Online Appendix for:
Local Projections and VARs Estimate the Same Impulse Responses

Mikkel Plagborg-Møller Christian K. Wolf

July 10, 2019

This online appendix contains supplemental material for the article "Local Projections and VARs Estimate the Same Impulse Responses". Specifically, we state and prove the sample asymptotic equivalence result for local projection and VAR impulse response estimators.

Any references to equations or sections that are not preceded by "B." refer to the main article.

B.1 Asymptotic equivalence of LP and VAR impulse response estimators

Complementing the informal discussion in Section 3, here we prove that local projections and recursively identified VARs estimate nearly the same impulse response functions *in sample*, provided the lag lengths used in the specifications are large enough. Assume we observe the data $w_1, w_2, \dots, w_T$ (recall the notation in Section 2.1). For all lag lengths $p \leq T$, define the following:

- Let $\hat{x}_t(p)$ be the residual from a regression of x_t on an intercept, r_t, and $w_{t-1}, \dots, w_{t-p}$.
- Let $\hat{\beta}_h(p)$ denote the OLS estimator of the local projection parameter β_h in the sample version of regression equation (1), where we include p lags of w_t on the right-hand side

instead of the infeasible infinite distibuted lag. By the Frisch-Waugh theorem,

$$\hat{\beta}_h(p) = \frac{\sum_{t=p+1}^{T-h} y_{t+h}\hat{x}_t(p)}{\sum_{t=p+1}^{T-h} \hat{x}_t(p)^2}.$$

- Let $\hat{\theta}_h(p)$ denote the horizon-h impulse response of y_t to an innovation in x_t in a Cholesky-identified VAR(p) model (with intercept) estimated by least squares on the data points $t = p+1, p+2, \dots, T$.

In detail, the VAR estimator $\hat{\theta}_h(p)$ is defined as follows. Let $\hat{A}_\ell(p)$ denote the usual least-squares VAR(p) coefficient matrix estimator at lag ℓ, and let $\hat{c}(p)$ denote the corresponding intercept vector estimator. Let $\hat{u}_t(p)$ denote the residual vector. Define the innovation covariance matrix estimator $\hat{\Sigma}(p) \equiv \frac{1}{T-p}\sum_{t=p+1}^{T} \hat{u}_t(p)\hat{u}_t(p)'$ and let $\hat{\Sigma}(p) = \hat{B}(p)\hat{B}(p)'$ denote its lower triangular Cholesky decomposition. Define the reduced-form impulse response matrices by $\hat{C}_0(p) = I_{n_w}$ and $\hat{C}_m(p) = \sum_{\ell=1}^{m} \hat{A}_\ell(p)\hat{C}_{m-\ell}(p)$ for $m = 1, \dots, h$. Then $\hat{\theta}_h(p)$ is given by the (n_r+2, n_r+1) element of $\hat{C}_h(p)\hat{B}(p)$.

Note that the VAR(p) residuals

$$\hat{u}_t(p) \equiv w_t - \hat{c}(p) - \sum_{\ell=1}^{p} \hat{A}_\ell(p) w_{t-\ell}, \quad t = p+1, p+2, \dots, T,$$

satisfy

$$\sum_{t=p+1}^{T} \hat{u}_t(p) = 0_{n_w \times 1}, \quad \sum_{t=p+1}^{T} \hat{u}_t(p) w_{t-\ell} = 0_{n_w \times n_w}, \quad \ell = 1, 2, \dots, p. \tag{B.1}$$

We adopt the convention that $\hat{u}_t(p) \equiv 0$ whenever $t \leq p$.

We are now ready to state the near-equivalence result for LP and VAR impulse response estimators. Let $\|\cdot\|$ denote the Frobenius norm.

Proposition B.1. *In the following, the lag length $p = p(T)$ used for estimation is implicitly a function of T. Assume the following:*

i) *$\{w_t\}$ is covariance stationary and has a VAR(∞) representation (2), where $\sum_{\ell=1}^{\infty}\|A_\ell\| < \infty$, and the Wold innovations u_t have finite and positive definite covariance matrix Σ. (We do not assume that the innovations are necessarily Gaussian.)*

ii) *$\|\hat{c}(p) - c\| = o_p(1)$, $\|\hat{A}(p) - A(p)\| = o_p(1)$, and $\|\hat{\Sigma}(p) - \Sigma\| = o_p(1)$, where we have defined $\hat{A}(p) \equiv (\hat{A}_1(p), \dots, \hat{A}_p(p))$ and $A(p) \equiv (A_1, \dots, A_p)$.*

Then

$$\hat{\theta}_h(p) = \frac{\frac{1}{T-p}\sum_{t=p+1}^{T-h} y_{t+h}\hat{x}_t(p)}{\left(\frac{1}{T-p}\sum_{t=p+1}^{T}\hat{x}_t(p)^2\right)^{1/2}} + O_p(\hat{R}(p)),$$

where

$$\hat{R}(p) \equiv \frac{\max\{1, \sup_{1\le t\le T}\|w_t\|\}^2}{T-p} + \left(\sum_{\ell=p-h+1}^{p}\|\hat{A}_\ell(p)\|^2\right)^{1/2}.$$

Thus, the VAR impulse response estimator $\hat{\theta}_h(p)$ approximately equals the LP impulse response estimator $\hat{\beta}_h(p)$ up to a scale factor that does not depend on the horizon h. The approximation error is of an order $O_p(\hat{R}(p))$ that is likely to be small unless the data is so persistent that the estimated VAR coefficients at the very longest lags are non-negligible.

Assumptions (i) and (ii) of the proposition are easily satisfied under standard nonparametric regularity conditions on the data generating process and a restriction on how quickly the lag length p can grow with T. See for example Lewis & Reinsel (1985) and Gonçalves & Kilian (2007).

B.1.1 Proof of Proposition B.1

We split the proof into several steps.

Step 1. We will show that $\sum_{\ell=1}^{p}\|\hat{A}_\ell(p)\| = O_p(1)$. The statement follows from

$$\sum_{\ell=1}^{p}\|\hat{A}_\ell(p)\| \le \sum_{\ell=1}^{p}\|A_\ell\| + \sum_{\ell=1}^{p}\|\hat{A}_\ell(p) - A_\ell\| \le \sum_{\ell=1}^{\infty}\|A_\ell\| + \|\hat{A}(p) - A(p)\|$$

and then exploiting assumptions (i) and (ii).

Step 2. We will show that $\sup_{p+1\le t\le T}\|\hat{u}_t(p)\| = \sup_{1\le t\le T}\|w_t\| \times O_p(1)$. Observe that

$$\begin{aligned}\sup_{p+1\le t\le T}\|\hat{u}_t(p)\| &= \sup_{p+1\le t\le T}\left\|w_t - \sum_{\ell=1}^{p}\hat{A}_\ell(p)w_{t-\ell}\right\| \\ &\le \left(\sup_{1\le t\le T}\|w_t\|\right)\left(1 + \sum_{\ell=1}^{p}\|\hat{A}_\ell(p)\|\right).\end{aligned}$$

Step 1 then gives the desired result.

Step 3. We will show that, for any $m = 0, 1, \dots, h$,

$$\frac{1}{T-p} \sum_{t=p+1}^{T} \hat{u}_{t-m}(p) = O_p\left(\frac{\sup_t \|w_t\|}{T-p}\right).$$

We have

$$\sum_{t=p+1}^{T} \hat{u}_{t-m}(p) = \sum_{t=p+1}^{T} \hat{u}_t(p) - \sum_{t=T-m+1}^{T} \hat{u}_t(p).$$

The first sum on the right-hand side is exactly zero by the orthogonality conditions (B.1). The second sum consists of m terms, each of which is $O_p(\sup_t \|w_t\|)$ by Step 2.

Step 4. We will show that, for any $m = 1, 2, \dots, h$,

$$\frac{1}{T-p} \sum_{t=p+1}^{T} \hat{u}_t(p)\hat{u}_{t-m}(p)' = O_p\left(\left(\sum_{\ell=p-h+1}^{p} \|\hat{A}_\ell(p)\|^2\right)^{1/2}\right).$$

$\hat{u}_{t-m}(p)$ is a linear function of $w_{t-m}, w_{t-1-m}, \dots, w_{t-p-m}$. By the orthogonality conditions (B.1), $\hat{u}_t(p)$ is orthogonal to $w_{t-m}, w_{t-1-m}, \dots, w_{t-p}$ (and a constant). Thus,

$$\begin{aligned}
\left\|\frac{1}{T-p} \sum_{t=p+1}^{T} \hat{u}_t(p)\hat{u}_{t-m}(p)'\right\| &= \left\|\frac{1}{T-p} \sum_{t=p+m+1}^{T} \hat{u}_t(p) \sum_{\ell=p-m+1}^{p} w'_{t-m-\ell}\hat{A}_\ell(p)'\right\| \\
&\leq \left(\sum_{\ell=p-m+1}^{p} \|\hat{A}_\ell(p)\|^2\right)^{1/2} \left(\sum_{\ell=p-m+1}^{p} \left\|\frac{1}{T-p} \sum_{t=p+m+1}^{T} \hat{u}_t(p)w'_{t-m-\ell}\right\|^2\right)^{1/2}.
\end{aligned}$$

Note that $\sum_{\ell=p-m+1}^{p} \|\hat{A}_\ell(p)\|^2 \leq \sum_{\ell=p-h+1}^{p} \|\hat{A}_\ell(p)\|^2$ since $m \leq h$. Finally,

$$\begin{aligned}
\left\|\frac{1}{T-p} \sum_{t=p+m+1}^{T} \hat{u}_t(p)w'_{t-m-\ell}\right\| &\leq \left(\frac{1}{T-p} \sum_{t=p+m+1}^{T} \|\hat{u}_t(p)\|^2\right)^{1/2} \left(\frac{1}{T-p} \sum_{t=p+m+1}^{T} \|w_{t-m-l}\|^2\right)^{1/2} \\
&\leq \left(\frac{1}{T-p} \sum_{t=p+1}^{T} \|\hat{u}_t(p)\|^2\right)^{1/2} \left(\frac{1}{T-p} \sum_{t=1}^{T} \|w_t\|^2\right)^{1/2} \\
&\leq \|\hat{\Sigma}(p)\| \left(\frac{1}{T-p} \sum_{t=1}^{T} \|w_t\|^2\right)^{1/2}.
\end{aligned}$$

The first factor on the right-hand side above is $O_p(1)$ by assumption (ii), while the second factor is $O_p(1)$ since $E\|w_t\|^2 < \infty$.

Step 5. Let $\ell, m \geq 0$ satisfy $m \leq h$ and $\ell \leq p$. If $m \leq \ell$, then

$$\frac{1}{T-p}\sum_{t=p+1}^{T} \hat{u}_{t-m}(p)w'_{t-\ell} = \frac{1}{T-p}\sum_{t=p+1}^{T} \hat{u}_t(p)w'_{t-(\ell-m)} + O_p\left(\frac{\sup_t \|w_t\|^2}{T-p}\right), \tag{B.2}$$

while if $m > \ell$, then

$$\frac{1}{T-p}\sum_{t=p+1}^{T} \hat{u}_{t-m}(p)w'_{t-\ell} = \frac{1}{T-p}\sum_{t=p+1}^{T} \hat{u}_{t-(m-\ell)}(p)w'_t + O_p\left(\frac{\sup_t \|w_t\|^2}{T-p}\right), \tag{B.3}$$

where the $O_p(\cdot)$ terms are uniform in ℓ and m. Claim (B.3) is proven in the same way as (B.2), so we only prove the latter. Simply note that

$$\sum_{t=p+1}^{T} \hat{u}_{t-m}(p)w'_{t-\ell} = \sum_{t=p+1}^{T} \hat{u}_t(p)w'_{t-(\ell-m)} - \sum_{t=T-m+1}^{T} \hat{u}_t(p)w'_{t-(\ell-m)},$$

and the second sum consists of m terms, each of which is $O_p(\sup_t \|w_t\|^2)$ by Step 2.

Step 6. We will show that, for any $\ell, m \geq 0$ such that $m \leq h$ and $m < \ell \leq p$,

$$\frac{1}{T-p}\sum_{t=p+1}^{T} \hat{u}_{t-m}(p)w'_{t-\ell} = O_p\left(\frac{\sup_t \|w_t\|^2}{T-p}\right),$$

where the $O_p(\cdot)$ term is uniform in ℓ and m. By Step 5,

$$\frac{1}{T-p}\sum_{t=p+1}^{T} \hat{u}_{t-m}(p)w'_{t-\ell} = \frac{1}{T-p}\sum_{t=p+1}^{T} \hat{u}_t(p)w'_{t-(\ell-m)} + O_p\left(\frac{\sup_t \|w_t\|^2}{T-p}\right).$$

Since $1 \leq \ell - m \leq p$, the sum on the right-hand side is precisely zero by the orthogonality conditions (B.1).

Step 7. Define for all $m = 0, 1, \ldots, h$ the matrix $\hat{H}_m(p) \equiv \frac{1}{T-p}\sum_{t=p+1}^{T} w_t \hat{u}_{t-m}(p)'$. We will show that

$$\hat{H}_m(p) = \sum_{\ell=1}^{m} \hat{A}_\ell(p)\hat{H}_{m-\ell}(p) + O_p(\hat{R}(p)), \quad m = 1, 2, \ldots, h.$$

Let $m = 1, \ldots, h$ be arbitrary. Since

$$w_t = \hat{c}(p) + \sum_{\ell=1}^{p} \hat{A}_\ell(p)w_{t-\ell} + \hat{u}_t(p),$$

we obtain

$$\begin{aligned}\hat{H}_m(p) &= \frac{1}{T-p}\sum_{t=p+1}^{T} w_t \hat{u}_{t-m}(p)' \\ &= \sum_{\ell=1}^{p} \hat{A}_\ell(p)\frac{1}{T-p}\sum_{t=p+1}^{T} w_{t-\ell}\hat{u}_{t-m}(p)' \\ &\quad + \hat{c}(p)\frac{1}{T-p}\sum_{t=p+1}^{T}\hat{u}_{t-m}(p)' \\ &\quad + \frac{1}{T-p}\sum_{t=p+1}^{T}\hat{u}_t(p)\hat{u}_{t-m}(p)'.\end{aligned}$$

By Step 3, the second term above is $O_p(\frac{1}{T-p}\sup_t \|w_t\|)$. By Step 4, the third term is $O_p((\sum_{\ell=p-h+1}^{p}\|\hat{A}_\ell(p)\|^2)^{1/2})$. As for the first term above, we split it up as follows:

$$\begin{aligned}\sum_{\ell=1}^{p}\hat{A}_\ell(p)\frac{1}{T-p}\sum_{t=p+1}^{T} w_{t-\ell}\hat{u}_{t-m}(p)' &= \sum_{\ell=1}^{m}\hat{A}_\ell(p)\frac{1}{T-p}\sum_{t=p+1}^{T} w_{t-\ell}\hat{u}_{t-m}(p)' \\ &\quad + \sum_{\ell=m+1}^{p}\hat{A}_\ell(p)\frac{1}{T-p}\sum_{t=p+1}^{T} w_{t-\ell}\hat{u}_{t-m}(p)'.\end{aligned}$$

By Steps 1 and 6, the second term above is $O_p(\frac{1}{T-p}\sup_t \|w_t\|^2)$. By Steps 1 and 5, the first term above equals

$$\begin{aligned}\sum_{\ell=1}^{m}\hat{A}_\ell(p)\frac{1}{T-p}\sum_{t=p+1}^{T} w_{t-\ell}\hat{u}_{t-m}(p)' &= \sum_{\ell=1}^{m}\hat{A}_\ell(p)\frac{1}{T-p}\sum_{t=p+1}^{T} w_t\hat{u}_{t-(m-\ell)}(p)' + O_p\left(\frac{\sup_t \|w_t\|^2}{T-p}\right) \\ &= \sum_{\ell=1}^{m}\hat{A}_\ell(p)\hat{H}_{m-\ell}(p) + O_p\left(\frac{\sup_t \|w_t\|^2}{T-p}\right).\end{aligned}$$

STEP 8. We will show that $\hat{H}_m(p) = \hat{C}_m(p)\hat{\Sigma}(p) + O_p(\hat{R}(p))$ for all $m = 0, \dots, h$. We proceed by induction on m. The claim is true by definition for $m = 0$. Assume the claim is true for all $m \leq \tilde{m} - 1$. Then Step 7 implies

$$\begin{aligned}\hat{H}_{\tilde{m}} &= \sum_{\ell=1}^{\tilde{m}}\hat{A}_\ell(p)\hat{H}_{\tilde{m}-\ell}(p) + O_p(\hat{R}(p)) \\ &= \sum_{\ell=1}^{\tilde{m}}\hat{A}_\ell(p)\{\hat{C}_{\tilde{m}-\ell}(p)\hat{\Sigma}(p) + O_p(\hat{R}(p))\} + O_p(\hat{R}(p))\end{aligned}$$

$$= \left(\sum_{\ell=1}^{\tilde{m}} \hat{A}_\ell(p) \hat{C}_{\tilde{m}-\ell}(p) \right) \hat{\Sigma}(p) + O_p(\hat{R}(p))$$
$$= \hat{C}_{\tilde{m}}(p) \hat{\Sigma}(p) + O_p(\hat{R}(p)).$$

Here the penultimate equality uses Step 1, and the last equality uses the recursive definition of $\hat{C}_{\tilde{m}}(p)$.

STEP 9. We will show that $\|\hat{B}(p)^{-1}\| = O_p(1)$. This follows from assumption (ii), the continuity of the Cholesky decomposition at any positive definite matrix, and the assumption (i) that Σ is positive definite.

STEP 10. Let e_x be the $(n_r + 2)$-th n_w-dimensional unit vector, i.e., $x_t = e_x' w_t$. Then

$$e_x' \hat{B}(p)^{-1} \hat{u}_t(p) = \frac{1}{(\frac{1}{T-p} \sum_{t=p+1}^{T} \hat{x}_t(p)^2)^{1/2}} \hat{x}_t(p)$$

for all $t = p+1, p+2, \dots, T$. This is just the sample analogue of the population result (5)–(6), so we refrain from giving the details of the proof.

STEP 11. We will show that

$$\hat{C}_h(p) \hat{B}(p) e_x = \frac{1}{(\frac{1}{T-p} \sum_{t=p+1}^{T} \hat{x}_t(p)^2)^{1/2}} \times \frac{1}{T-p} \sum_{t=p+1}^{T} w_t \hat{x}_{t-h}(p)' + O_p(\hat{R}(p)).$$

By Steps 8 and 9,

$$\hat{C}_h(p) \hat{B}(p) = \hat{C}_h(p) \hat{\Sigma}(p) \hat{B}(p)^{-1\prime} = \hat{H}_m \hat{B}(p)^{-1\prime} + O_p(\hat{R}_p).$$

Hence,

$$\hat{C}_h(p) \hat{B}(p) e_x = \frac{1}{T-p} \sum_{t=p+1}^{T} w_t \left(e_x' \hat{B}(p)^{-1} \hat{u}_{t-h}(p) \right)' + O_p(\hat{R}_p),$$

so the claim follows from Step 10.

STEP 12. The statement of the proposition follows from Step 11 and the fact that $\hat{\theta}_h(p)$ by definition equals the $(n_r + 2)$-th element of $\hat{C}_h(p) \hat{B}(p) e_x$. □

References

Gonçalves, S. & Kilian, L. (2007). Asymptotic and bootstrap inference for AR(∞) processes with conditional heteroskedasticity. *Econometric Reviews*, *26*(6), 609–641.

Lewis, R. & Reinsel, G. C. (1985). Prediction of Multivariate Time Series by Autoregressive Model Fitting. *Journal of Multivariate Analysis*, *16*(3), 393–411.