96.32	78.49	74.70	81.32	100.89	137.11
p-value	0.08	0.46	0.59	0.38	0.04	0
F Statistic	1.23	1.01	0.96	1.04	1.29	1.76
p-value	0.20	0.49	0.57	0.44	0.15	0.01

The F-test is based on $(n-1)p = 78$ nominator degrees of freedom (dof), and $H - np - 1 = 406$ denominator dofs (Lütkepohl, 2005).

[10] We also test locally for residual autocorrelation and find no evidence thereof, see Appendix D.

4.2 Results

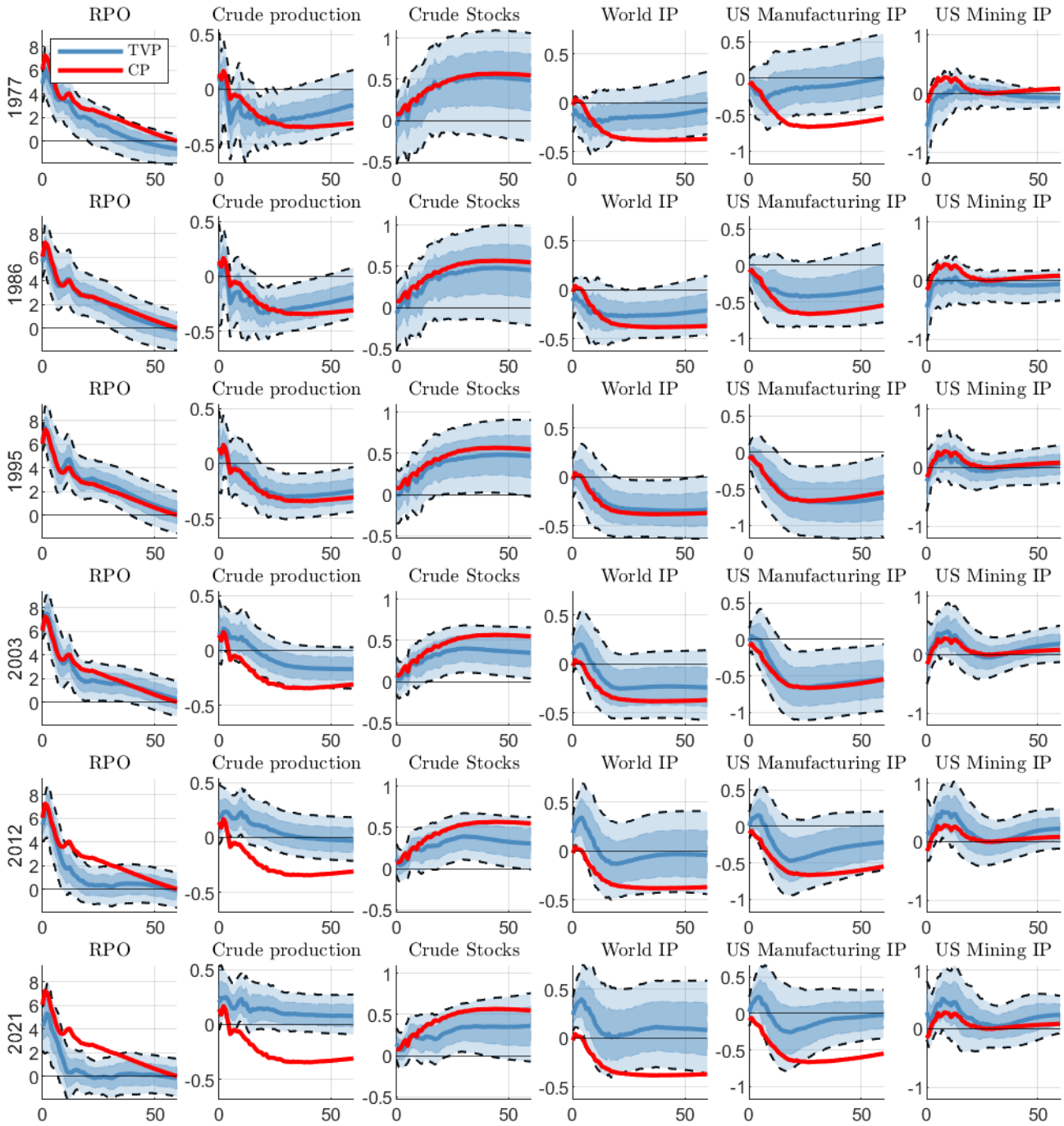

Figure 7: Time-varying impulse response functions to an oil-supply shock of unit variance.

Figure 7 displays estimates of time-varying IRFs 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 replicate those of Känzig (2021). A supply

news shock raises the oil price for up to 5 years. Crude oil production is declining gradually, while crude 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 in large part reflects the extraction of oil and gas, increases slightly but with some lag.[11]

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 2/3 of the price effect towards the end of the sample, compared to estimates from 1977-2003. Furthermore, the price effect is notably less persistent. Despite the overall more muted price signal, 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 micro-evidence on larger price elasticity of shale-oil producers (Aastveit et al., 2022).

Since 2012, world crude oil production is no longer estimated to decline significantly in reaction to an oil-supply news shock, but instead increases somewhat in the short-run. This may reflect, in part, that increasing output by non-OPEC oil producers is able to offset OPEC production declines. The IRF of the world industrial production (IP) index is no longer estimated to decline but instead increases temporarily. However, since the world IP index includes mining output it is difficult to disentangle how much of the response reflects increased oil and gas output by non-OPEC states. Indeed, for US manufacturing output, the time-varying effects are less pronounced. Still, there is striking evidence that the oil price shock no longer triggers a significant decline in US manufacturing output, questioning if it still resembles a cost-push shock for the United States.

To shed more light on the drivers of the aggregate time-variation documented for US

[11]Relative importance weights for the US industrial production index suggest that extraction of oil and natural gas (NAICS 2111) reflects currently about 70% of US mining output.

minining and manufacturing output, we further study responses for three digit (manufacturing) and four digit (mining) industries according to the North American Industry Classification System (NAICS). The collection of responses are obtained by augmenting the model with one industry at the time, while netting out the respective industry variation of the total manufacturing or mining index included in the baseline VAR.

Figure 8 provides an overview of IRFs for selected industries that display large degrees of time-variation between 1995 and 2021. For a complete picture of each industry, we refer to Appendix D.

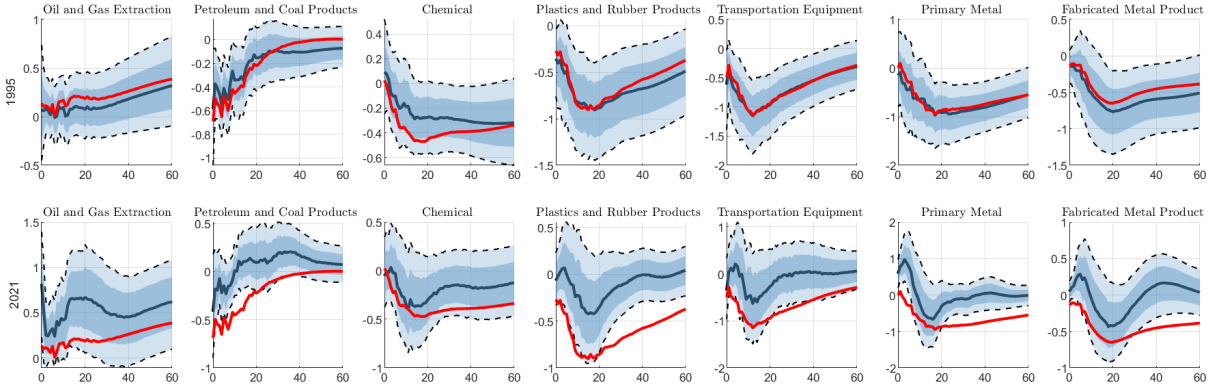

Figure 8: Estimated impulse response functions to an oil-supply shock of unit variance for selected industries at three-(manufacturing) and four- (mining) digit NAICS level. For comparison, the red line shows point estimates from a constant parameter IV-SVAR.

Within US mining output, oil and gas extraction (NAICS 2111) shows the strongest pattern. For the mid 90's, our methodology points to a very slow and imprecisely estimated response. However, in recent times, we find a strong and rapid response which is significant at the 90% confidence interval.

Within manufacturing, the response of Petroleum and Coal Products (NAICS 324) shows a strong pattern of time-variation. Here, the dominant process is petroleum refining which is downstream to Oil and Gas extraction. Our estimates for 1995 show a response close to a constant parameter model, where the industry output declines significantly up to 20 months as the cost of crude increases. Nowadays, our estimates point towards no significant decline, but instead a slight increase in output after 2-3 years.

Estimates for two other downstream industries, that is Chemicals and Plastic and Rubber Products, also have varied significantly, moving from contractionary to insignificant territory. Four other industry responses stand out. For Transportation Equipment (NAICS 336), which reflects mostly the production of motor vehicles and aircrafts, oil supply news shocks tended to be strongly contractionary in line with larger operating costs of the produced goods. There is no statistical evidence this is still the case, as the industry output has become insensitive to the oil price increase. Finally, and somewhat surprisingly, the responses of Primary Metal (NAICS 331) and Fabricated Metal products (NAICS 332) have switched in sign over the short run. As they currently represent a combined 12% of manufacturing output in the US, the time-variation in those sectors is likely contributing strongly to the total manufacturing response.

Summing up, our findings suggest that OPEC oil-supply news shocks transmit differently in recent times than a constant parameter model would suggest. The oil price effects seem to have declined and are less persistent, while there is no evidence that global crude oil production still declines. Within the US, mining output responds more positive and at a faster pace, while there is no evidence that the shock is still contractionary for the US manufacturing sector.

5 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.

We illustrate the methodology revisiting the influential paper by Känzig (2021) on the transmission of oil-supply news shocks. We find strong patterns of time-variation, mostly aligning in time with the shale-oil revolution. While for much of the sample an oil price increase due to oil-supply news resulted in clear headwinds to US manufacturing output, this is no longer the case in more recent history.

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*.

Amir-Ahmadi, P., C. Matthes, and M.-C. Wang (2023): "Understanding Instruments in Macroeconomics-A Study of High-Frequency Identification," Tech. rep., Working paper.

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*.

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.

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," Working Papers 2023-03, Center for Research in Economics and Statistics.

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 Y. Li (2024a): "Regression Modelling under General Heterogeneity," *Preprint, No. 983, Queen Mary University of London.*

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.

Giraitis, L., Y. Li, and P. C. B. Phillips (2024b): "Robust inference on correlation under general heterogeneity," *Journal of Econometrics*, 240, 105–120.

Gonçalves, S. and L. Kilian (2004): "Bootstrapping autoregressions with conditional heteroskedasticity of unknown form," *Journal of Econometrics*, 123, 89–120.

Hipp, R. (2020): "On causal networks of financial firms: Structural identification via non-parametric heteroskedasticity," Tech. rep., Bank of Canada.

Inoue, A., B. Rossi, and Y. Wang (2024a): "Has the Phillips Curve Flattened?" *CEPR Discussion Paper No. 18846*.

——— (2024b): "Local projections in unstable environments," *Journal of Econometrics*, 105726.

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.

Kapetanios, G., M. Marcellino, and F. Venditti (2019): "Large time-varying parameter VARs: A nonparametric approach," *Journal of Applied Econometrics*, 34, 1027–1049.

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.

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.

——— (2005): "New Introduction to Multiple Time Series Analysis," *Springer Books*.

Mertens, K. and J. L. Montiel Olea (2018): "Marginal tax rates and income: New time series evidence," *The Quarterly Journal of Economics*, 133, 1803–1884.

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., S. H. Hoke, K. Bluwstein, et al. (2024): "Patents, News, and Business Cycles," *Staff Working Paper No. 788, Bank of England*.

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.

Müller, U. K. and P.-E. Petalas (2010): "Efficient estimation of the parameter path in unstable time series models," *The Review of Economic Studies*, 77, 1508–1539.

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.

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.

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

Appendix A Proofs

A.1 Proof of Theorem 1

Theorem 1 states that under Assumption 3-5 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_σ 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

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

where

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

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

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

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

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

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 $\frac{1}{\sqrt{H}} \sum_{j=1}^{T} w_{t,j}\xi_j \to \mathcal{N}(0, \Pi_{ww,t})$.

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 3 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 4.

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 ε_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_{ψ,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 ith 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 $\sqrt{H}\left(\hat{\lambda}_{k,t} - \lambda_{k,t}\right) \xrightarrow{d} \mathcal{N}\left(0, \Omega_{k,t}\right)$, where $\Omega_{k,t} = J_k\left(\beta_t, \Gamma_t, \sigma_t\right) V_{\theta_t} J_k\left(\beta_t, \Gamma_t, \sigma_t\right)'$ 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 $n \times (n^2 p + n + n(n+1)/2)$ dimensional gradient. The corresponding derivatives are stated in the following. First, note that $C_k\left(A_t\right) = J_s \mathbf{A}_t^k J_s'$ 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 $k > 1$:

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

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

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

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

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

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

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

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 ijth 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 $\text{CS}^{AR}\{\lambda_{k,i,t} | q(\lambda_{k,i,t}) \leq \chi^2_{1,1-a}\}$. 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:

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

where $\hat{\tilde{\lambda}}_{k,i,t,t_b} = e_{1+i}'\hat{\tilde{\lambda}}_{k,t,t_b}$. Here, $\hat{\tilde{A}}_t$, $\hat{\tilde{P}}_{\bullet 1,t} = e_1'\text{chol}(\hat{\tilde{\Sigma}}_t)$ 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 $\tilde{\Omega}_{k,t,t_b}$ in $\sqrt{H}\left(\hat{\tilde{\lambda}}_{k,t,t_b} - \tilde{\lambda}_{k,t,t_b}\right) \xrightarrow{d} \mathcal{N}\left(0, \tilde{\Omega}_{k,t,t_b}\right)$, an application of the Delta method yields $\Omega_{k,t,t_b} = \tilde{J}_k\left(\tilde{\beta}_t, \tilde{\sigma}_{t,t_b}\right)\text{diag}(V_1, V_2)\tilde{J}_k\left(\tilde{\beta}_t, \tilde{\sigma}_{t,t_b}\right)'$ for

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

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

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

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 $\frac{\partial\tilde{\lambda}_{k,t}}{\partial\tilde{\sigma}_{t,t_b}'}$ 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_{σ_t} be selection matrix such that $\tilde{\sigma}_t = S_{\sigma_t}\tilde{\sigma}_{t,t_b}$ and S_{σ_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:

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

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

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

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

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

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

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

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

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

Finally:

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

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

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

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 30 and $\sqrt{30}$ respectively, interpolating coefficients from the same data-generating process described in section 3. Figure C.9 shows the true underlying impulse response functions, which display the same dynamics just over a larger time frame.

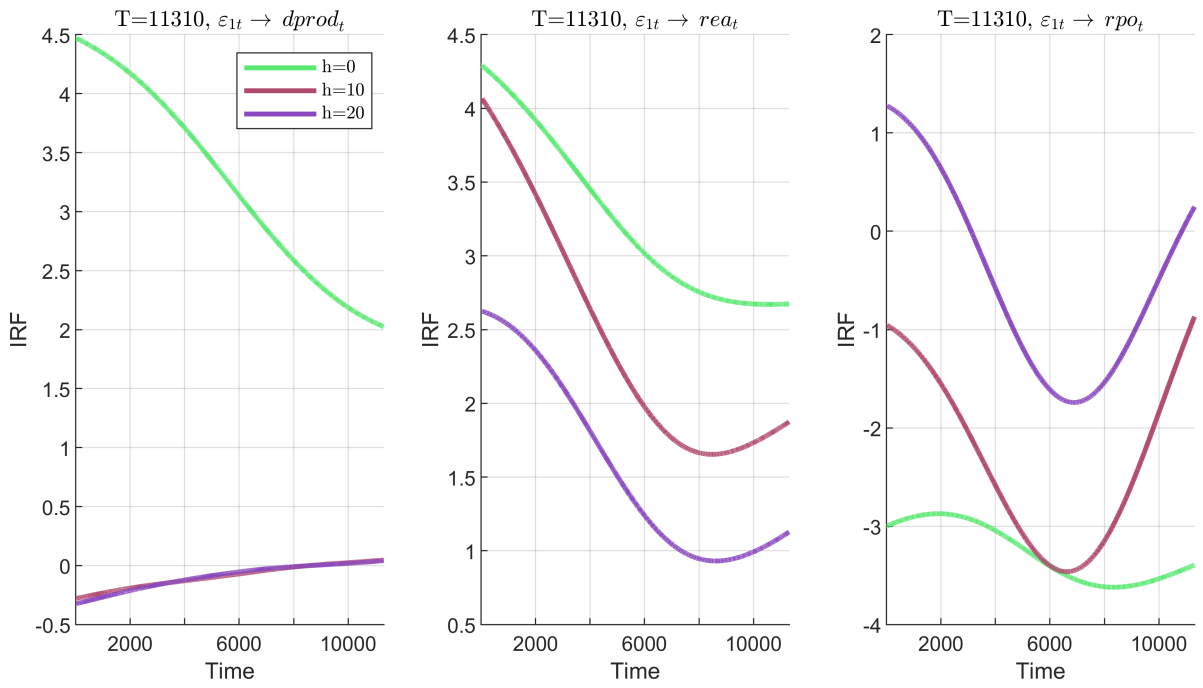

Figure C.9: True impulse response functions $\lambda_{h,i,t}$.

As reported in Figure C.10, estimated empirical coverage gets very close to the nominal 95% confidence level, in the DGP that considers a strong instrument and H to be known.

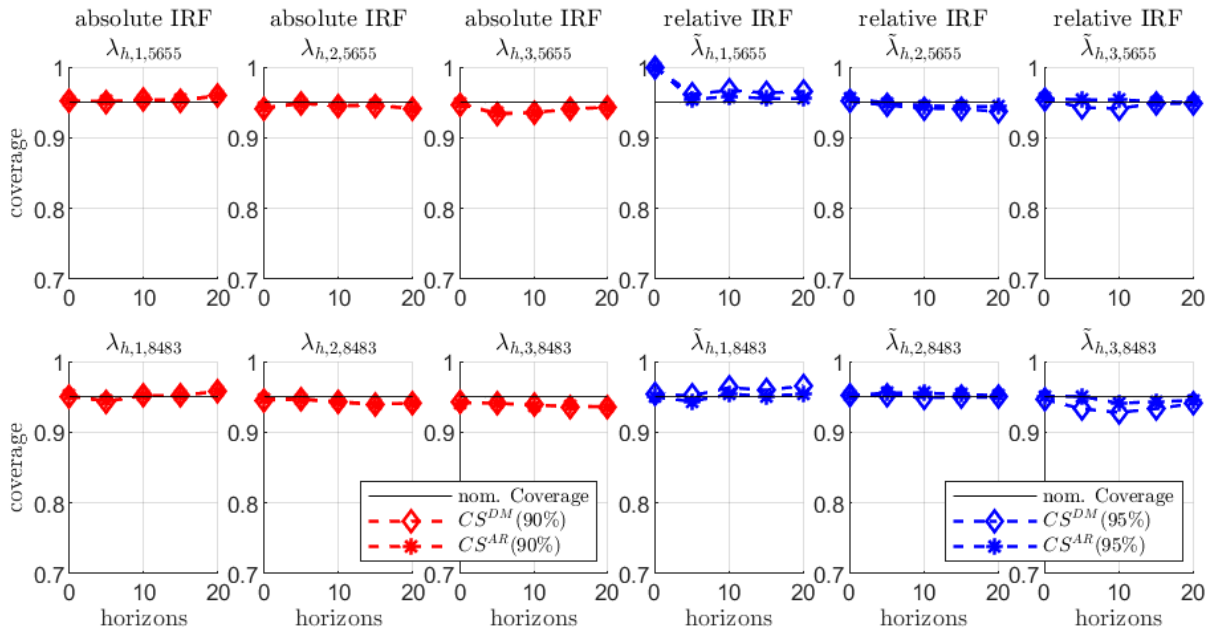

Figure C.10: Estimated empirical coverage at 95% confidence level obtained for $\lambda_{h,i,t}$ (red) and $\tilde{\lambda}_{h,i,t}$ (blue) at $t = 1/2T = 5655$ (first row) and $t = 3/4T = 8483$ (second row). $\theta^{strong} = \{\phi_z = 0.86, \sigma_z = 0.06\}$ and H known. Confidence Sets (CS) based on the Delta Method (DM) are highlighted by diamonds, while Anderson Rubin confidence sets (AR) by stars.

Appendix D Supplementary empirical results

D.1 Test of residual autocorrelation

In this part of the supplementary material, we test if the iid assumption on the residuals may be violated during some periods of the sample. To do so, we follow the textbook treatment in Lütkepohl (2005) leveraging simple multivariate Portmanteau Tests adjusted for the TVP case. At this point, we warn that we have not verified the asymptotic validity of the test in the time-varying case, nor studied its finite sample properties. A thorough analysis thereof is beyond the scope of our paper.

Let the time t sample auto-covariances be $\hat{C}_{it} = \frac{1}{H}\sum_{j=i+1}^{T} w_t(H)\hat{u}_j\hat{u}_{j-i}, i = 1, \ldots, h_q < T$ and the corresponding sample autocorrelation be $\hat{R}_{it} = \hat{D}_t^{-1}\hat{C}_{it}\hat{D}_t^{-1}$ where D_t is a diagonal matrix with elements of $\hat{C}_{0t}$ on the diagonal. Then, we aim to test the null hypothesis that $H_0 : \mathbf{R}_{h_q,t} = (R_{1,t}, \ldots, R_{h_q,t}) = 0$ vs $H_1 : \mathbf{R}_{h_q,t} \neq 0$. The corresponding multivariate Portmanteau test statistic and the approximate distribution is given by

$Q_{h_q,t} = H \sum_{i=1}^{h_q} \text{tr}(\hat{C}_{it}\hat{C}_{0t}^{-1}\hat{C}_{it}\hat{C}_{0t}^{-1}) \approx \chi^2(n^2(h_q - p))$ for large sample sizes H and h_q.

Table 2 summarizes the results from a sequence of Portmanteau Tests for residual autocorrelation up to lag $h_q = 26$ in our empirical application. The test points to no statistical evidence of remaining residual autocorrelation throughout the sample.

Table 2: Portmanteau test results computed at different points of time for the null hypothesis that $\mathbf{R}_{h_q,t} = (R_{1,t}, \ldots, R_{h_q,t}) = 0$.

date t	July 77	May 86	Feb 95	Dec 03	Sep 12	Jun 21
Portmanteau Statistic	332.47	228.76	164.28	141.41	161.57	212.82
p-value	1	1	1	1	1	1

The Portmanteau test is based on $n\,(h_q - p) = 468$ degrees of freedom (Lütkepohl, 2005).

D.2 Robustness of the main results

In this part of the supplementary material, we briefly study the robustness of our results with respect to TVP estimates obtained by relative IRFs, and the bandwidth.

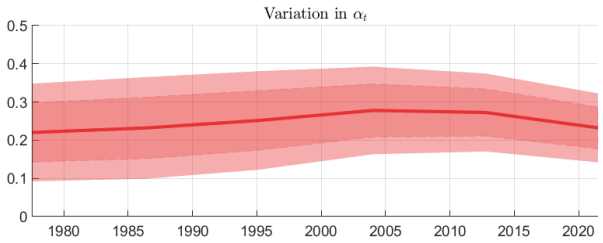

Figure D.11: Estimate of α_t obtained in the IV-SVAR.

First, Figure D.11 displays point estimates alongside 90% confidence intervals of α_t obtained in the IV-SVAR model. Assuming shock invertibility, there is no strong evidence that α_t varies over time. This allows to fix the shock size across time and study the relative IRF $\tilde{\lambda}_{h,i,t,t_b}$. We choose t_b such that we have a reasonable local instrument strength. In our case this is December 2003, which is when the Wald test statistic for the null hypothesis of $e_2'\tilde{P}_{\bullet 1,t} = 0$ is the largest, also guaranteeing finite length of the AR confidence sets.

The top two rows of Figure D.12 show the estimates obtained for relative IRFs $\tilde{\lambda}_{h,i,t,t_b}$ throughout time, standardized to increase the real oil price by 6.2% in December 2003

($H = 150$), hence aligning the shock size with $\lambda_{h,i,t}$ at that month. While uncertainty is fairly large for our estimates of $\tilde{\lambda}_{h,i,t,t_b}$, the baseline point estimates are mostly included in the 90% confidence sets. The results are qualitatively similar, in that oil-supply news shocks are no longer clearly contractionary for the US manufacturing sector, and more recently lead to a large boost in the US mining output.

The bottom two rows of Figure D.12 show estimates obtained for relative IRFs $\lambda_{h,i,t}$ obtained under different bandwidths. As expected, for a larger bandwidth ($H = 190$) the response of manufacturing IP is less time-varying and remains negative towards the end of the sample. Estimates obtained with a lower bandwidth $H = 110$, instead, are more erratic and turn positive towards the beginning and end of the sample. In any case, the alternative point estimates remain within the 68% confidence intervals obtained in our baseline model. Interestingly, IRF estimates of US mining output are less sensitive to the choice of the bandwidth.

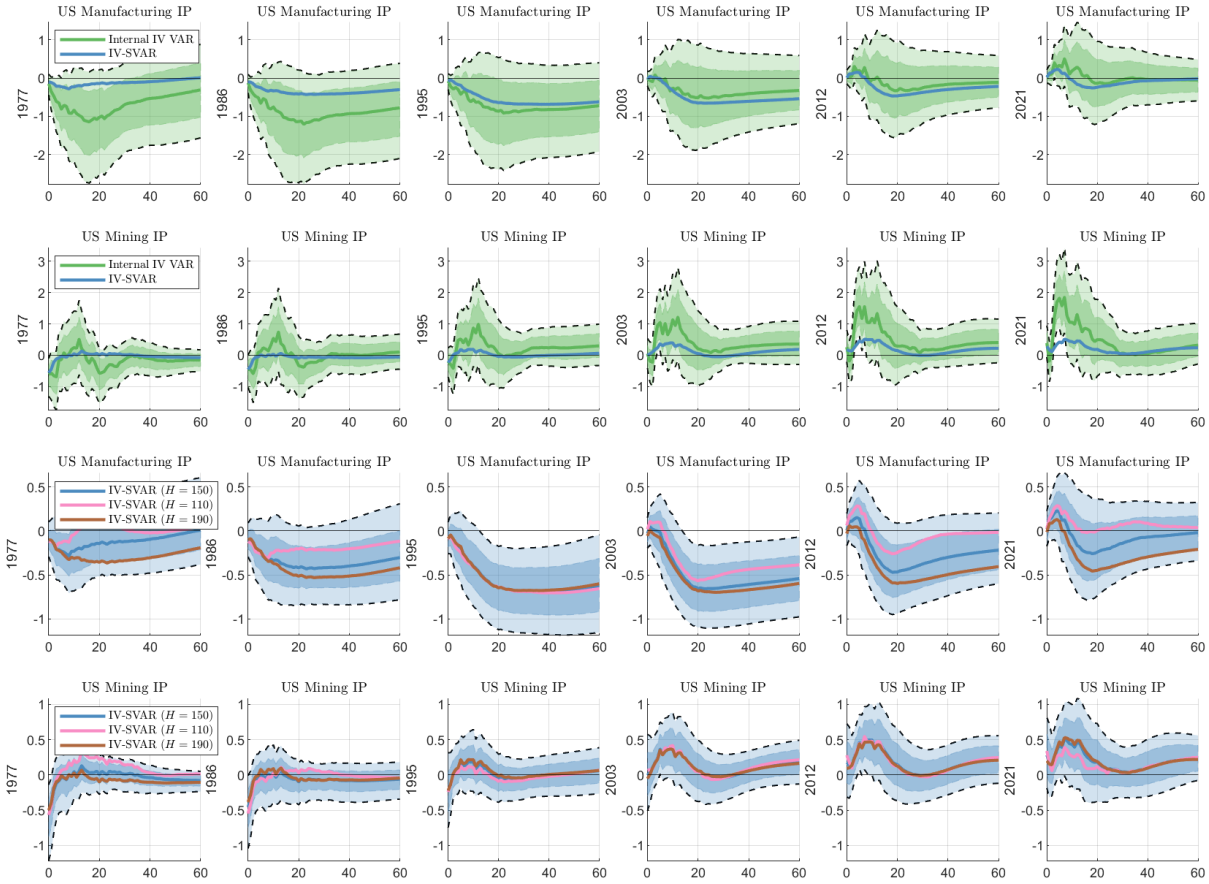

Figure D.12: Time-varying impulse response functions of US manufacturing- and mining output to an oil-supply shock. The top two rows contrast baseline estimates obtained by the IV-SVAR (blue) to estimates of relative IRFs $\tilde{\lambda}_{h,i,t,t_b}$ obtained by the internal instrument VAR (green), standardized to increase the real oil price by 6.2% in December 2003 ($H = 150$). The bottom two rows contrast the baseline results (blue) obtained under $H = 150$ to estimates under different bandwidths ($H = 110$ and $H = 190$). Shaded areas indicate 68% and 90% confidence intervals.

Finally, Figure D.13 shows the entire set of estimates of $\tilde{\lambda}_{h,i,t,t_b}$ for all six variables in the model, and contrasts it to results obtained under the constant parameter case (red). Two findings are worth highlighting. First, the time-varying IRF estimates of the oil-market variables are qualitatively similar to those obtained using an IV-SVAR model. Second, the confidence sets for $\tilde{\lambda}_{h,i,t,t_b}$ are broadly comparable to those obtained in the fix parameter model, and at certain times even narrower. Given the smaller effective sample size for the TVP estimator, this means that estimates of the asymptotic covariance are substantially lower.

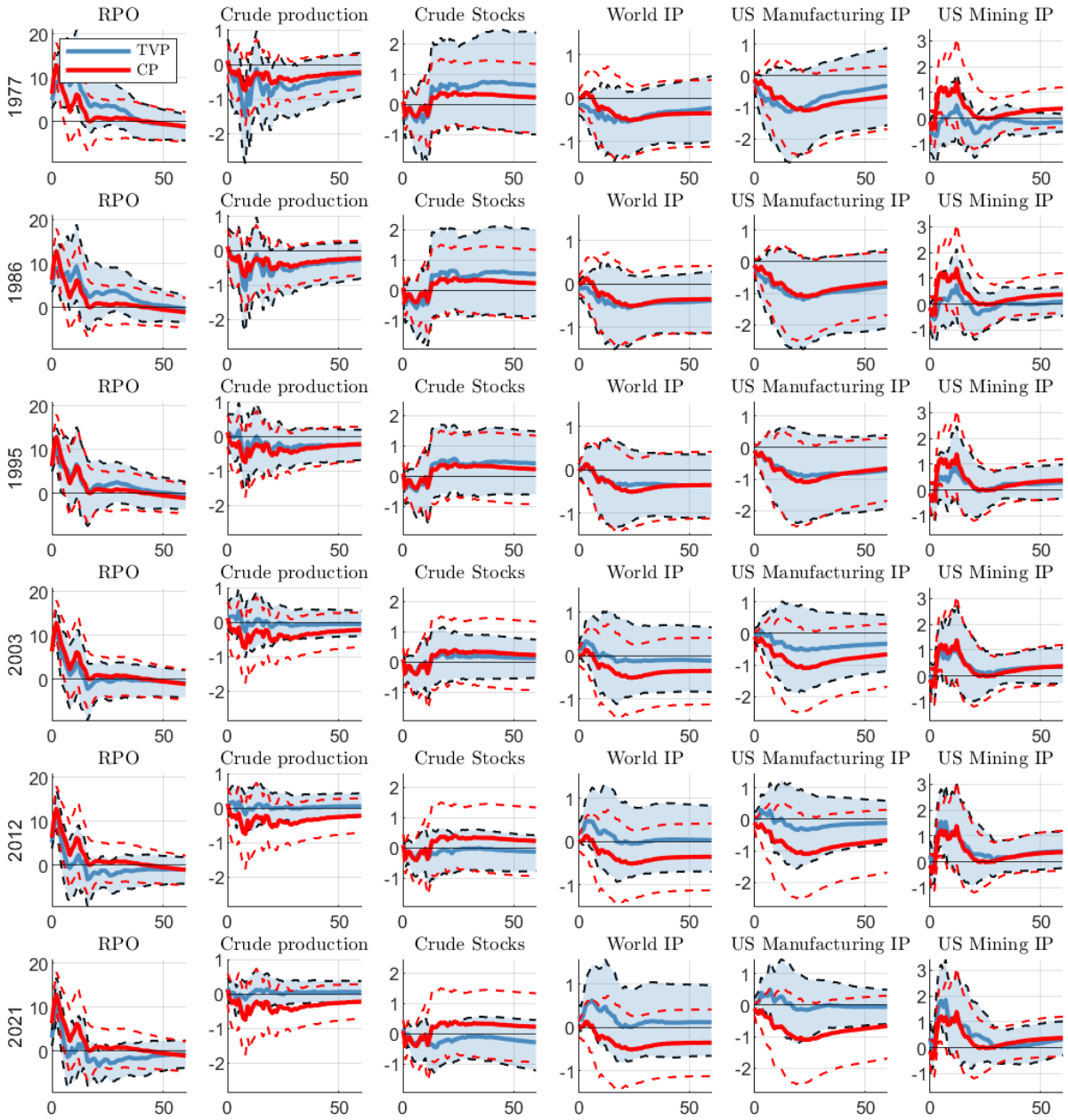

Figure D.13: Time-varying relative impulse response functions $\tilde{\lambda}_{h,i,t,t_b}$ to an oil-supply shock. Estimates are obtained using the internal instrument VAR, standardized to increase the real oil price by 6.2% in December 2003 ($H = 150$). For comparison, the red thick line indicates relative IRF estimates while dashed lines and the shaded area show corresponding 90% confidence sets.

D.3 Complementary results: TVP IRF estimates by industry

This part of the appendix includes complementary estimates of time-varying IRFs for various industries in the manufacturing and mining sector, sorted by durable goods producing industries (Figures D.14 and D.15), nondurable goods producing industries (Figures D.16

and D.17) and mining (Figure D.18)

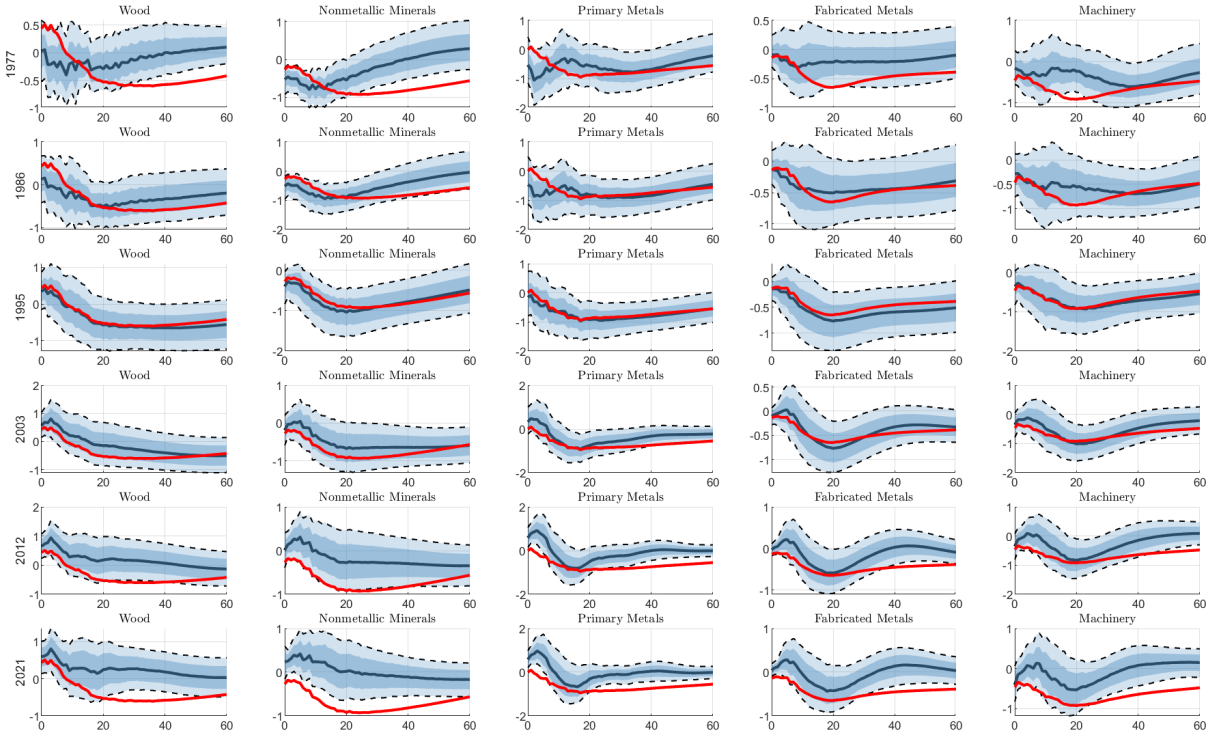

Figure D.14: Time-varying impulse response functions to an oil-supply shock of unit variance. Shaded areas indicate 65% and 90% confidence sets, while the red line indicates point estimates obtained under fix parameters.

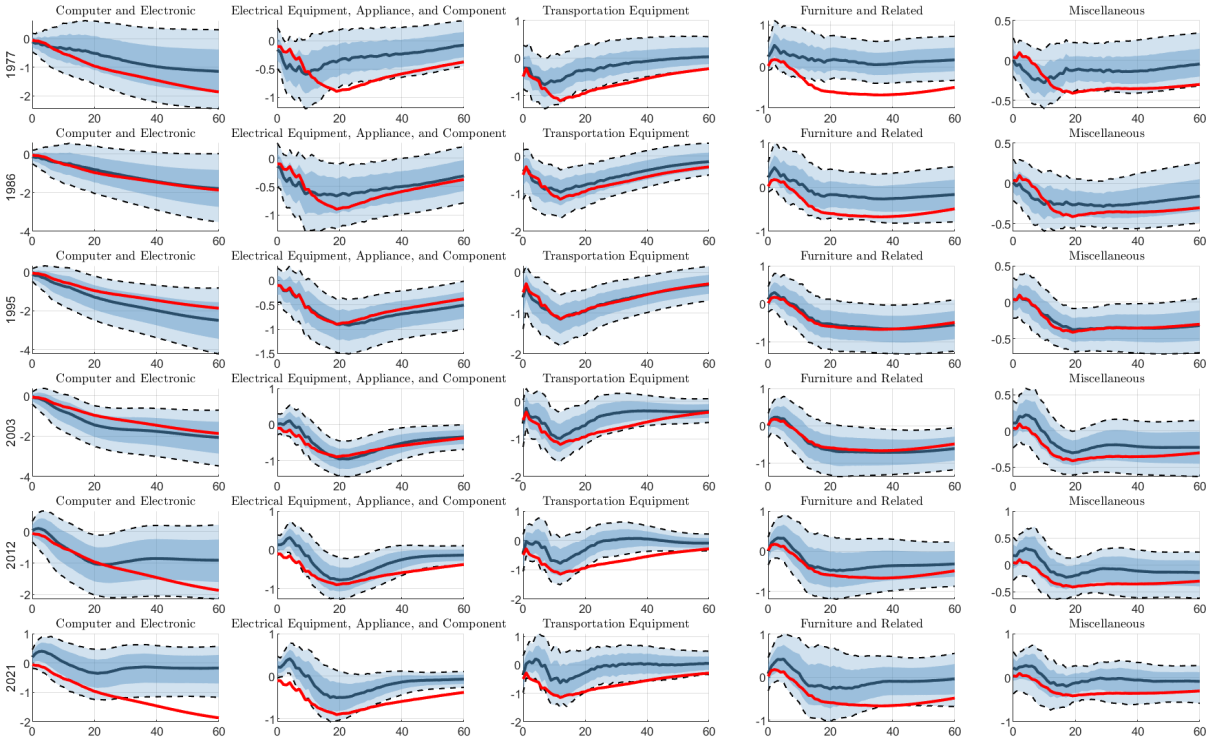

Figure D.15: Time-varying impulse response functions to an oil-supply shock of unit variance. Shaded areas indicate 65% and 90% confidence sets, while the red line indicates point estimates obtained under fix parameters.

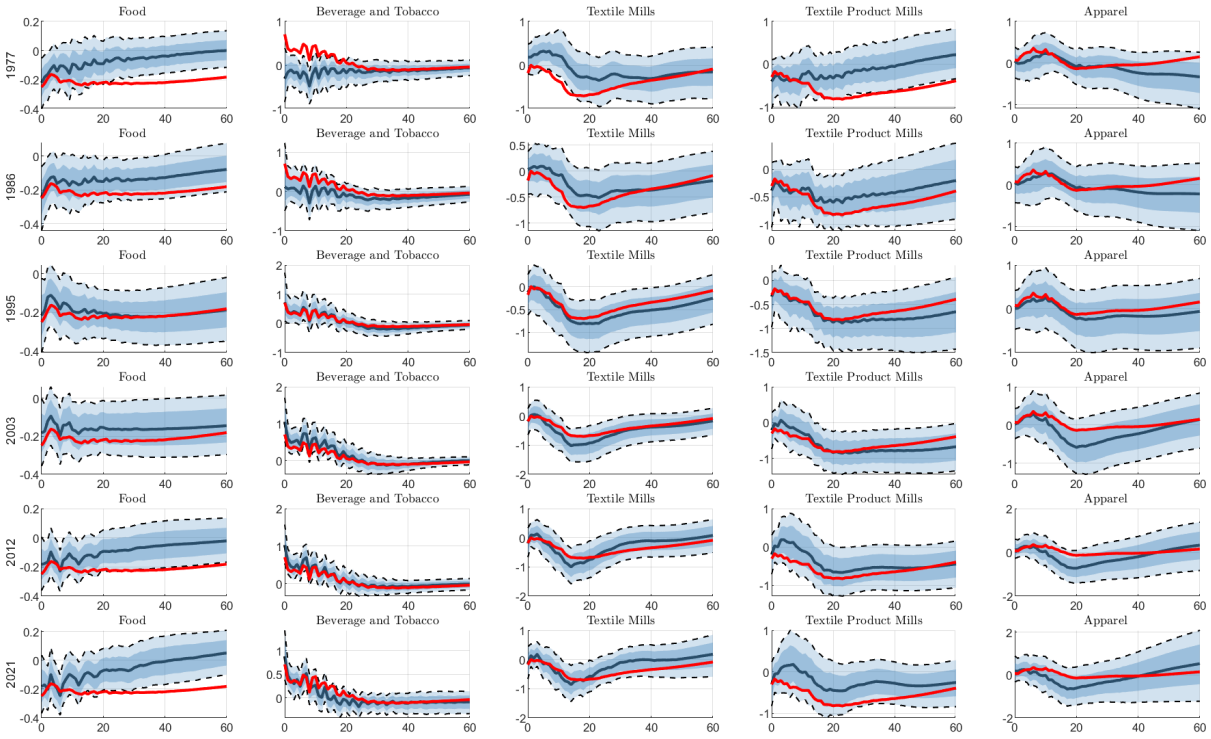

Figure D.16: Time-varying impulse response functions to an oil-supply shock of unit variance. Shaded areas indicate 65% and 90% confidence sets, while the red line indicates point estimates obtained under fix parameters.

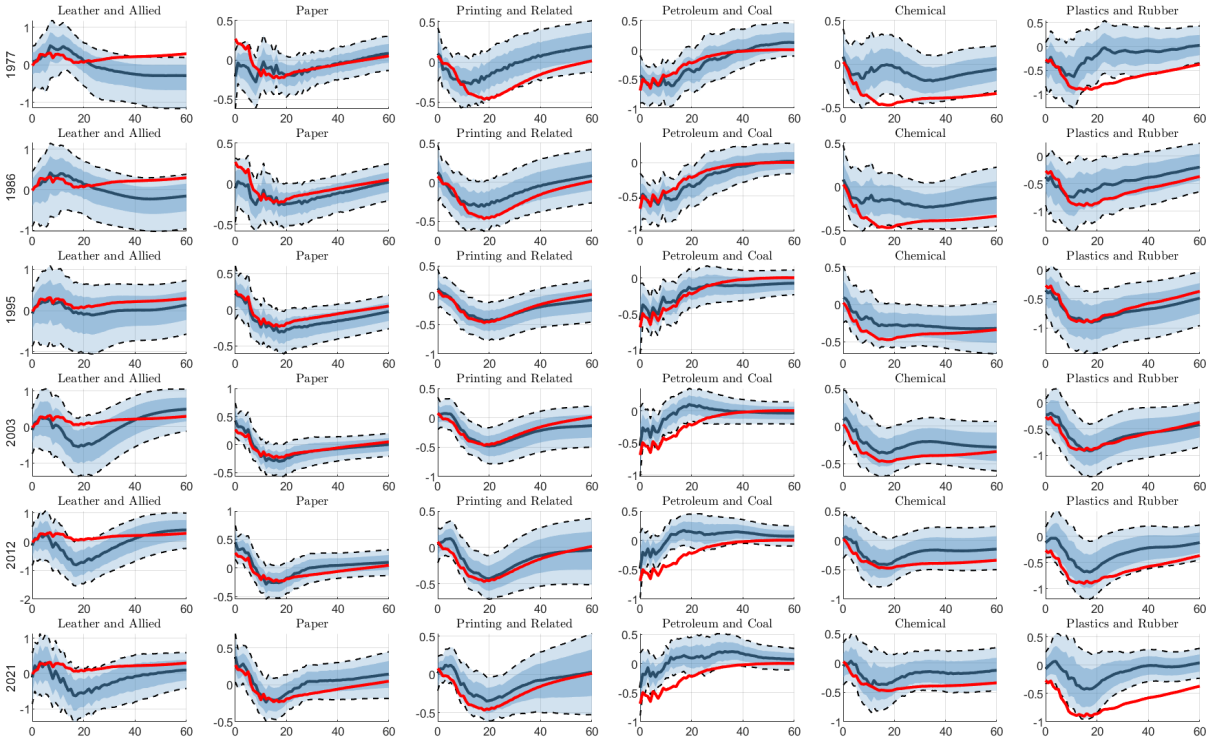

Figure D.17: Time-varying impulse response functions to an oil-supply shock of unit variance. Shaded areas indicate 65% and 90% confidence sets, while the red line indicates point estimates obtained under fix parameters.

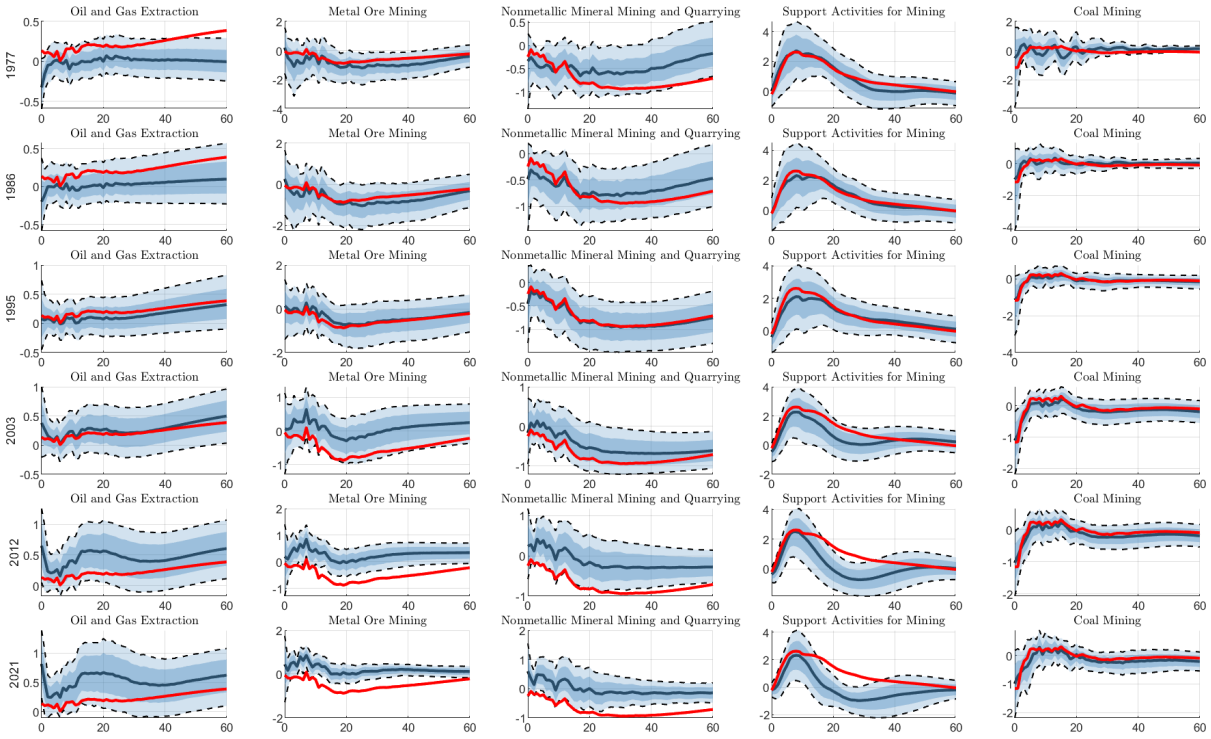

Figure D.18: Time-varying impulse response functions to an oil-supply shock of unit variance. Shaded areas indicate 65% and 90% confidence sets, while the red line indicates point estimates obtained under fix parameters.

Appendix E A comparison to alternative time-varying IRF estimators

A major benefit of the kernel estimator is its simplicity, computational efficiency and the ability to easily include very persistent time-series into the VAR model. Existing methods, including the Bayesian VAR-X model of Paul (2020) and path estimators based on the framework in Müller and Petalas (2010) (see e.g. Inoue et al. (2024b,a)) require to loop through each time t estimate, and hence are computationally more demanding. Furthermore, the Bayesian estimator of IRFs may struggle with explosive posterior draws when persistent time-series are included in the VAR without further transformation. Finally, the methodology in Müller and Petalas (2010) does not offer a joint distribution of parameters at time t and t_b, hence making it difficult to compute IRFs of the same shock size based on an internal instrument VAR. We also encountered that estimates based on the methodology of Müller and Petalas (2010) tend to favor very unstable parameters if we used the proposed default equal-probability mixture for the random walk weighting functions.

Since our empirical applications include long, very persistent time-series we weren't able to compare our method to alternatives in the empirical application.[12] However, in the following, we offer a simple comparison of IRF estimates in a controlled environment, generating the time-series and instrumental variable from a bivariate VAR. We find that all estimator yield fairly similar results.

Our experiment is based on time-series of size $T = 500$, simulated from a Model where $y_t = A_1 y_{t-1} + B_t \varepsilon_t$ where $\varepsilon_t \sim \mathcal{N}(0, I_2)$, $B_t = B_1$ for $t = 1, \dots, \frac{T}{2}$ and $B_t = B_2$ for $t = \frac{T}{2} + 1, \dots, T$. We set $A_1 = \begin{bmatrix} 0.8 & -0.05 \\ 0.2 & 0.7 \end{bmatrix}$, $B_1 = \begin{bmatrix} 0.1 & 0.2 \\ 1 & 1 \end{bmatrix}$, $B_2 = \begin{bmatrix} 0.5 & 0.4 \\ -1 & 1 \end{bmatrix}$. The instrument is simulated from $z_t = 0.2\varepsilon_{1t} + 0.1\eta_t$ where $\eta_t = \mathcal{N}(0, 1)$.

[12] Another challenge we face are the large outliers based on the pandemic, which we simply dummy out in the kernel-based methods. It's less clear how to best treat those observations in the alternative methods.

We compute estimates of impulse responses at $t = [1/8, 2/8, 3/8, 4/8, 5/8, 6/8, 7/8]T$, comparing the kernel based estimator of an IV-SVAR to two alternative methods: the path estimator of an IV-SVAR as proposed in Müller and Petalas (2010), and the VAR-X estimator of Paul (2020). The amount of time-variation is obtained as follows. Our kernel-based estimator relies on a bandwidth of $H = T^{0.5}$, while the path estimator uses the default equal probability mixture proposed in Müller and Petalas (2010). The Bayesian VAR-X is based on a series of independent random walks for the intercepts, the regression coefficients of the instrument, and each of the autoregressive coefficients. Its variances are treated as random and given a conjugate inverse gamma prior with a mean of 0.01^2 and five degrees of freedom. Note that unlike the IV-SVAR estimators, estimates based on the VAR-X are unable to identify the shock variance. For that reason, we standardize the IRFs of the VAR-X estimator to increase the first variable by one at $t = 1/4T$, hence matching the true effects of a unit variance shock.

Figures E.19, E.20 and E.21 show the results of the exercise for the kernel, path and Bayesian estimator respectively. All estimators yield fairly similar point estimates and are able to correctly detect the break at the middle of the sample. Naturally, since all methods assume that parameters are smooth, they fail around the break-point at $T = 1/2T$. However, as they move away from the break estimates get fairly accurate.

The width of the confidence intervals seem close for the kernel and VAR-X estimator. Considerable wider intervals are obtained for the path estimator, which puts a lot of weight on very unstable parameters. However, setting up a path estimator with a tighter specification that favors more stable parameters yields a width of the confidence interval that is comparable to the kernel and VAR-X estimator.[13]

[13]The path estimator suggested in Müller and Petalas (2010) is based on a multivariate Gaussian random walk with a variance proportional to the inverse Hessian of the likelihood. Their default method choice is based on minimizing weighted average risk (WAR) relative to an equal-probability mixture of 11 values for the constant of proportionality c^2/T^2, with $c \in \{0, 5, \ldots, 50\}$. When we set $c \in \{0, 5/10, \ldots, 50/10\}$, we obtain confidence widths that are similar to that of the other two estimators.

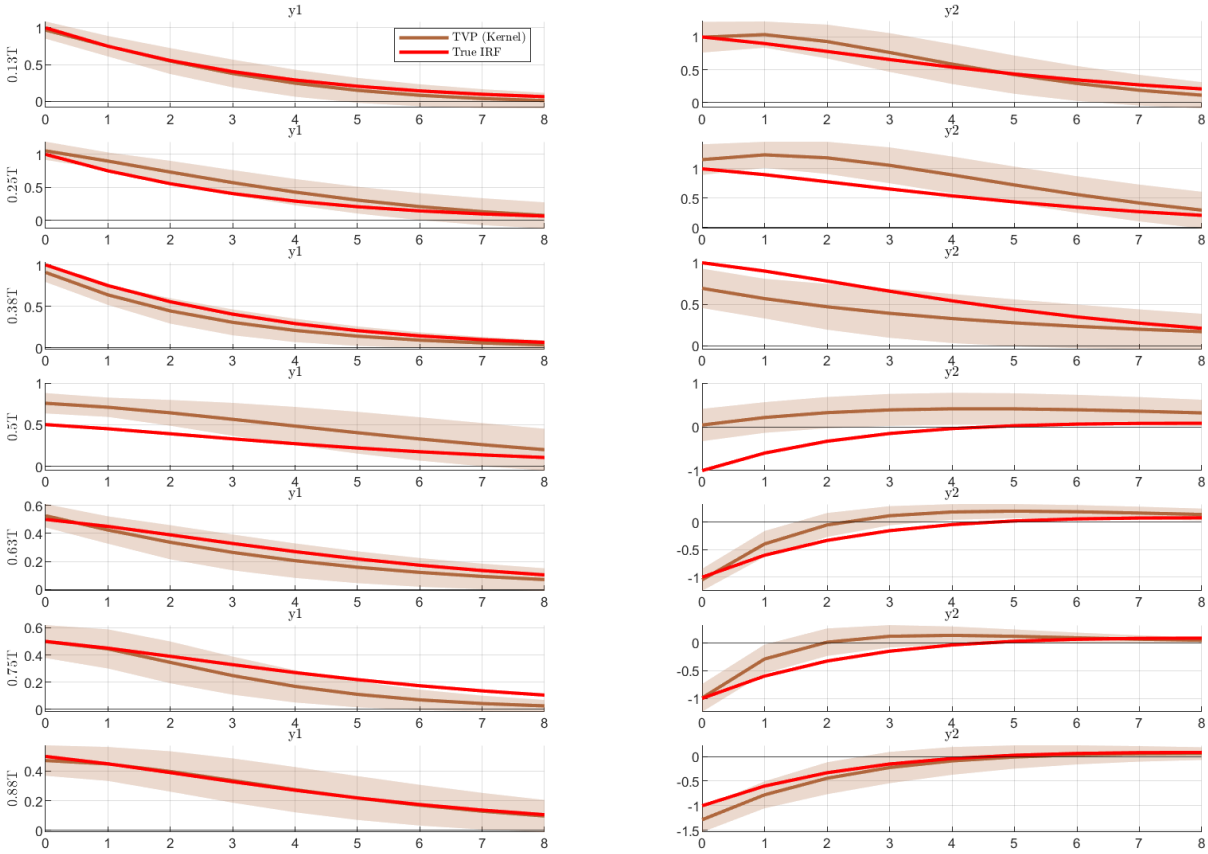

Figure E.19: Kernel based estimator: Point estimates of IRFs alongside 90% confidence intervals at $t = [1/8, 2/8, 3/8, 4/8, 5/8, 6/8, 7/8]T$ for y_1 (first column) and y_2 (second column). The true IRFs are plotted as red lines.

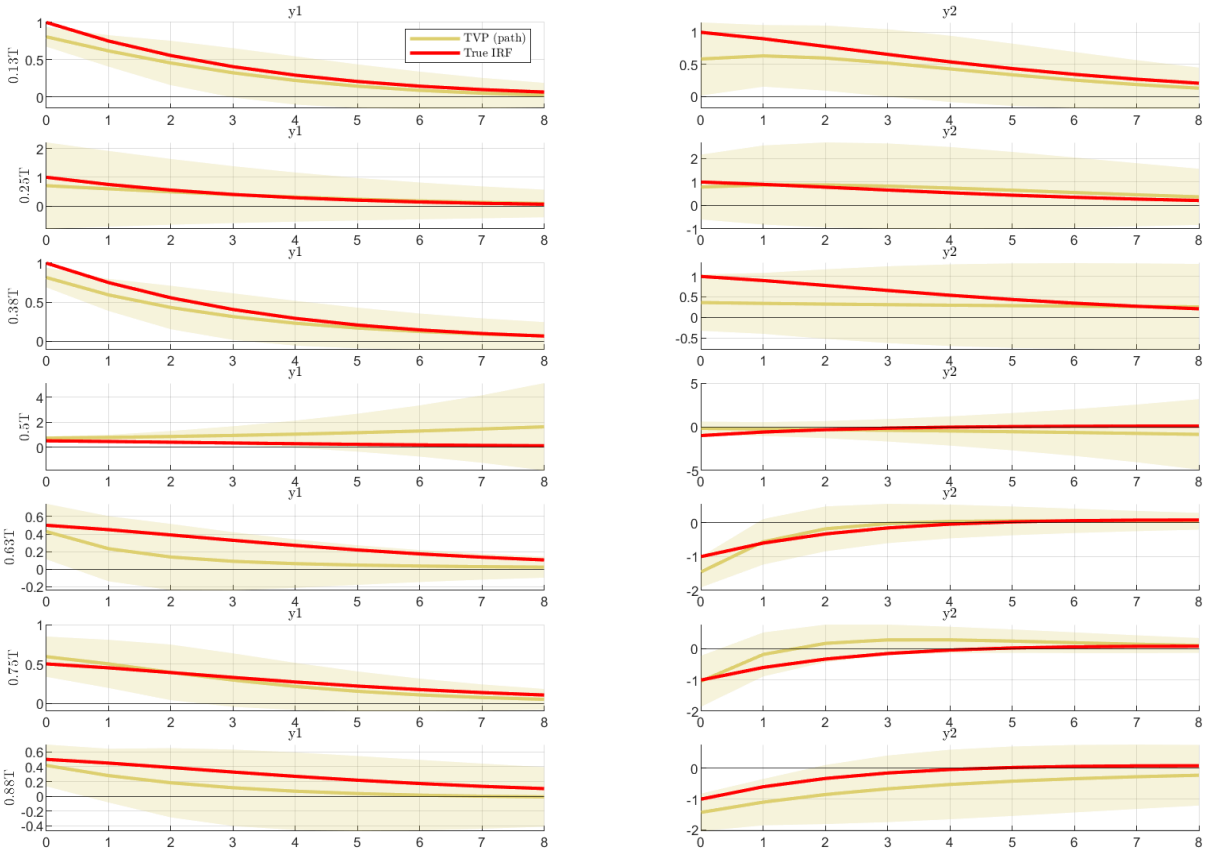

Figure E.20: Path estimator: Point estimates of IRFs alongside 90% confidence intervals at $t = [1/8, 2/8, 3/8, 4/8, 5/8, 6/8, 7/8]T$ for y_1 (first column) and y_2 (second column). The true IRFs are plotted as red lines.

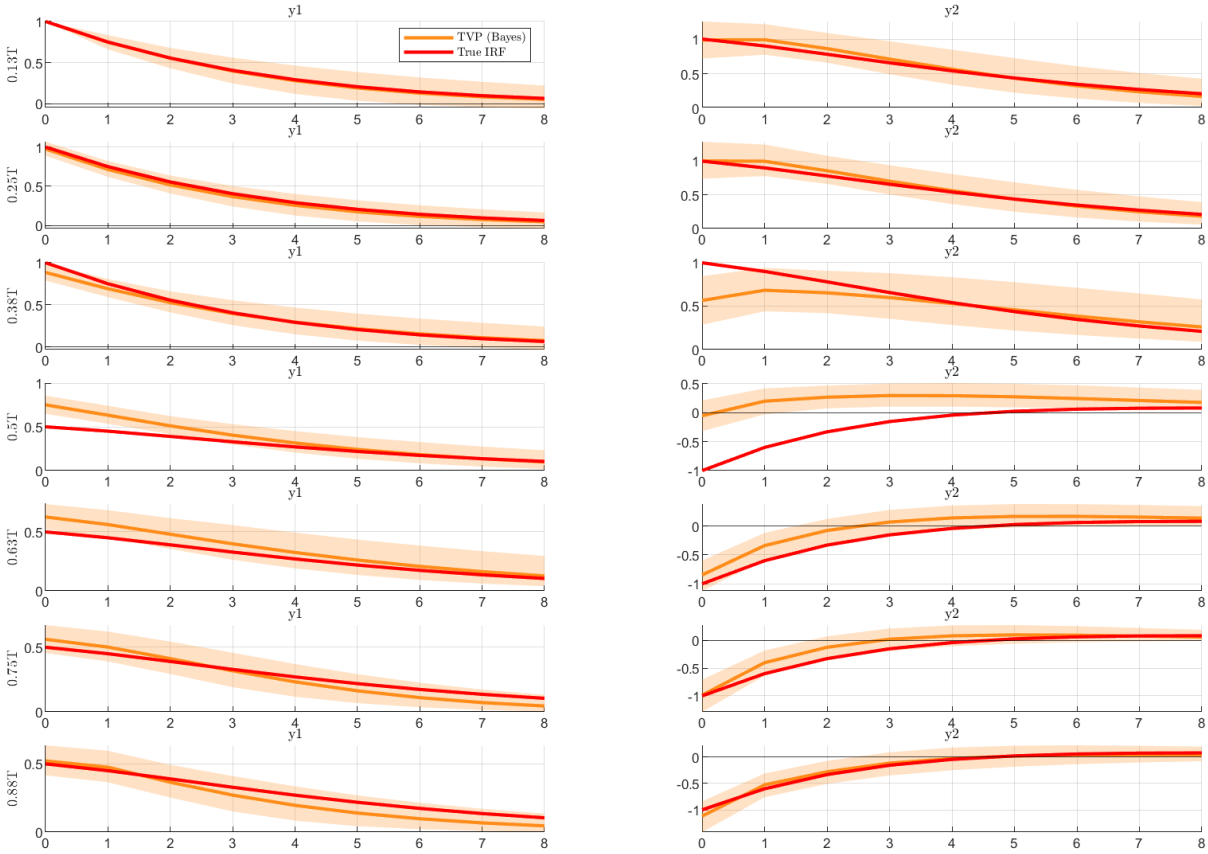

Figure E.21: Bayesian VAR-X: Posterior median estimates of IRFs alongside 90% posterior credible intervals at $t = [1/8, 2/8, 3/8, 4/8, 5/8, 6/8, 7/8]T$ for y_1 (first column) and y_2 (second column). The true IRFs are plotted as red lines.