More on VARs and Local Projections Equivalence: Unit Roots and Multiple Instruments*

Francesco Fusari [†], Joe Marlow[‡§] and Alessio Volpicella[¶]

This draft: February 2026

Abstract

We show that the equivalence in population between impulse responses in Vector Autoregressions (VARs) and Local Projections (LPs) can be extended to (possibly cointegrated) unit roots with unrestricted lag structure. We also prove that structural estimation with multiple instruments for multiple endogenous regressors (LP-IV) is equivalent to a recursively block-identified Structural VAR, where the block of instruments is ordered first. Simulations and two applications illustrate our results.

Keywords: cointegration, impulse response function, local projection, multiple instruments, vector autoregression, unit roots

JEL: C32, C36

*This paper previously circulated with the title "VARs and Local Projections Equivalence for Impulse Responses: Unit Roots and Multiple Instruments". We benefited from comments of seminar participants at Catholic University of the Sacred Heart (Milan) and Newcastle University Business School. This paper reflects the views of the authors and does not necessarily reflect the views of the Central Bank of Ireland.

[†]Newcastle University Business School, 5 Barrack Road, Newcastle upon Tyne NE1 4SE, UK. Email: francesco.fusari@newcastle.ac.uk

[‡]University of Surrey, Elizabeth Fry Building (AD), Guildford GU2 7XH, UK. Email: j.marlow@surrey.ac.uk

[§]Central Bank of Ireland, New Wapping Street, North Wall Quay, Dublin 1, D01 F7X3.

[¶]Corresponding Author, University of Pavia, Via S. Felice Al Monastero, 5, 27100 Pavia, Italy. Email: alessio.volpicella@unipv.it

1 Introduction and Related Literature

The dynamic effects of structural shocks are of paramount importance for macroeconomics, e.g. Frisch (1933) and Ramey (2016). Specifically, the pivotal object for this analysis are Impulse Response Functions (IRFs), i.e. the dynamic response of a macroeconomic variable to a structural shock. Structural Vector Autoregressions (SVARs), popularized by Sims (1980), are the dominant approach for estimation of IRFs. Local Projections (LPs) have been increasingly common as an alternative since Jordà (2005). In *finite samples*, the choice between LP and VAR implies the classical bias-variance trade-off between direct and iterative methods, respectively. The latter are more efficient, but more subject to bias under model misspecification, e.g. heteroskedasticity, non-linearities, omitted covariates, lag order etc. On the other hand, direct methods are more robust to misspecification, but are characterized by higher estimation uncertainty due to serial correlation for residuals and over-parametrization, especially at long horizons. In common macroeconomic applications, time series can be short and serially correlated, implying that the advantages of LPs can be easily exceeded by high sample uncertainty for policy analysis (Li et al., 2024; Kilian and Kim, 2011) and in forecasting exercises (Marcellino et al., 2006; Pesaran et al., 2011; Chevillon, 2007). This bias-variance trade-off has been challenged by Ludwig (2024), who showed that the LP impulse response can be written exactly as a function of VAR impulse responses with varying lag lengths regardless the sample size.

In *population*, Plagborg-Møller and Wolf (2021) proved that LPs and VARs estimate the same IRFs. This result requires the data to be weakly stationary and the lag structure to be unrestricted. The intuition being that weak stationarity allows to apply the Wold decomposition theorem, showing that a VAR with sufficiently large lag length captures all covariance properties of the data.

This paper advances the VAR-LP literature through two contributions. Firstly, under unrestricted lag structures, we establish the in-population equivalence between IRFs from VARs and LPs for (possibly cointegrated) unit roots. For finite lag lengths, approximate equivalence holds in the short-run only. Many economic variables exhibit stochastic trends and long-term co-movements that may be of interest to the researcher, so it is important to investigate the equivalence between VARs and LPs in this setting. We employ the Granger's representation theorem, which allows us to decompose the inverse of a matrix polynomial around the unit roots, to map the permanent component implied by the unit roots onto the space of structural shocks. Since the covariance between the observable of interest and unexpected changes in the regressor being shocked can be interpreted as the estimand in LPs, IRFs coincide in population for VARs and LPs. Appendix A generalizes our result to i) higher order of integration (Stock and Watson, 1993; Johansen, 1995b; McNown and Wallace, 1994) ii) fractional integration (Johansen, 2008; Johansen and Nielsen, 2012)[1] and iii) deterministic trends.

[1]Fractionally integrated VARs allow to relax the assumption of exact $I(1)$ process when long-run identification

Secondly, valid structural estimation with multiple instruments (IVs or proxies) has become increasingly common in the form of augmenting the VAR with proxies ("internal-instrument" VAR), constraining the covariance between external instruments and VAR observables (Proxy SVAR or SVAR-IV) and running the LP with multiple endogenous regressors (LP-IV), e.g. Mertens and Ravn (2013); Lunsford (2015); Mertens and Montiel Olea (2018); Bruns et al. (2025); Huber et al. (2024); Hou (2024); Lakdawala (2019); Fanelli and Marsi (2022); Yang and Zhang (2025). While Plagborg-Møller and Wolf (2021) illustrated the equivalence of IRFs between LP-IV and "internal-instrument" VAR for a *single* instrument, we generalize their finding by showing that IRFs of LP-IV with a *multiplicity* of instruments (for a multiplicity of endogenous regressors) can be obtained by running a recursively block-identified SVAR, where the block of instruments is ordered first. This holds even if the instrumented shocks are non-invertible, unlike the "external-instrument" Proxy SVAR approach.

Simulations from a RBC model, where output and consumption share a common stochastic trend, illustrate the equivalence between LPs and VARs IRFs. For unit roots, an application to oil price shocks shows that IRFs from LP(p) and VAR(p) coincide up to horizon p. A further application based on the effects of monetary policy and central bank information shocks illustrate the equivalence of IRFs between a LP-IV, where two external instruments are used to identify the shocks, and a recursively block-identified SVAR, where the two instruments are ordered first.

The paper is organized as follows. Section 2 presents the results for unit roots; Section 3 investigates the setting with multiple instruments; Section 4 and 5 provide some Monte-Carlo evidence and the empirical applications; Section 6 concludes. The Appendix contains generalizations to i) higher order of (possibly fractional) integration and deterministic trends (Appendix A) and ii) over-identified LP-IV (Appendix B).

2 Integrated Processes

This section provides the in-population equivalence for $I(1)$ processes (Section 2.1) and extends the results to non-recursive VAR identification (Section 2.2) and finite lag lengths (Section 2.3). For our results we employ orthogonal complement of matrices. Given two scalars $\bar{\alpha}$ and $\bar{\beta}$, an orthogonal complement of the $\bar{\alpha} \times \bar{\beta}$ matrix $\boldsymbol{M}$ with $rank(\boldsymbol{M}) = \bar{\beta}$ is denoted by $\boldsymbol{M}_{\perp}$. Put it another way, $\boldsymbol{M}_{\perp}$ is any $\bar{\alpha} \times (\bar{\alpha} - \bar{\beta})$ matrix with $rank(\boldsymbol{M}_{\perp}) = \bar{\alpha} - \bar{\beta}$ and $\boldsymbol{M}'\boldsymbol{M}_{\perp} = \boldsymbol{0}$. The matrix $[\boldsymbol{M}, \boldsymbol{M}_{\perp}]$ is non-singular. Also, if $\boldsymbol{M}$ is a non-singular square matrix ($\bar{\alpha} = \bar{\beta}$), then $\boldsymbol{M}_{\perp} = \boldsymbol{0}$, and, if $\boldsymbol{M} = \boldsymbol{0}$, then $\boldsymbol{M}_{\perp} = \boldsymbol{I}_{\bar{\alpha}}$.

is being used and are relatively popular in that literature, e.g. Chapter 10.5.1 in Kilian and Lütkepohl (2017) and Tschernig et al. (2013); Bollerslev et al. (2013); Carlini and Santucci de Magistris (2019).

2.1 Unit Roots

Assume that the researcher observes $\boldsymbol{z}_t \equiv (\boldsymbol{r}_t', x_t, y_t, \boldsymbol{q}_t')'$, where x_t and y_t are scalar time series; $\boldsymbol{r}_t$ and $\boldsymbol{q}_t$ consist of n_r and n_q controls, respectively. $n \equiv n_r + n_q + 2$ is the number of observables. We are interested in the response of y_t to a shock to x_t. We make the following assumptions.

Assumption A1 *(Unit Roots) Individual time series in $\boldsymbol{z}_t$ are $I(0)$ or $I(1)$ with c_r linearly independent cointegrating relationships, where $0 \leq c_r < n$.*

Assumption A2 *(Gaussianity) $\boldsymbol{z}_t$ is jointly Gaussian.*

The results in this paper do not rely on Gaussianity; however, Assumption A2 is instrumental to employ conditional expectations.

LPs are represented by

$$y_{t+h} = \beta_h x_t + \boldsymbol{\omega}_h' \boldsymbol{r}_t + \sum_{j=1}^{\infty} \boldsymbol{\gamma}_{hj}' \boldsymbol{z}_{t-j} + \epsilon_{ht} \tag{2.1}$$

for each $h = 0, 1, \ldots$. Without loss of generality, we do not include any intercept. Note that we contemporaneously control for $\boldsymbol{r}_t$, not for $\boldsymbol{q}_t$. While we do not require x_t to be an economically interpretable shock, that is often the case in practice. The above formulation is indeed general enough to cover all the common empirical implementations of LPs: projecting y_{t+h} on an exogenous shock x_t; projecting y_{t+h} on an endogenous regressor x_t while controlling for confounding factors in $\boldsymbol{r}_t$; estimating (reduced-form) IRFs from regression (2.1) and rotating them by using the impact IRF matrix from some SVARs.

Definition 2.1 (IRFs LP) *$\{\beta_h\}_{h\geq 0}$ is the LP impulse response of y_t to x_t.*

In practice, this is equivalent to $\beta_h \equiv \mathbb{E}\left(y_{t+h}|x_t = 1, \boldsymbol{r}_t, \{\boldsymbol{z}_\tau\}_{\tau<t}\right) - \mathbb{E}\left(y_{t+h}|x_t = 0, \boldsymbol{r}_t, \{\boldsymbol{z}_\tau\}_{\tau<t}\right)$.

We consider a VAR(∞):

$$\boldsymbol{z}_t = \sum_{j=1}^{\infty} \boldsymbol{A}_j \boldsymbol{z}_{t-j} + \boldsymbol{u}_t, \tag{2.2}$$

where $\boldsymbol{u}_t$ collects the reduced-form shocks, $\boldsymbol{\Sigma}_u \equiv \mathbb{E}(\boldsymbol{u}_t \boldsymbol{u}_t')$ is the variance-covariance matrix and its Cholesky decomposition is $\boldsymbol{\Sigma}_u \equiv \boldsymbol{B}\boldsymbol{B}'$. The corresponding SVAR(∞) is

$$\boldsymbol{z}_t = \sum_{j=1}^{\infty} \boldsymbol{A}_j \boldsymbol{z}_{t-j} + \boldsymbol{B}\boldsymbol{\eta}_t, \tag{2.3}$$

where structural shocks are denoted by $\boldsymbol{\eta}_t \equiv \boldsymbol{B}^{-1}\boldsymbol{u}_t$ and partitioned the same as $\boldsymbol{z}_t$: $\boldsymbol{\eta}_t \equiv (\boldsymbol{\eta}_{rt}', \eta_{xt}, \eta_{yt}, \boldsymbol{\eta}_{qt}')'$. Without loss of generality, $\boldsymbol{B}$ is assumed lower triangular, i.e. the SVAR is recursively identified, and $\boldsymbol{r}_t$ ($\boldsymbol{q}_t$) is ordered first (last). Note that Section 2.2 extends our findings to non-recursive identification.

Definition 2.2 (IRFs VAR) *Let $\{\theta_h\}_{h\geq 0}$ denote the SVAR impulse response of y_t to unitary shock to x_t: $\theta_h \equiv \frac{\partial y_{t+h}}{\partial \eta_{xt}}$.*

For $c_r > 0$, subtracting $\boldsymbol{z}_{t-1}$ on both sides of equation (2.3) yields the (Vector) Error-Correction Model (VECM) representation:

$$\Delta \boldsymbol{z}_t = \boldsymbol{\Pi} \boldsymbol{z}_{t-1} + \sum_{j=1}^{\infty} \boldsymbol{\Gamma}_j \Delta \boldsymbol{z}_{t-j} + \boldsymbol{B} \boldsymbol{\eta}_t, \tag{2.4}$$

where $\boldsymbol{\Pi} = -(\boldsymbol{I} - \sum_{j=1}^{\infty} \boldsymbol{A}_j)$ and $\boldsymbol{\Gamma}_j = -\sum_{j+1}^{\infty} \boldsymbol{A}_j$. $\boldsymbol{\Pi}$ has rank c_r, i.e. the number of linearly independent cointegrating relationships, and drives the long-run relationships among variables. $\boldsymbol{\Pi}$ can be decomposed into $\boldsymbol{\Pi} \equiv \boldsymbol{\alpha}\boldsymbol{\beta}'$, where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}'$ are the loading and cointegrating matrix, respectively. Both $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are $n \times c_r$ matrices. The decomposition is not unique, so a convenient normalization is

$$\boldsymbol{\beta} \equiv \begin{bmatrix} \boldsymbol{I}_{c_r} \\ \boldsymbol{\beta}_{(n-c_r)} \end{bmatrix}, \tag{2.5}$$

where we arrange the elements in $\boldsymbol{\beta}$ such that $\boldsymbol{\beta}_{(n-c_r)}$ is $(n - c_r) \times c_r$.

For $c_r = 0$, $\boldsymbol{\Pi} = \boldsymbol{0}$ and the error-correction framework is reduced to a VAR in first differences.

Next proposition illustrates the equivalence between β_h and θ_h up to a scale, where the scaling factor depends on the variance of the implicit LP innovation.

Proposition 2.1 *Suppose that Assumption A1 and A2 hold, and $\{\boldsymbol{A}_j\}_j$, $\{\boldsymbol{\Gamma}_j\}_j$ are absolutely summable. Set $\tilde{x}_t \equiv x_t - \mathbb{E}(x_t | \boldsymbol{r}_t, \{\boldsymbol{z}_\tau\}_{-\infty<\tau<t})$.*

(i) $c_r > 0$: if $\boldsymbol{\alpha}_\perp' \left(\boldsymbol{I} - \sum_{j=1}^{\infty} \boldsymbol{\Gamma}_j\right) \boldsymbol{\beta}_\perp$ is of full rank, i.e. $n - c_r$, then $\theta_h = \sqrt{\mathbb{E}(\tilde{x}_t^2)}\beta_h$ for $h = 1, 2, \ldots$.

(ii) $c_r = 0$: if $\left(\boldsymbol{I} - \sum_{j=1}^{\infty} \boldsymbol{\Gamma}_j\right)$ is of full rank, i.e. n, then $\theta_h = \sqrt{\mathbb{E}(\tilde{x}_t^2)}\beta_h$ for $h = 1, 2, \ldots$.

Weak stationarity is therefore not required for the equivalence result. On the one hand, under the rank conditions, the Granger's representation theorem maps the permanent component onto the space of structural shocks for VAR. Put it differently, the Granger's representation plays the role of the Wold decomposition in a non-stationary framework. On the other hand, the presence of unit roots does not affect the fact that the estimand in LPs is the covariance between the observable of interest and unexpected changes in the regressor. This result, under further regularity conditions, also holds for higher order of (possibly fractional) integration and deterministic trends (Corollary A.1 and A.2 in Appendix A, respectively).

Proof.

Let us start from the LPs. According to the Frisch-Waugh theorem, we obtain

$$\beta_h = \frac{\text{cov}(y_{t+h}, \tilde{x}_t)}{\mathbb{E}(\tilde{x}_t^2)}. \tag{2.6}$$

Moving everything other than $\boldsymbol{B}\boldsymbol{\eta}_t$ on the left-hand side in equation (2.3) yields

$$\begin{aligned} \boldsymbol{z}_t - \sum_{j=1}^{\infty} \boldsymbol{A}_j \boldsymbol{z}_{t-j} &= \boldsymbol{B}\boldsymbol{\eta}_t \\ \boldsymbol{A}(L)\boldsymbol{z}_t &= \boldsymbol{B}\boldsymbol{\eta}_t, \end{aligned} \tag{2.7}$$

where L is the standard lag operator and $\boldsymbol{A}(z)$ is an infinite-order matrix polynomial: $\boldsymbol{A}(z) = \boldsymbol{I} - \boldsymbol{A}_1 z - \boldsymbol{A}_2 z^2 - \cdots = \boldsymbol{I} - \sum_{j=1}^{\infty} \boldsymbol{A}_j z^j$. The corresponding (infinite-order) characteristic polynomial is $|\boldsymbol{A}(z)|$. For $I(1)$ process, $|\boldsymbol{A}(1)| = 0$, where $z = 1$ signifies the unit root or, mathematically speaking, the singularity, i.e. where $|\boldsymbol{A}(z)| = 0$.

The Laurent series expansion theory, which is used to derive $\boldsymbol{A}(z)^{-1}$ around the singularity, allows us to invoke the Granger's representation theorem.[2] That is a multivariate version of the Beveridge-Nelson decomposition of $\boldsymbol{z}_t$ (Beveridge and Nelson, 1981):

$$\boldsymbol{z}_t = \boldsymbol{\Lambda} \sum_{j=1}^{t} \boldsymbol{B}\boldsymbol{\eta}_j + \sum_{j=0}^{\infty} \boldsymbol{\Omega}_j \boldsymbol{B}\boldsymbol{\eta}_{t-j} + \boldsymbol{z}_0^*, \tag{2.8}$$

where $\sum_{j=0}^{\infty} \boldsymbol{\Omega}_j \boldsymbol{B}\boldsymbol{\eta}_{t-j}$ is $I(0)$ and $\boldsymbol{z}_0^*$ depends on deterministic elements only.

For $c_r > 0$, $\boldsymbol{\Lambda} = \boldsymbol{\beta}_\perp \left[\boldsymbol{\alpha}'_\perp \left(\boldsymbol{I} - \sum_{j=1}^{\infty} \boldsymbol{\Gamma}_j\right) \boldsymbol{\beta}_\perp\right]^{-1} \boldsymbol{\alpha}'_\perp$ is a $n \times n$ matrix. $\boldsymbol{\Lambda}$ exists if and only if the $(n - c_r) \times (n - c_r)$ matrix $\boldsymbol{\alpha}'_\perp \left(\boldsymbol{I} - \sum_{j=1}^{\infty} \boldsymbol{\Gamma}_j\right) \boldsymbol{\beta}_\perp$ is of full rank.[3]

For $c_r = 0$, the rank condition reduces to full rank of $\boldsymbol{I} - \sum_{j=1}^{\infty} \boldsymbol{\Gamma}_j$, in which case

$$\boldsymbol{\Lambda} = \left(\boldsymbol{I} - \sum_{j=1}^{\infty} \boldsymbol{\Gamma}_j\right)^{-1}. \tag{2.9}$$

Equation (2.8) decomposes $\boldsymbol{z}_t$ into $I(1)$ and $I(0)$ components. In particular, $\sum_{j=1}^{t} \boldsymbol{\eta}_j$ is a random-walk process. It follows that

$$\theta_h = \boldsymbol{\Lambda}_{n_r+2,\bullet} \boldsymbol{B}_{\bullet,n_r+1} + \boldsymbol{\Omega}_{h,n_r+2,\bullet} \boldsymbol{B}_{\bullet,n_r+1}, \tag{2.10}$$

where $\boldsymbol{\Lambda}_{n_r+2,\bullet}$ is the $n_r + 2$-th row of $\boldsymbol{\Lambda}$, $\boldsymbol{\Omega}_{h,n_r+2,\bullet}$ is the $n_r + 2$-th row of $\boldsymbol{\Omega}_h$ and $\boldsymbol{B}_{\bullet,n_r+1}$ is the $n_r + 1$-th column of $\boldsymbol{B}$.

[2] See Theorem 4.3 in Johansen (1995a) and Theorem 4.3 and Corollary 4.4 in Franchi and Paruolo (2019).

[3] This result from the Laurent series expansion has been formally shown by Howlett (1982) in Theorem 3, Johansen (1991) in Theorem 4.1 and Franchi and Paruolo (2019) in Theorem 3.3.

Consider the following:

$$\begin{aligned}\operatorname{cov}(y_{t+h}, \eta_{xt}) &= \operatorname{cov}(\mathbf{\Lambda}_{n_r+2,\bullet}\boldsymbol{B}_{\bullet,n_r+1}\eta_{xt} + \mathbf{\Omega}_{h,n_r+2,\bullet}\boldsymbol{B}_{\bullet,n_r+1}\eta_{xt} + y_0, \eta_{xt}) \\ &= \mathbf{\Lambda}_{n_r+2,\bullet}\boldsymbol{B}_{\bullet,n_r+1}\operatorname{var}(\eta_{xt}) + \mathbf{\Omega}_{h,n_r+2,\bullet}\boldsymbol{B}_{\bullet,n_r+1}\operatorname{var}(\eta_{xt}) \\ &= \theta_h,\end{aligned} \tag{2.11}$$

where y_0 denotes the terms uncorrelated with η_{xt} and $\operatorname{var}(\eta_{xt}) = 1$. Thus, we obtain

$$\theta_h = \operatorname{cov}(y_{t+h}, \eta_{xt}). \tag{2.12}$$

Recalling $\boldsymbol{\eta}_t \equiv \boldsymbol{B}^{-1}\boldsymbol{u}_t$ and the Cholesky decomposition delivers:

$$\begin{aligned}&\boldsymbol{B}_{n_r+1,1:n_r}\boldsymbol{\eta}_{rt} + \boldsymbol{B}_{n_r+1,n_r+1}\eta_{xt} = u_{xt} \\ &\boldsymbol{B}_{n_r+1,n_r+1}\eta_{xt} = u_{xt} - \mathbb{E}(u_{xt}|\boldsymbol{\eta}_{rt}) = u_{xt} - \mathbb{E}(u_{xt}|\boldsymbol{u}_{rt}) \equiv \tilde{u}_{xt},\end{aligned} \tag{2.13}$$

where we partition $\boldsymbol{u}_t$ as follows: $\boldsymbol{u}_t \equiv \left(\boldsymbol{u}'_{rt}, u_{xt}, u_{yt}, \boldsymbol{u}'_{qt}\right)'$. $\mathbb{E}(\eta^2_{xt}) = 1$ yields

$$\eta_{xt} = \frac{\tilde{u}_{xt}}{\sqrt{\mathbb{E}(\tilde{u}^2_{xt})}}. \tag{2.14}$$

We observe that

$$\begin{aligned}u_{xt} - \tilde{x}_t &= \mathbb{E}(x_t|\boldsymbol{r}_t, \{\boldsymbol{z}_\tau\}_{-\infty<\tau<t}) - \mathbb{E}(x_t|\{\boldsymbol{z}_\tau\}_{-\infty<\tau<t}) = \mathbb{E}(u_{xt}|\boldsymbol{r}_t, \{\boldsymbol{z}_\tau\}_{-\infty<\tau<t}) \\ &= \mathbb{E}(u_{xt}|\boldsymbol{u}_{rt}, \{\boldsymbol{z}_\tau\}_{-\infty<\tau<t}) = \mathbb{E}(u_{xt}|\boldsymbol{u}_{rt}).\end{aligned} \tag{2.15}$$

Thus, we note that $\tilde{u}_{xt} \equiv u_{xt} - \mathbb{E}(u_{xt}|\boldsymbol{u}_{rt})$ and get

$$\tilde{u}_{xt} = \tilde{x}_t. \tag{2.16}$$

Combining the previous equation with (2.10) and (2.14) yields

$$\theta_h = \frac{\operatorname{cov}(y_{t+h}, \tilde{x}_t)}{\sqrt{\mathbb{E}(\tilde{x}^2_t)}}. \tag{2.17}$$

Proposition 2.1 follows. ■

2.2 Non-recursive Identification

Proposition 2.1 is stated for a recursively identified SVAR for expositional convenience. However, as in Plagborg-Møller and Wolf (2021) the equivalence between VAR- and LP-based impulse responses does not rely on recursive identification. The key observation is that any scalar identified shock can be written (up to normalization) as a linear combination of reduced-form innovations. This is not affected by being in a non-stationary setting.

Fix an arbitrary vector $\boldsymbol{b} \in \mathbb{R}^n$ and define the scalar innovation $v_t \equiv \boldsymbol{b}'\boldsymbol{u}_t$. Define the "unexpected" component of v_t after partialling out contemporaneous controls $\boldsymbol{r}_t$ and the infinite lag history:

$$\tilde{v}_t \equiv v_t - \mathbb{E}(v_t \,|\, \boldsymbol{r}_t, \{\boldsymbol{z}_\tau\}_{-\infty<\tau<t}), \qquad \eta_t^{(\boldsymbol{b})} \equiv \frac{\tilde{v}_t}{\sqrt{\mathbb{E}(\tilde{v}_t^2)}}, \tag{2.18}$$

so that $\operatorname{var}(\eta_t^{(\boldsymbol{b})}) = 1$. Consider the LP regression

$$y_{t+h} = \bar{\beta}_h \, v_t \; + \; \bar{\boldsymbol{\omega}}_h' \boldsymbol{r}_t \; + \; \sum_{j=1}^{\infty} \bar{\boldsymbol{\gamma}}_{hj}' \boldsymbol{z}_{t-j} \; + \; \bar{\epsilon}_{ht}, \tag{2.19}$$

and define $\bar{\beta}_h$ as the associated population coefficient. Since $\operatorname{var}(\eta_t^{(\boldsymbol{b})}) = 1$, the same steps as in the proof of Proposition 2.1 deliver $\theta_h^{(\boldsymbol{b})} = \sqrt{\mathbb{E}(\tilde{v}_t^2)}\, \bar{\beta}_h$.

Remark. A generic (not necessarily recursive) SVAR identification can be written as $\boldsymbol{u}_t = \boldsymbol{B}\boldsymbol{\eta}_t$ with $\mathbb{E}(\boldsymbol{\eta}_t\boldsymbol{\eta}_t') \equiv \boldsymbol{I}$ and an invertible $\boldsymbol{B}$ not required to be triangular. Any scalar identified shock $\eta_{1t} \equiv \boldsymbol{e}_1'\boldsymbol{\eta}_t$ is therefore of the form $\eta_{1t} = \boldsymbol{b}'\boldsymbol{u}_t$ with $\boldsymbol{b} = \boldsymbol{B}^{-\prime}\boldsymbol{e}_1$ (up to scale). Hence, the LP coefficient on $\boldsymbol{b}'\boldsymbol{u}_t$ (with the same controls) recovers the corresponding VAR impulse response.

2.3 Finite Lag Length

Finite-p estimands for LP under unit roots. Empirical works employ a finite lag length p. Define the LP(p) population regression

$$y_{t+h} = \beta_{hp} x_t + \boldsymbol{\omega}_{hp}' \boldsymbol{r}_t + \sum_{j=1}^{p} \boldsymbol{\gamma}_{hj,p}' \boldsymbol{z}_{t-j} + \epsilon_{ht}^{(p)}, \tag{2.20}$$

and the corresponding residualized regressor

$$\tilde{x}_t^{(p)} \equiv x_t - \mathbb{E}(x_t \mid \boldsymbol{r}_t, \boldsymbol{z}_{t-1}, \ldots, \boldsymbol{z}_{t-p}). \tag{2.21}$$

By the Frisch-Waugh theorem,

$$\beta_{hp} = \frac{\operatorname{cov}(y_{t+h}, \tilde{x}_t^{(p)})}{\mathbb{E}\left[(\tilde{x}_t^{(p)})^2\right]}. \tag{2.22}$$

The finite-p LP estimand β_{hp} is well-defined in our (possibly) $I(1)$ setting because it depends on the projection residual $\tilde{x}_t^{(p)} \equiv x_t - \mathbb{E}(x_t \mid \boldsymbol{r}_t, \boldsymbol{z}_{t-1}, \ldots, \boldsymbol{z}_{t-p})$, which has finite variance and is $I(0)$, and on y_{t+h}, which admits a linear representation in structural shocks under the Granger's representation theorem.

Finite-p estimands for VAR under unit roots. To connect LP(p) and VAR(p) impulse-response estimands, we adapt the projection argument in Plagborg-Møller and Wolf (2021), Appendix A.1, to unit roots by working with a covariance-stationary state. Under Assumptions A1–A2 and the summability/rank conditions in Proposition 2.1, define the stationary state

$$\boldsymbol{s}_t \equiv \begin{cases} \left(\Delta \boldsymbol{z}_t',\ (\boldsymbol{\beta}'\boldsymbol{z}_{t-1})'\right)' & \text{if } c_r > 0, \\ \Delta \boldsymbol{z}_t & \text{if } c_r = 0. \end{cases} \tag{2.23}$$

This delivers a pseudo-true *state* VAR(p) for $\boldsymbol{s}_t$: $\boldsymbol{s}_t = \sum_{j=1}^{p} \boldsymbol{\Phi}_j^{(p)} \boldsymbol{s}_{t-j} + \boldsymbol{e}_t^{(p)}$. When written back in terms of $\boldsymbol{z}_t$, it corresponds to a VECM($p-1$) representation for $\boldsymbol{z}_t$ (with regressors $(\boldsymbol{z}_{t-1}, \Delta\boldsymbol{z}_{t-1}, \ldots, \Delta\boldsymbol{z}_{t-p+1})$ and, when $c_r > 0$, the error-correction term $\boldsymbol{\beta}'\boldsymbol{z}_{t-1}$). By the standard VECM$\leftrightarrow$VAR mapping, that VECM($p-1$) can in turn be reparameterized as a VAR(p) in levels for $\boldsymbol{z}_t$, i.e. the order-p finite-lag approximation:

$$\boldsymbol{z}_t = \sum_{j=1}^{p} \boldsymbol{A}_{jp} \boldsymbol{z}_{t-j} + \boldsymbol{u}_t^{(p)}, \tag{2.24}$$

with $\boldsymbol{\Sigma}_u^{(p)} \equiv \mathbb{E}(\boldsymbol{u}_t^{(p)} \boldsymbol{u}_t^{(p)\prime})$, a recursive identification (ordering $\boldsymbol{r}_t$ first and $\boldsymbol{q}_t$ last) given by a Cholesky factor $\boldsymbol{\Sigma}_u^{(p)} = \boldsymbol{B}_p \boldsymbol{B}_p'$ and structural shocks $\boldsymbol{\eta}_t^{(p)} \equiv \boldsymbol{B}_p^{-1} \boldsymbol{u}_t^{(p)}$.

In this formulation, the counterfactual covariance operator $\mathrm{cov}_p(\cdot,\cdot)$ used by Plagborg-Møller and Wolf (2021), Appendix A.1, i.e. the covariance computed under the pseudo-true state VAR(p) approximation, is well-defined, and the key step (that the pseudo-true VAR(p) matches autocovariances out to lag p) applies to the state $\boldsymbol{s}_t$. Since $\tilde{x}_t^{(p)}$ is $I(0)$, it is an L^2 linear functional of the stationary shocks (and hence of $\boldsymbol{s}_t$). Therefore, the conclusion of Plagborg-Møller and Wolf (2021), Proposition 2, carries over: the LP(p) and VAR(p) impulse-response estimands coincide for horizons $h \leq p$.

For $h > p$, the link need not hold because VAR(p) constructs h-step responses by iterating the one-step transition implied by the finite-lag companion form; with unit roots, this iteration governs the stochastic-trend component, so the truncation (remainder) accumulates and the resulting cross-horizon restrictions generally differ from the direct h-step projection defining LP(p), unless the DGP is exactly VAR(p).

3 Multiple Instrumental Variables

In order to discuss structural analysis, we assume a general Structural Vector Moving Average (SVMA) setting, which covers all (linearized) DSGE models and SVARs, as Data-Generating Process (DGP). Without loss of generality, we keep considering processes with no intercept. To simplify notation, this section imposes weak stationarity on the covariates; that can be relaxed by invoking Proposition 2.1.

Assumption A3 *(DGP) Data* $\boldsymbol{z}_t$ *are driven by* n_ξ *exogenous structural shocks* $\boldsymbol{\xi}_t \equiv (\xi_{1t}, \ldots, \xi_{n_\xi t})$*:*

$$\boldsymbol{z}_t = \boldsymbol{\Theta}(L)\boldsymbol{\xi}_t, \tag{3.1}$$

$$\boldsymbol{\Theta}(L) \equiv \sum_{l=0}^{\infty} \boldsymbol{\Theta}_l L^l, \tag{3.2}$$

$$\boldsymbol{\xi}_t \sim N(\boldsymbol{0}, \boldsymbol{I}), \tag{3.3}$$

where $\boldsymbol{\Theta}_l \in \mathbb{R}^{n \times n_\xi}$*,* $\{\boldsymbol{\Theta}_l\}_l$ *is absolutely summable and* $\boldsymbol{\Theta}(x)$ *has full row rank.*

Thus, $\boldsymbol{z}_t$ is a non-singular, strictly stationary jointly Gaussian time series. The (i, j)-th element $\boldsymbol{\Theta}_{ijl}$ of $\boldsymbol{\Theta}_l$ is the IRF of variable i to shock j at horizon l. If the researcher is interested in the effect of shock ξ_{1t} on y_{t+h}, consistently with applied literature we consider the relative IRF $\tilde{\boldsymbol{\Theta}}_{n_r+2\ 1h} \equiv \frac{\boldsymbol{\Theta}_{n_r+2\ 1h}}{\boldsymbol{\Theta}_{n_r+1\ 10}}$: the response of y_{t+h} after a shock on ξ_{1t} that increases x_t by one unit on impact.

Instruments (known as proxy variables in macroeconometric literature) are commonly used in semistructural analysis. This section establishes the in-population equivalence between the LP instrumental variable estimation with more instruments for more endogenous regressors and the IRFs of a recursively block-identified SVAR(∞), with the block of instruments ordered first. Our framework nests the one-instrument-for-one-regressor scenario analyzed by Plagborg-Møller and Wolf (2021).
An instrument (instrumental variable, IV) is an observable w_t that is contemporaneously correlated with the shock being instrumented (relevance condition), contemporaneously uncorrelated with other shocks (exogeneity condition) and uncorrelated with all shocks at all leads and lags (lead-lag exogeneity). For m instruments collected by the vector $\boldsymbol{w}_t$, those conditions, which correspond to the validity requirements in Stock and Watson (2018), can be represented as follows.

Assumption A4 *(Valid Instruments) For simplicity, assume that the* $k = m$ *shocks of interest are ordered last in* $\boldsymbol{\xi}_t$*.*

$$\boldsymbol{w}_t = \sum_{l=1}^{\infty} \left(\boldsymbol{\Upsilon}_l \boldsymbol{w}_{t-l} + \tilde{\boldsymbol{\Omega}}_l \boldsymbol{z}_{t-l} \right) + \tilde{\boldsymbol{\Lambda}} \boldsymbol{\xi}_t + \boldsymbol{v}_t, \tag{3.4}$$

where $\tilde{\boldsymbol{\Lambda}} \equiv \left[\boldsymbol{0}_{m \times (n_\xi - m)}, \boldsymbol{D} \right]$*,* $\boldsymbol{D}$ *is a* $m \times m$ *diagonal matrix with full rank,* $\boldsymbol{v}_t \sim N(\boldsymbol{0}, \boldsymbol{\Sigma}_v)$*,* $\boldsymbol{v}_t$ *is independent of* $\boldsymbol{\xi}_t$ *at all leads and lags and* $\boldsymbol{\Sigma}_v$ *is diagonal.* $\boldsymbol{I} - \sum_{l=1}^{\infty} \boldsymbol{\Upsilon}_l L^l$ *have all the roots outside the unit circle and* $\{\tilde{\boldsymbol{\Omega}}_l\}_l$ *is absolutely summable.*

Consider the following LP-OLS regression with k endogenous covariates:

$$y_{t+h} = \boldsymbol{\beta}_h' \boldsymbol{X}_t + \sum_{j=1}^{\infty} \boldsymbol{\gamma}_{hj}' \boldsymbol{z}_{t-j} + \epsilon_{ht}, \tag{3.5}$$

where $\boldsymbol{X}_t$ is a column-vector containing k endogenous regressors and $\boldsymbol{\beta}_h'$ is the corresponding row-vector of k coefficients.

The LP Instrumental Variable (LP-IV) framework estimates the IRFs $\boldsymbol{\beta}_h$ using a two-stage least squares version of LP. In practice, the regressors $\boldsymbol{X}_t$ are instrumented by $\boldsymbol{w}_t$. Let $\boldsymbol{\beta}_{IVh}$ denote the LP-IV estimand for the k endogenous regressors for two-stage least-squares estimation of regression (3.5), where the first-stage regression is

$$\boldsymbol{X}_t = \boldsymbol{\beta}_{FS}'\boldsymbol{w}_t + \sum_{j=1}^{\infty} \boldsymbol{\gamma}_{FSj}'\boldsymbol{Z}_{t-j} + \boldsymbol{\epsilon}_{FSt}, \tag{3.6}$$

with $\boldsymbol{Z}_t \equiv (\boldsymbol{w}_t', \boldsymbol{r}_t', \boldsymbol{X}_t, y_t, \boldsymbol{q}_t')'$ being the extended dataset.

The following proposition claims that the LP-IV IRFs can be estimated from a recursively block-identified SVAR(∞) that orders the block of instruments first and the block of covariates last, where the order within the two blocks is irrelevant.

Proposition 3.1 *Suppose that $\boldsymbol{Z}_t$ is weakly stationary, $m = k$, Assumption A2 holds and $\boldsymbol{\beta}_{FS}$ is of full rank. Consider a recursively 2-block-identified SVAR(∞), where $\boldsymbol{w}_t$ is ordered first and $(\boldsymbol{r}_t', \boldsymbol{X}_t', y_t, \boldsymbol{q}_t')'$ is ordered last. Let $\tilde{\boldsymbol{\theta}}_{x0}$ denote a $m \times m$ matrix collecting the contemporaneous responses of $\boldsymbol{X}_t$ to the first m shocks; let $\tilde{\boldsymbol{\theta}}_{yh}$ denote a column-vector of m rows collecting the responses at horizon h of y_t to the first m shocks.*

(i) We obtain

$$\boldsymbol{\beta}_{IVh} = \tilde{\boldsymbol{\theta}}_{x0}^{-1}\tilde{\boldsymbol{\theta}}_{yh}; \tag{3.7}$$

(ii) under Assumption A3 and A4, $[\boldsymbol{\beta}_{IVh}]_j = \left[\left(\tilde{\boldsymbol{\theta}}_{x0}\right)^{-1}\tilde{\boldsymbol{\theta}}_{yh}\right]_j = \tilde{\boldsymbol{\Theta}}_{n_r+2\ j^*h}$ *for* $j = 1, \ldots, m$ *and* $j^* = n_\xi - m + j, \ldots,$ *where* $[\bullet]_j$ *indicates the* j*-th element of the vector* $\bullet$.

Appendix B extends the proposition to over-identification, i.e. $m > k$. We do not require any structure, i.e. Assumption A3 and A4, to obtain the in-population equivalence (i). They are only required to interpret the LP-IV and the IRFs from the recursively 2-block-identified SVAR(∞) as (relative) structural responses. The latter are valid even when the m shocks of interest $\boldsymbol{\xi}_{n_\xi-m+1}, \ldots, \boldsymbol{\xi}_{n_\xi}$ are not invertible. The intuition for one instrument case (Plagborg-Møller and Wolf, 2021) can be generalized to $m > 1$ IVs: “internal-instrument” recursive SVAR, where we put the (noisily measured) shocks (IVs) first, can estimate the relative IRF correctly, even if there is no invertible SVAR.

On a related note, our setting nests the result in Plagborg-Møller and Wolf (2021) for $m = k = 1$ as a particular case: $\beta_{IV} = \frac{\tilde{\theta}_{yh}}{\tilde{\theta}_{x0}}$, i.e. the LP-IV is equivalent to a recursively ordered SVAR(∞), with the (single) instrument (for the single endogenous regressor) ordered first.

Proof of Proposition 3.1.

$\boldsymbol{\beta}_{FS}$ collects the first-stage coefficients of interest:

$$\boldsymbol{\beta}_{FS} = \left(\mathbb{E}\left[\tilde{\boldsymbol{w}}_t \tilde{\boldsymbol{w}}_t'\right]\right)^{-1} \mathbb{E}\left[\tilde{\boldsymbol{w}}_t \tilde{\boldsymbol{X}}_t'\right], \tag{3.8}$$

where $\tilde{\boldsymbol{w}}_t \equiv M_Z \boldsymbol{w}_t$, $\tilde{\boldsymbol{X}}_t \equiv M_Z \boldsymbol{X}_t$ and M_Z is the residual-maker (annihilator) operator, i.e. it projects the objects into the space orthogonal to $\boldsymbol{Z}_{t-j}$. The fitted values from this regression are $\widehat{\boldsymbol{X}_t} = P_{wZ}\boldsymbol{X}_t$, where P_{wZ} is the projection onto the column space of $\boldsymbol{w}_t$ and $\boldsymbol{Z}_{t-j}$.

Consider the reduced-form IV regression:

$$y_{t+h} = \boldsymbol{\beta}_{RFh}' \boldsymbol{w}_t + \sum_{j=1}^{\infty} \boldsymbol{\gamma}_{RFh,j}' \boldsymbol{Z}_{t-j} + \epsilon_{RFh,t}. \tag{3.9}$$

The reduced-form coefficients of interest are collected by

$$\boldsymbol{\beta}_{RFh} = \left(\mathbb{E}\left[\tilde{\boldsymbol{w}}_t \tilde{\boldsymbol{w}}_t'\right]\right)^{-1} \mathbb{E}\left[\tilde{\boldsymbol{w}}_t \tilde{y}_{t+h}\right], \tag{3.10}$$

where $\tilde{\boldsymbol{y}}_{t+h} \equiv M_Z \boldsymbol{y}_{t+h}$. Standard 2SLS theory delivers

$$\boldsymbol{\beta}_{IVh} = \left(\mathbb{E}\left[P_{WZ}\boldsymbol{X}_t \boldsymbol{X}_t'\right]\right)^{-1} \mathbb{E}\left[P_{WZ}\boldsymbol{X}_t y_{t+h}\right]. \tag{3.11}$$

Assumption of invertibility for $\boldsymbol{\beta}_{FS}$ yields

$$\boldsymbol{\beta}_{IVh} = \boldsymbol{\beta}_{FS}^{-1} \boldsymbol{\beta}_{RFh}. \tag{3.12}$$

For $k = m = 1$ (a single instrument for a single endogenous regressor), $\beta_{IVh} = \frac{\beta_{RFh}}{\beta_{FS}}$, i.e. the standard result in Angrist and Pischke (2009), page 122.

Consider a recursively identified 2-block SVAR(∞), where $\boldsymbol{w}_t$ is ordered first and $(\boldsymbol{r}_t', \boldsymbol{X}_t', y_t, \boldsymbol{q}_t')'$ is ordered last. Invoking Proposition 1 in Plagborg-Møller and Wolf (2021) for each element of $\boldsymbol{\beta}_{FS}$ and $\boldsymbol{\beta}_{RFh}$ in equation (3.6) and (3.9), respectively, delivers

$$\begin{aligned} \tilde{\boldsymbol{\theta}}_{x0} &= \boldsymbol{\Lambda}_m \boldsymbol{\beta}_{FS} \\ \tilde{\boldsymbol{\theta}}_{yh} &= \boldsymbol{\Lambda}_m \boldsymbol{\beta}_{RFh}, \end{aligned} \tag{3.13}$$

where $\boldsymbol{\Lambda}_m$ is a diagonal $m \times m$ matrix containing the contemporaneously uncorrelated innovations (after partialling out controls) to the m instruments in equation (3.6) and (3.9). Specifically, let $\Lambda_m(i,i)$ denote the (i,i)-th element of $\boldsymbol{\Lambda}_m$. Set $\Lambda_m(i,i) \equiv \sqrt{\mathbb{E}(\tilde{w}_t(i)^2)}$, where $\tilde{w}_t(i) \equiv w_t(i) - \mathbb{E}(w_t(i)|\boldsymbol{r}_t, \{\boldsymbol{Z}_\tau\}_{-\infty<\tau<t})$, with $w_t(i)$ being the i-th instrument for $i = 1, \ldots, m$. As a result, we get

$$\begin{aligned} \boldsymbol{\beta}_{FS} &= \boldsymbol{\Lambda}_m^{-1} \tilde{\boldsymbol{\theta}}_{x0} \\ \boldsymbol{\beta}_{RFh} &= \boldsymbol{\Lambda}_m^{-1} \tilde{\boldsymbol{\theta}}_{yh}. \end{aligned} \tag{3.14}$$

Replacing $\boldsymbol{\beta}_{FS}$ and $\boldsymbol{\beta}_{RFh}$ in equation (3.12) with equation (3.14) yields

$$\begin{aligned}\boldsymbol{\beta}_{IVh} &= \left(\boldsymbol{\Lambda}_m^{-1}\tilde{\boldsymbol{\theta}}_{x0}\right)^{-1}\boldsymbol{\Lambda}_m^{-1}\tilde{\boldsymbol{\theta}}_{yh} \\ &= \tilde{\boldsymbol{\theta}}_{x0}^{-1}\tilde{\boldsymbol{\theta}}_{yh}.\end{aligned} \tag{3.15}$$

We can obtain (*ii*) by noting that Stock and Watson (2018) prove that $[\boldsymbol{\beta}_{IVh}]_j = \tilde{\boldsymbol{\Theta}}_{n_r+2\ j*h}$ under Assumption A3 and A4. ■

4 Simulation

To show the equivalence between LPs and VARs under non-stationarity, we simulate 1,000,000 data points for output and consumption from an RBC model (King and Rebelo, 1999) with technology, government spending, and preference shocks. We assume a unit root in the AR(1) process for technology, that results in a common stochastic trend for output and consumption. We focus on the output response to a technology shock and compare LP(p) with a bi-variate VAR(p) representation in levels. The shock is treated as observed and included (with its p lags) as an exogenous regressor in the VAR (Paul, 2020). We consider two cases: $p = 4$, and $p = 250$, with the latter used as an approximation of infinite lags. Figure 1 displays our results.

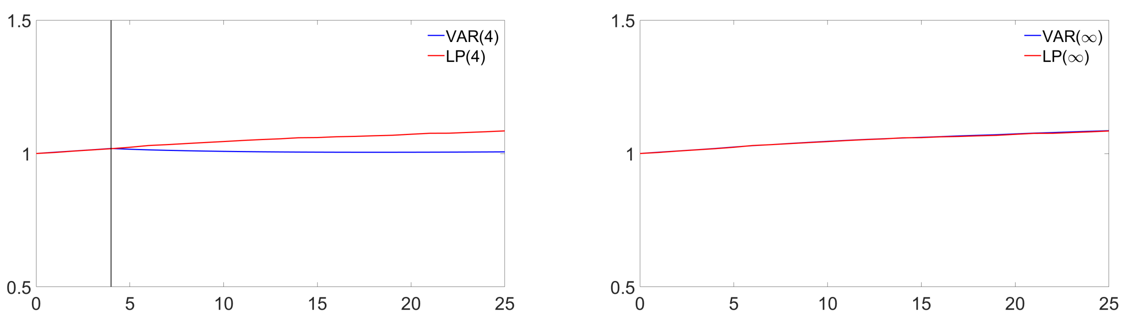

Figure 1: Output Response to a Positive Total Factor Productivity Shock

Note: LP and VAR impulse response estimands using simulated data from a RBC model. The vertical line in the left panel marks the horizon p after which the finite-lag-length LP(p) and VAR(p) estimands diverge. Shocks normalized to induce a one unit impact increase in output.

As formalized in Proposition 2.1, the VAR(∞) and LP(∞) estimands agree at all horizons. For finite lag lengths, the equivalence holds exactly up to horizon $h = p$.

5 Empirical Application

The following section illustrates the theoretical results of this paper through two applications, which document the approximate equivalence up to $h = p$ between IRFs estimated using LPs and VARs. For non-stationary variables, we evaluate the effects of oil price shocks. For multiple instruments, we focus on monetary policy and central bank information shocks.

5.1 Oil Price Shocks

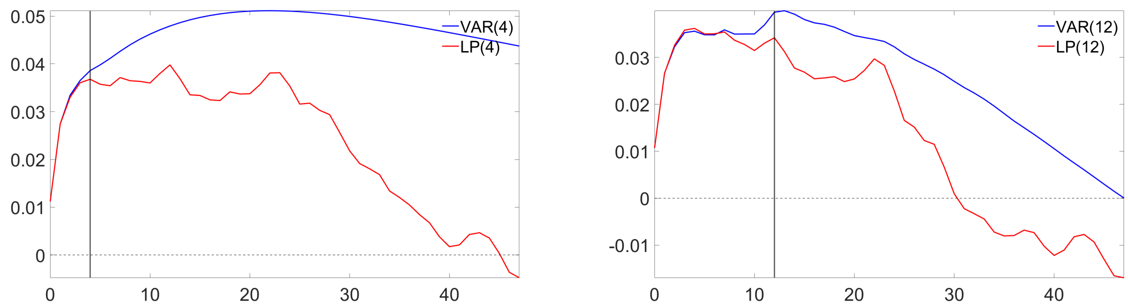

Figure 2: CPI Response to a Positive Oil Price Shock

Note: The vertical lines mark the horizon p after which the LP(p) and VAR(p) may diverge substantially. Shocks normalized to induce a one unit impact increase in oil price.

We consider the same specification as Känzig (2021), that includes the real price of oil, world oil production, world industrial production, US industrial production and US CPI. All series are $I(1)$ and enter the model in log-levels. Data are monthly and span the period 1974M1-2017M12. We estimate a VAR in levels with 3 cointegrating relations, as suggested by Johansen's (1988) test. We depart from Känzig (2021) and identify an oil price shock as the first innovation in the Cholesky decomposition. The IRFs from the VAR(p) are compared to those from LP(p), obtained by imposing the recursive ordering. For the sake of brevity, Figure 2 only shows the responses for CPI.

The left-hand plot considers $p = 4$, while the right-hand plot refers to the case with $p = 12$. As expected, a positive oil price shock increases consumer prices. In line with our theoretical argument, the responses are similar until horizon p, while differ substantially for $h > p$. For $p = 12$, the VAR response gets negative at longer horizons than what is being shown.

5.2 Monetary Policy and Information Shocks

We propose a monetary policy application based on Gertler and Karadi's (2015) specification, that includes output growth (log growth rate of industrial production), inflation (log growth rate of CPI), 1-year bond yield, SP500 and the Excess Bond Premium of Gilchrist and Zakrajšek (2012). Data are monthly and span the period 1990M1-2012M6. We employ two high-frequency proxies: one for a conventional monetary policy shock, the other for a central bank information

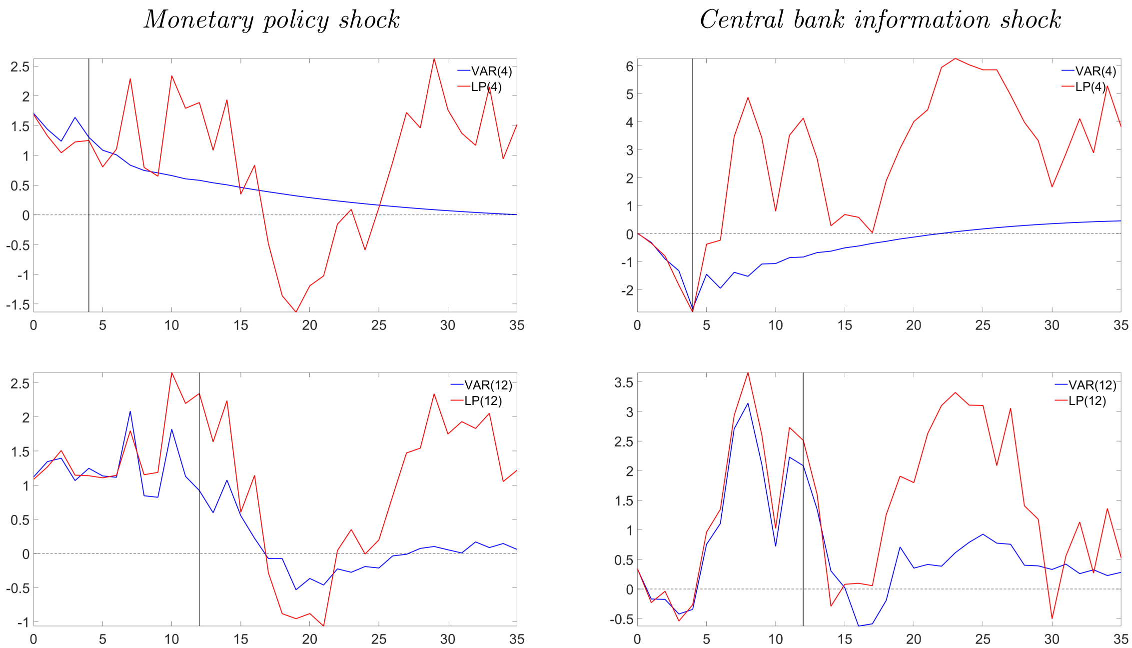

Figure 3: Excess Bond Premium Responses to a Contractionary Monetary Policy Shock (Left Panel) and to a Positive Central Bank Information Shock (Right Panel).

Note: The vertical lines mark the horizon p after which the LP(p) and VAR(p) may diverge substantially. Shocks normalized to induce a 100bp increase in the 1-year bond yield.

shock (Jarociński and Karadi, 2020).[4] The regressors in $\boldsymbol{X}_t$ that we instrument with the monetary policy and central bank information proxies are the 1-year bond yield and the SP500, respectively. Our focus is on the response of Excess Bond Premium.

As further discussed in Jarociński and Karadi (2020), the two shocks trigger different short-term effects on financial markets (Figure 3). A contractionary monetary policy shock leads to an increase in the excess bond premium, while a positive central bank information shock pushes it in the opposite direction. In line with the content of Proposition 3.1, the impulse responses estimated by LP-IV and "internal-instrument" VAR (computed by using equation 3.15) agree at short horizons, with the equivalence that holds approximately up to horizon p.

6 Conclusion

Firstly, this paper shows that the assumption of stationarity is not required for the asymptotic equivalence between LP IRFs and VAR IRFs (Plagborg-Møller and Wolf, 2021). We establish this result under (possibly cointegrated) unit roots by employing the Granger's representation

[4] The two proxies are available on Marek Jarocinski's website and computed by using the median rotation-based decomposition described in Jarociński (2022).

theorem.

Secondly, while Plagborg-Møller and Wolf (2021) described the equivalence of IRFs between LP-IV and “internal-instrument” VAR for a *single* instrument, we extend their finding by proving that IRFs of LP-IV with a *multiplicity* of instruments for multiple endogenous regressors can be obtained by running a recursively block-identified SVAR, where the block of instruments is ordered first.

The results hold for unconstrained lag structures and are independent of the identification strategy. For a given lag length p, equivalence holds up to horizon p.

Simulations from a RBC model illustrate the equivalence. For unit roots, an application to oil price shocks shows that IRFs from LP(p) and VAR(p) coincide up to horizon p. A further application based on the effects of monetary policy and central bank information shocks display the equivalence of IRFs between a LP-IV, where two external instruments are used to identify the shocks, and a recursively block-identified SVAR, where the two instruments are ordered first.

References

Angrist, J. D. and J.-S. Pischke (2009): *Mostly harmless econometrics: An empiricist's companion*, Princeton university press.

Beveridge, S. and C. R. Nelson (1981): "A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the 'business cycle'," *Journal of Monetary economics*, 7, 151–174.

Bollerslev, T., D. Osterrieder, N. Sizova, and G. Tauchen (2013): "Risk and return: Long-run relations, fractional cointegration, and return predictability," *Journal of Financial Economics*, 108, 409–424.

Bruns, M., H. Lütkepohl, and J. McNeil (2025): "Avoiding unintentionally correlated shocks in proxy vector autoregressive analysis," *Journal of Business & Economic Statistics*, 1–21.

Carlini, F. and P. Santucci de Magistris (2019): "On the identification of fractionally cointegrated VAR models with the F (d) condition," *Journal of Business & Economic Statistics*, 37, 134–146.

Chevillon, G. (2007): "Direct multi-step estimation and forecasting," *Journal of Economic Surveys*, 21, 746–785.

Fanelli, L. and A. Marsi (2022): "Sovereign spreads and unconventional monetary policy in the euro area: A tale of three shocks," *European Economic Review*, 150, 104281.

Franchi, M. and P. Paruolo (2019): "A general inversion theorem for cointegration," *Econometric Reviews*.

Frisch, R. (1933): "Propagation Problems and Impulse Problems in Dynamic Economics 1," in *Business Cycle Theory, Part I Volume 4*, Routledge, 305–336.

Gertler, M. and P. Karadi (2015): "Monetary Policy Surprises, Credit Costs, and Economic Activity," *American Economic Journal: Macroeconomics*, 7, 44–76.

Gilchrist, S. and E. Zakrajšek (2012): "Credit Spreads and Business Cycle Fluctuations," *American Economic Review*, 102, 1692–1720.

Gregoir, S. (1999): "Multivariate time series with various hidden unit roots, part I," *Econometric Theory*, 15, 435–468.

Hou, C. (2024): "Large Bayesian SVARs with linear restrictions," *Journal of Econometrics*, 244, 105850.

Howlett, P. G. (1982): "Input retrieval in finite dimensional linear systems," *The ANZIAM Journal*, 23, 357–382.

Huber, F., C. Matthes, and M. Pfarrhofer (2024): "General Seemingly Unrelated Local Projections," *arXiv preprint arXiv:2410.17105*.

Jarociński, M. (2022): "Central bank information effects and transatlantic spillovers," *Journal of International Economics*, 139, 103683.

Jarociński, M. and P. Karadi (2020): "Deconstructing Monetary Policy Surprises—The Role of Information Shocks," *American Economic Journal: Macroeconomics*, 12, 1–43.

Johansen, S. (1988): "Statistical analysis of cointegration vectors," *Journal of Economic Dynamics and Control*, 12, 231–254.

——— (1991): "Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models," *Econometrica: journal of the Econometric Society*, 1551–1580.

——— (1995a): *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press.

——— (1995b): "A stastistical analysis of cointegration for I (2) variables," *Econometric theory*, 11, 25–59.

——— (2008): "A representation theory for a class of vector autoregressive models for fractional processes," *Econometric Theory*, 24, 651–676.

Johansen, S. and M. Ø. Nielsen (2012): "Likelihood inference for a fractionally cointegrated vector autoregressive model," *Econometrica*, 80, 2667–2732.

Jordà, Ò. (2005): "Estimation and inference of impulse responses by local projections," *American economic review*, 95, 161–182.

Kilian, L. and Y. J. Kim (2011): "How reliable are local projection estimators of impulse responses?" *Review of Economics and Statistics*, 93, 1460–1466.

Kilian, L. and H. Lütkepohl (2017): *Structural vector autoregressive analysis*, Cambridge University Press.

King, R. G. and S. T. Rebelo (1999): "Resuscitating real business cycles," in *Handbook of macroeconomics*, ed. by J. Taylor and M. Woodford, Elsevier Science.

Känzig, D. R. (2021): "The Macroeconomic Effects of Oil Supply News: Evidence from OPEC Announcements," *American Economic Review*, 111, 1092–1125.

Lakdawala, A. (2019): "Decomposing the effects of monetary policy using an external instruments SVAR," *Journal of applied econometrics*, 34, 934–950.

Li, D., M. Plagborg-Møller, and C. K. Wolf (2024): "Local projections vs. vars: Lessons from thousands of dgps," *Journal of Econometrics*, 105722.

Ludwig, J. F. (2024): "Local Projections Are VAR Predictions of Increasing Order," *Available at SSRN 4882149*.

Lunsford, K. G. (2015): "Identifying structural VARs with a proxy variable and a test for a weak proxy," Tech. rep.

Marcellino, M., J. H. Stock, and M. W. Watson (2006): "A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series," *Journal of econometrics*, 135, 499–526.

McNown, R. and M. S. Wallace (1994): "Cointegration tests of the monetary exchange rate model for three high-inflation economies," *Journal of Money, Credit and Banking*, 26, 396–411.

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, 1212–47.

Paul, P. (2020): "The Time-Varying Effect of Monetary Policy on Asset Prices," *The Review of Economics and Statistics*, 102, 690–704.

Pesaran, M. H., A. Pick, and A. Timmermann (2011): "Variable selection, estimation and inference for multi-period forecasting problems," *Journal of Econometrics*, 164, 173–187.

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

Ramey, V. A. (2016): "Macroeconomic shocks and their propagation," *Handbook of macroeconomics*, 2, 71–162.

Sims, C. A. (1980): "Macroeconomics and reality," *Econometrica: journal of the Econometric Society*, 1–48.

Stock, J. H. and M. W. Watson (1993): "A simple estimator of cointegrating vectors in higher order integrated systems," *Econometrica*, 783–820.

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

Tschernig, R., E. Weber, and R. Weigand (2013): "Long-run identification in a fractionally integrated system," *Journal of Business & Economic Statistics*, 31, 438–450.

Yang, Y. and R. Zhang (2025): "Combining Proxies and Narrative Sign Restrictions: Revisiting the Effects of Technology Shocks," *Journal of Applied Econometrics*.

Appendix to: "More on VARs and Local Projections Equivalence: Unit Roots and Multiple Instruments"

The Appendix contains generalizations of the equivalence between IRFs of VARs and LPs to higher order of (possibly fractional) integration and deterministic trends (Appendix A) and shows how a LP-IV with over-identification can be estimated by a recursively identified SVAR(∞) (Appendix B).

A Higher order of integration, fractional integration and deterministic trends

In the following extensions, we extend the LP-VAR equivalence to higher order of (possibly fractional) integration for $\boldsymbol{z}_t$ (Corollary A.1) and deterministic trends (Corollary A.2). The intuition being that, while the Frisch-Waugh theorem holds in those settings, we borrow formulations of the Granger's representation theorem from existing literature, e.g. Franchi and Paruolo (2019), for higher order of (possibly fractional) integration and processes with deterministic trends.

In order to simplify some algebra, and without loss of generality, we (i) employ the Laurent series expansion around $z = 1$ in powers of $1 - z$ rather than $z - 1$ as in the main text and (ii) focus on $c_r > 0$.

As in Section 2.1, partition $\boldsymbol{z}_t \equiv (\boldsymbol{r}_t', x_t, y_t, \boldsymbol{q}_t')'$ and define the LP and VAR impulse responses $\{\beta_h\}_{h\geq 0}$ and $\{\theta_h\}_{h\geq 0}$ as in Definition 2.1 and 2.2, respectively. In particular, θ_h refers to the response of y_{t+h} to the unit-variance structural shock η_{xt} in the structural representation $A(L)\boldsymbol{z}_t = \boldsymbol{B}\boldsymbol{\eta}_t$, where $\mathbb{E}(\boldsymbol{\eta}_t\boldsymbol{\eta}_t') \equiv \boldsymbol{I}$ and η_{xt} is the component of $\boldsymbol{\eta}_t$ corresponding to x_t under the recursive ordering (with $\boldsymbol{r}_t$ first and $\boldsymbol{q}_t$ last), as in Section 2.1.

Corollary A.1 *(Higher order of possibly fractional integration) Consider generic LP* $y_{t+h} = \beta_h x_t + \boldsymbol{\omega}_h'\boldsymbol{r}_t + \sum_{j=1}^{\infty}\boldsymbol{\gamma}_{hj}'\boldsymbol{z}_{t-j} + \epsilon_{ht}$ *and SVAR* $\boldsymbol{A}(L)\boldsymbol{z}_t = \boldsymbol{B}\boldsymbol{\eta}_t$*, with* $\boldsymbol{z}_t \sim I(d)$*,* $d \in \mathbb{R}$ *and* $c_r > 0$*. Under the regularity conditions in Theorem 4.3 of Franchi and Paruolo (2019),* $\theta_h = \sqrt{\mathbb{E}(\tilde{x}_t^2)}\beta_h$ *for* $h = 1, 2, \dots,$ *where* $\tilde{x}_t \equiv x_t - \mathbb{E}(x_t|\boldsymbol{r}_t, \{\boldsymbol{z}_\tau\}_{-\infty<\tau<t})$*.*

Proof.

First, assume that $d \in \mathbb{N}_{\geq 1}$, where $\mathbb{N}_{\geq 1} \equiv \{1, 2, \dots\}$. According to the Frisch-Waugh theorem, we obtain

$$\beta_h = \frac{\text{cov}(y_{t+h}, \tilde{x}_t)}{\mathbb{E}(\tilde{x}_t^2)}. \tag{A.1}$$

Consider the problem of inversion of the matrix function

$$\boldsymbol{A}(z) = \sum_{k=0}^{\infty} \boldsymbol{A}^{(k)}(1-z)^k, \ \boldsymbol{A}^{(k)} \in \mathbb{R}^{n\times n}, \ \boldsymbol{A}^{(0)} = \boldsymbol{A}(1) \neq \boldsymbol{0}, \ |\boldsymbol{A}^{(0)}| = 0, \tag{A.2}$$

around the singularity $z = 1$. The inversion of $\boldsymbol{A}(z)$ around the singularity delivers an inverse with a pole of order $d = 1, 2, \ldots,$ at $z = 1$. Unlike the result in the main text, the order of the pole is allowed to be non-simple, i.e. $d > 1$; in a SVAR setting, that corresponds to order of integration higher than one.

Furthermore, let us introduce the integral operator $\mathcal{S}$, i.e. the inverse of the difference operator Δ up to a constant. In particular, for a generic process g_t, we define $\mathcal{S}$ as follows:

$$\mathcal{S}g_t \equiv \mathbf{1}_{t\geq 1}\sum_{i=1}^{t} g_i - \mathbf{1}_{t\leq -1}\sum_{i=t+1}^{0} g_i, \tag{A.3}$$

where $\mathbf{1}_{\bullet}$ is the indicator function. $\mathcal{S}$ satisfies[5]

$$\Delta\mathcal{S}g_t = g_t, \ \mathcal{S}\Delta g_t = g_t - g_0, \ \mathcal{S}_1 = t, \ \mathcal{S}^{\tilde{\gamma}}g_t \text{ means applying } \mathcal{S} \ \tilde{\gamma} \text{ times.} \tag{A.4}$$

Given the generic SVAR $\boldsymbol{A}(L)\boldsymbol{z}_t = \boldsymbol{B}\boldsymbol{\eta}_t$ with $\boldsymbol{z}_t \sim I(d)$, according to the Theorem 4.3 in Franchi and Paruolo (2019), we obtain the following Granger's representation for $\boldsymbol{z}_t$:

$$\boldsymbol{z}_t = \sum_{k=0}^{d-1} \boldsymbol{C}_k\boldsymbol{S}_{d-k,t} + \boldsymbol{C}_d(L)\boldsymbol{B}\boldsymbol{\eta}_t + \tilde{\boldsymbol{z}}_0^*, \tag{A.5}$$

where $\boldsymbol{S}_{ht} \equiv \mathcal{S}^h\boldsymbol{B}\boldsymbol{\eta}_t$ and $\tilde{\boldsymbol{z}}_0^*$ is a deterministic component. Note that $\boldsymbol{A}(z)^{-1} = \boldsymbol{C}(z) = \sum_{k=0}^{d-1}\boldsymbol{C}_k(1-z)^{k-d} + \boldsymbol{C}_d(z)$ and $\boldsymbol{C}_k$ are the Laurent coefficients. Following (Franchi and Paruolo, 2019) yields i) $\boldsymbol{C}_d(L)\boldsymbol{B}\boldsymbol{\eta}_t \sim I(0)$ and ii) $\sum_{k=0}^{d-1}\boldsymbol{C}_k\boldsymbol{S}_{d-k,t}$ illustrates the permanent component. Let us apply the argument of Proposition 2.1 to equation (A.1) and (A.5):

- the Granger-type representation implies $\theta_h = \text{cov}(y_{t+h}, \eta_{xt})$;
- the identified shock equals the normalized unexpected component $\eta_{xt} = \frac{\tilde{x}_t}{\sqrt{\mathbb{E}(\tilde{x}_t^2)}}$.

Combining the last equation with A.1 delivers the desired result.

Second, assume that $d \in \mathbb{R}$, i.e. fractional integration. The fractional difference operator is $\Delta^d \equiv (1-L)^d$ and we assume $\Delta^d\boldsymbol{z}_t \sim I(0)$. We need to replace $\boldsymbol{z}_t$ with $(1-L)^d\boldsymbol{z}_t$ in the SVAR specification. Granger's representation in equation (A.5) applies with the appropriate fractional operators (formally, see Section 4.5 in Franchi and Paruolo (2019)), implying that the reasoning above holds for fractional integration as well. ■

[5]Formally, see Properties 2.1, 2.2 in Gregoir (1999) and Lemma A.2 in Appendix 2 of Franchi and Paruolo (2019).

Next corollary establishes the result under deterministic trends. We build it upon previous notation and introduce the class of vector polynomials of order u in t, denoted by $\mathcal{P}_{un}(t)$: $\boldsymbol{p}_u(t) = \sum_{i=0}^{u} \boldsymbol{c}_i t^i \in \mathcal{P}_{un}$ with $\boldsymbol{c}_i \in \mathbb{R}^n$.

Consider the following set-up with deterministic trends:

$$\begin{cases} \boldsymbol{\mu}_t \in \mathcal{P}_{u+\tilde{j},n}, \ \boldsymbol{z}_t \sim I(d), \ d \in \mathbb{N}_{\geq 1} \\ y_{t+h} = \beta_h x_t + \boldsymbol{\omega}_h' \boldsymbol{r}_t + \sum_{j=1}^{\infty} \boldsymbol{\gamma}_{hj}' \boldsymbol{z}_{t-j} + \boldsymbol{\mu}_t + \epsilon_{ht}, \\ \boldsymbol{A}(L)\boldsymbol{z}_t = \boldsymbol{B}\boldsymbol{\eta}_t + \boldsymbol{\mu}_t, \end{cases} \tag{A.6}$$

where $\tilde{j} = 0, 1, \ldots, d-1$ is the integration order of cointegrating relationships.

Corollary A.2 *(Deterministic Trends) Consider generic LP-VAR setting in system (A.6) and $c_r > 0$. Under the regularity conditions in Theorem 4.8 of Franchi and Paruolo (2019), $\theta_h = \sqrt{\mathbb{E}(\tilde{x}_t^2)}\beta_h$ for $h = 1, 2, \ldots,$ where $\tilde{x}_t \equiv x_t - \mathbb{E}(x_t | \boldsymbol{r}_t, \{\boldsymbol{z}_\tau\}_{-\infty<\tau<t}, \boldsymbol{\mu}_t)$.*

Proof. By Frisch-Waugh, $\beta_h = \text{cov}(y_{t+h}, \tilde{x}_t)/\mathbb{E}[(\tilde{x}_t)^2]$. Theorem 4.8, ii.1) and ii.2) of Franchi and Paruolo (2019) yields a Granger-type decomposition for $\boldsymbol{z}_t$ with (i) a stochastic component driven by the structural shocks $\boldsymbol{\eta}_t$, (ii) a stationary element and (iii) deterministic terms collected in $\boldsymbol{\mu}_t$ (and, if present, in deterministic components of the cointegrating relationships). Hence the same argument as in Proposition 2.1 implies $\theta_h = \text{cov}(y_{t+h}, \eta_{x,t})$. Under the recursive ordering, $\eta_{x,t} = \tilde{x}_t / \sqrt{\mathbb{E}[(\tilde{x}_t)^2]}$, so combining these displays gives the claim. ■

B Over-identification

Consider the following LP regression:

$$y_{t+h} = \boldsymbol{\beta}_h' \boldsymbol{X}_t + \sum_{j=1}^{\infty} \gamma_{hj}' \boldsymbol{z}_{t-j} + \epsilon_{ht}, \tag{B.1}$$

where $\boldsymbol{X}_t \in \mathbb{R}^k$ contains k endogenous regressors instrumented by $\boldsymbol{w}_t \in \mathbb{R}^m$, with $m > k$. Let $\boldsymbol{M}_Z$ denote the residual-maker associated with the controls in the LP-IV setup and define the residualized variables $\tilde{\boldsymbol{w}}_t \equiv \boldsymbol{M}_Z \boldsymbol{w}_t$, $\tilde{\boldsymbol{X}}_t \equiv \boldsymbol{M}_Z \boldsymbol{X}_t$, $\tilde{y}_{t+h} \equiv \boldsymbol{M}_Z y_{t+h}$.

We construct a k-dimensional instrument by forming a linear combination of $\boldsymbol{w}_t$:

$$\boldsymbol{w}_t^* \equiv \boldsymbol{C}' \boldsymbol{w}_t, \qquad \boldsymbol{C} \in \mathbb{R}^{m \times k}, \ \text{rank}(\boldsymbol{C}) \equiv k. \tag{B.2}$$

Since $\boldsymbol{M}_Z$ is linear, $\tilde{\boldsymbol{w}}_t^* \equiv \boldsymbol{M}_Z \boldsymbol{w}_t^* = \boldsymbol{C}' \tilde{\boldsymbol{w}}_t$.

To preserve identification strength, $\boldsymbol{C}$ should capture the component of $\tilde{\boldsymbol{w}}_t$ that is most predictive of $\tilde{\boldsymbol{X}}_t$. Consider the (population) first-stage regression:

$$\boldsymbol{X}_t = \boldsymbol{\beta}_{FS}' \boldsymbol{w}_t + \sum_{j=1}^{\infty} \gamma_{FSj}' \boldsymbol{Z}_{t-j} + \boldsymbol{\epsilon}_{FSt}, \tag{B.3}$$

so that, after partialling out controls, the first-stage coefficient is

$$\boldsymbol{\beta}_{FS} = \Big(\mathbb{E}[\tilde{\boldsymbol{w}}_t \tilde{\boldsymbol{w}}_t']\Big)^{-1} \mathbb{E}[\tilde{\boldsymbol{w}}_t \tilde{\boldsymbol{X}}_t'], \tag{B.4}$$

where $\mathbb{E}[\tilde{\boldsymbol{w}}_t \tilde{\boldsymbol{w}}_t']$ is assumed non-singular.

Define

$$\boldsymbol{C} \equiv \mathbb{E}[\tilde{\boldsymbol{w}}_t \tilde{\boldsymbol{X}}_t'] \in \mathbb{R}^{m \times k}, \tag{B.5}$$

and assume $\operatorname{rank}(\boldsymbol{C}) = k$ (equivalently, $\operatorname{rank}(\boldsymbol{\beta}_{FS}) = k$).

Why the reduction $\boldsymbol{w}_t \mapsto \boldsymbol{w}_t^*$ is w.l.o.g. Stack $\tilde{\boldsymbol{W}} \equiv (\tilde{\boldsymbol{w}}_1, \ldots, \tilde{\boldsymbol{w}}_T)' \in \mathbb{R}^{T \times m}$ and $\tilde{\boldsymbol{X}} \equiv (\tilde{\boldsymbol{X}}_1, \ldots, \tilde{\boldsymbol{X}}_T)' \in \mathbb{R}^{T \times k}$. The 2SLS procedure depends on instruments only through the first-stage fitted values $\widehat{\boldsymbol{X}} = P_{\tilde{W}} \tilde{\boldsymbol{X}}$. Let

$$\tilde{\boldsymbol{W}}^* \equiv (\tilde{\boldsymbol{w}}_1^*, \ldots, \tilde{\boldsymbol{w}}_T^*)' = \tilde{\boldsymbol{W}} \boldsymbol{C} \in \mathbb{R}^{T \times k}.$$

If $\operatorname{col}(\tilde{\boldsymbol{W}}^*) = \operatorname{col}(P_{\tilde{W}} \tilde{\boldsymbol{X}})$, then $P_{\tilde{W}^*} \tilde{\boldsymbol{X}} = P_{\tilde{W}} \tilde{\boldsymbol{X}}$ and therefore replacing $\tilde{\boldsymbol{w}}_t$ with $\boldsymbol{w}_t^*$ leaves the LP–IV estimand unchanged.

A convenient sufficient choice is to take any $\boldsymbol{C}$ such that $\operatorname{col}(\boldsymbol{C}) = \operatorname{col}(\boldsymbol{\beta}_{FS})$. In particular, using $\boldsymbol{C} \equiv \mathbb{E}[\tilde{\boldsymbol{w}}_t \tilde{\boldsymbol{X}}_t']$ satisfies $\operatorname{col}(\boldsymbol{C}) = \operatorname{col}(\boldsymbol{\beta}_{FS})$ because left-multiplication by the invertible matrix $\mathbb{E}[\tilde{\boldsymbol{w}}_t \tilde{\boldsymbol{w}}_t']^{-1}$ does not change column space. Hence $\tilde{\boldsymbol{W}} \boldsymbol{C}$ spans the same k-dimensional fitted-value space as the original first stage, and we may work with the exactly-identified instrument vector $\boldsymbol{w}_t^*$ without loss of generality.

Consequently, Proposition 3.1 can be applied to the k-dimensional instrument $\boldsymbol{w}_t^*$ (ordered first in the recursively block-identified SVAR), yielding the same LP–IV impulse responses as in the over-identified case.