Online Appendix for "Doubly Robust Local Projections and Some Unpleasant VARithmetic"

José Luis Montiel Olea Mikkel Plagborg-Møller
Eric Qian Christian K. Wolf

August 4, 2024

This online appendix contains supplemental material on: the empirically relevant range of relative LP and VAR standard errors (Supplemental Appendix C), simulation results (Supplemental Appendix D), and proofs (Supplemental Appendix E).

Appendix C Empirically relevant standard error range

We here describe the construction of the empirically relevant range of ratios of VAR and LP standard errors reported in Figure 4.1. We consider four applications in which the researcher has access to a direct measure of a macroeconomic shock, see the descriptions below. We estimate the dynamic causal effects of those four shocks using LPs and the corresponding recursive VAR (Plagborg-Møller and Wolf, 2021), with both specifications including the same set of observables and same number of controls. We then construct bootstrap standard errors with 2,000 bootstrap iterations, assuming homoskedasticity.

The four applications are described below. The data and series mnemonics come from the replication files of Ramey (2016). Our choices of shock measures, observables, samples, data treatment, trends, and lag lengths follow those in Ramey (2016).

1. *Monetary policy:* We use the high-frequency surprises of Gertler and Karadi (2015) (`ff4_tc`) as the observed shock series, and as macro observables we include log industrial production (`lip`), log prices (`lcpi`), the one-year rate (`gs1`), and the excess bond premium (`ebp`). The data are monthly from 1990:1 to 2012:6. We include two lags, and consider impulse responses of all macro variables at horizons of 1–4 years.

2. *Taxes:* The tax shock is the Romer and Romer (2010) series `rrtaxu`. As macro observables we include GDP (`rgdp`), federal tax revenue (`rfedtaxrev`), and government spending (`rgov`), all real, per capita, and in logs. The data are quarterly from 1950:1 to 2007:4. Before estimation, the data are residualized with respect to a quadratic time trend and a dummy variable for 1975:2. We include four lags, and consider impulse responses of all macro variables at horizons of 1–5 years.

3. *Government purchases:* We use the Ramey (2011) military news series (`rameynews`). The macro observables are GDP (`rgdp`), government spending (`rgov`), and the average tax rate (`taxrate`); the first two series are in real terms, logs, and per capita. The data are quarterly from 1947:2 to 2013:4. Before estimation, the data are residualized with respect to a quadratic time trend. We include two lags, and consider impulse responses of all macro variables at horizons of 1–5 years.

4. *Technology:* We use the unanticipated TFP shock series of Francis, Owyang, Roush, and DiCecio (2014) (`ford_tfp`). The macro observables are GDP (`rgdp`), stock prices (`stockp_sh`), and labor productivity (`rgdp/tothours`), all in logs (and in real per capita terms for GDP). The data are quarterly from 1949:2 to 2009:4. Before estimation, the data are residualized with respect to a quadratic time trend. We include two lags, and consider impulse responses of all macro variables at horizons of 1–5 years.

Further details on data construction are provided in the replication codes. Aggregating across shocks, outcome variables, and horizons, we compute 301 ratios of VAR to LP standard errors. The mean ratio is 0.394, the median is 0.367, the 10th percentile is 0.168, and the 90th percentile is 0.638.

Appendix D Further simulation results

D.1 Illustrative univariate model

We begin with a pedagogical illustration based on a univariate ARMA(1,1) model:

$$y_t = \rho y_{t-1} + \varepsilon_t + \psi \varepsilon_{t-1}, \quad \varepsilon_t \overset{i.i.d.}{\sim} N(0, \sigma^2). \tag{D.1}$$

Throughout this section we set $\rho = 0.9$, $\sigma^2 = 1$, and $\psi \in \{0, 0.25, 0.5, 0.75\}$. The lagged MA term thus accounts for up to 36 per cent of the overall variance of the error term in (D.1).

p	M			$\frac{M^2}{1+M^2}$		
	$\psi = 0.25$	$\psi = 0.50$	$\psi = 0.75$	$\psi = 0.25$	$\psi = 0.50$	$\psi = 0.75$
1	3.622	7.396	11.234	0.929	0.982	0.992
2	0.882	3.337	6.821	0.437	0.918	0.979
3	0.220	1.631	4.682	0.046	0.727	0.956
4	0.055	0.811	3.361	0.003	0.397	0.919

Table D.1: M and $\frac{M^2}{1+M^2}$ as a function of p and ψ in the ARMA(1,1) model (D.1), with the researcher estimating an AR(p).

We first quantify the amount of misspecification, after representing the model as the best-fitting ARMA(p, ∞) model (2.1) for a given number of VAR lags p, as we did in Section 5 (see Footnote 8 and Table 5.1). We see that, even for moderate p and ψ, the degree of misspecification—and accordingly the optimal model averaging weight on LP—can be large.

Next we show that the implied misspecification polynomial $\alpha(L)$ can in fact be very close to the least favorable polynomial $\alpha^\dagger(L; h)$ derived in Section 4, which represents the theoretical worst case for AR bias and coverage. Figure D.1 shows $\alpha(L)$ (solid) as well as $\alpha^\dagger(L; h)$ (dashed and dotted) for horizons $h \in \{1, 5, 10\}$, and throughout setting $\psi = 0.5$. For both $p = 1$ as well as $p = 4$, the actual MA polynomial $\alpha(L)$ implied by the ARMA(1,1) model (D.1) is very close—though not quite identical—to the worst-case $\alpha^\dagger(L; 1)$ at horizon $h = 1$. Hence, at least in the particular DGP considered here, the least favorable lag polynomial is not some practically immaterial theoretical curiosity.

Consistent with our theory, we find that coverage can be poor for VAR confidence intervals, while LP intervals are robust to the presence of the MA term. Figure D.2 reports coverage rates and median confidence interval lengths for the cases $\psi = 0$ (no misspecification) and $\psi = 0.25$ (moderate misspecification, with the lagged MA term accounting for around 6% of the variance of the error term). We throughout set $p = 1$ and simulate 5,000 samples of size $T = 240$. The top panel reveals that, when the AR(1) model is in fact correctly specified (i.e., for $\psi = 0$), then both LP and AR confidence intervals attain the nominal coverage probability of 90 per cent (left panel); furthermore, and also as expected, the AR confidence intervals are meaningfully shorter (right panel). In the misspecified case in the bottom panel, the AR confidence intervals instead substantially undercover, and particularly so at short horizons. LP, on the other hand, exhibits at worst mild undercoverage.

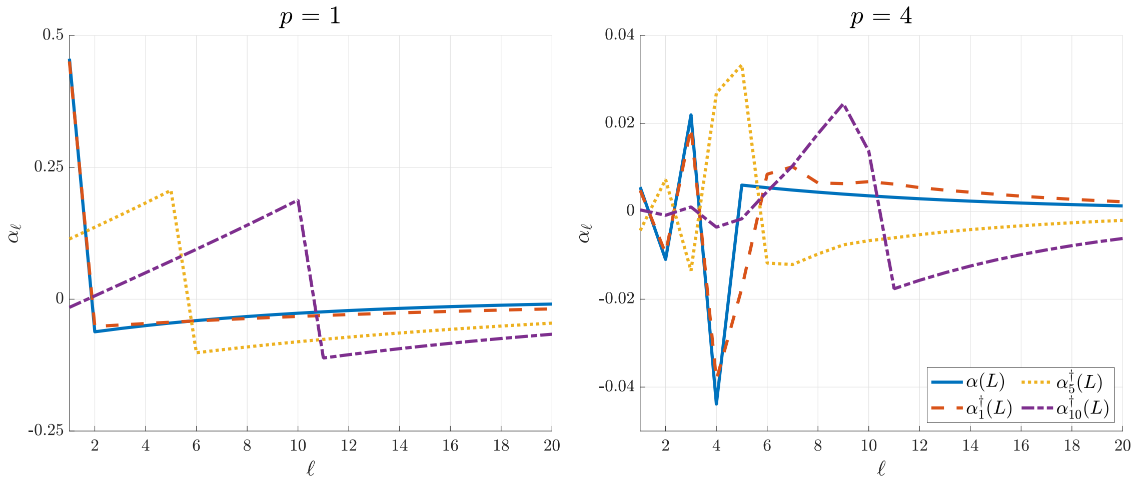

Figure D.1: $\alpha(L)$ (black, solid) and least favorable $\alpha^\dagger(L)$ (colored, dashed and dotted) at various horizons for the simple ARMA(1,1) model (D.1), with the researcher estimating an AR(1) (left) and an AR(4) (right).

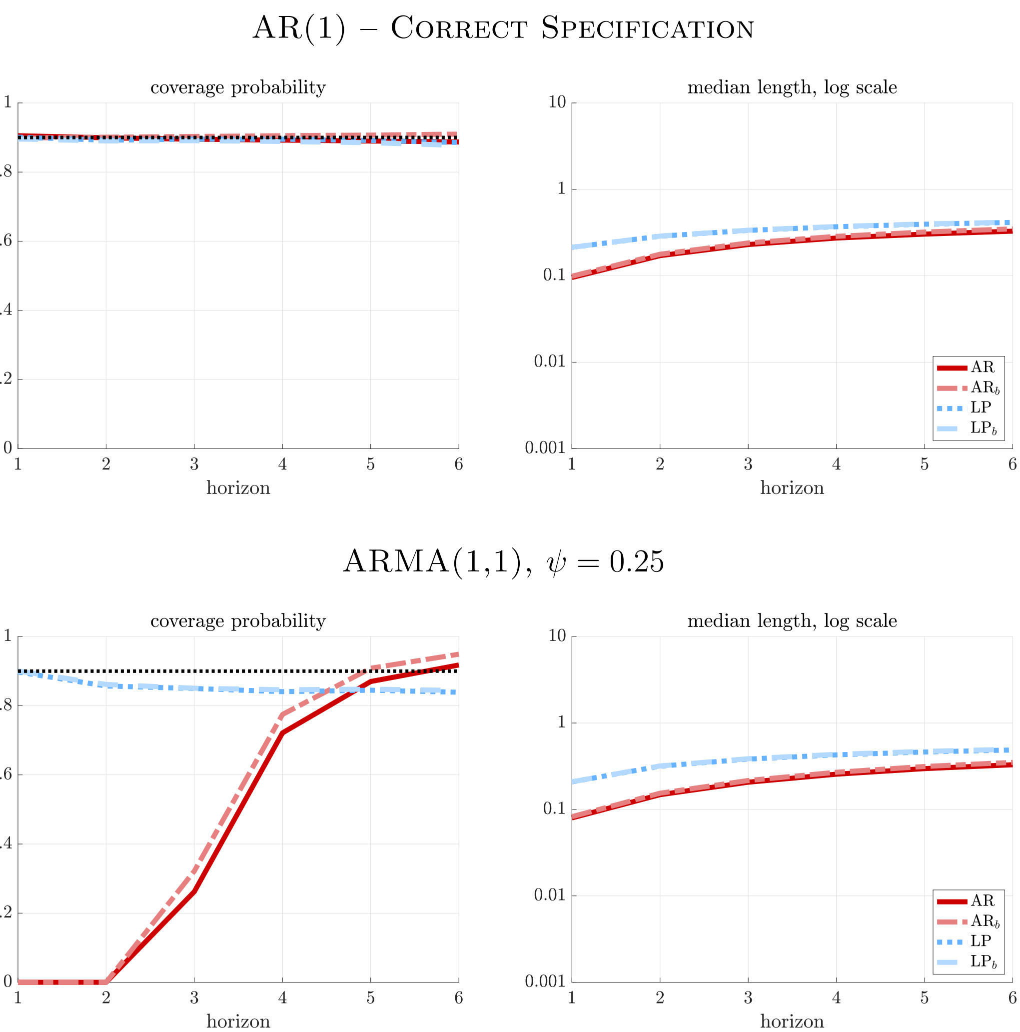

Figure D.2: Coverage probabilities (left) and median confidence interval length (right) for AR (red) and LP (blue) confidence intervals computed via the delta method or bootstrap (the latter are indicated with subscript "b" in the figure legends). The DGP is (D.1) with $\psi = 0$ (top) and $\psi = 0.25$ (bottom), with estimation lag length fixed at $p = 1$.

D.2 Smets and Wouters model

We now provide additional simulation results based on the Smets and Wouters (2007) DGP, as in Section 5. The sample size and inference methods are the same as in Section 5, unless otherwise noted.

Monetary shock. Here we assume that the econometrician observes two variables: the monetary policy shock and total output. The impulse response function of interest is that of output with respect to the monetary shock. We note that, since the monetary shock follows a simple AR(1) process, this setting is likely to be less challenging for VAR inference than the cost-push shock experiment in Section 5.

Figure D.3 shows that, just as in our main exercise, VAR confidence intervals can severely undercover, while LP intervals remain robust. The top and middle panels show results for lag length p selected by AIC or fixed at 4, respectively, and the results are comparable to those in Section 5. Finally, in the bottom panel, we show what happens if we slightly perturb the DGP by replacing the actual lag polynomial $\alpha(L)$ with the least favorable one at horizon 4, $\alpha^{\dagger}(L; 4)$, without changing the amount $\|\alpha(L)\|$ of overall misspecification (see Section 4).[D.1] VAR undercoverage is now severe even at shorter horizons. Overall, however, the magnitudes of undercoverage at medium and long horizons are broadly comparable with those obtained under the actual $\alpha(L)$ implied by the Smets and Wouters (2007) model, confirming that the least favorable MA polynomial $\alpha^{\dagger}(L)$ is not particularly pathological in general.

Cost-push shock. To complement our simulation evidence in Section 5, we repeat the cost-push shock exercise of that section for a larger sample size of $T = 2{,}000$. We fix $p = 2$, in line with the median AIC lag length selection in our main exercise. The results shown in Figure D.4 are similar both qualitatively and quantitatively to our main findings in the top panel of Figure 5.1, especially for the bootstrap confidence intervals. Hence, our results with $T = 240$ in Section 5 are not driven by small-sample phenomena. Figure D.4 also plots the theoretically predicted VAR coverage probability (orange dashed line) from Corollary 3.1. We see that this asymptotic coverage is very close to the actual one.

[D.1] To be precise, we first set $p = 1$, derive the VARMA$(1,\infty)$ as discussed in Footnote 8, and then switch out the implied lag polynomial $\alpha(L)$. The estimation lag length is selected by AIC.

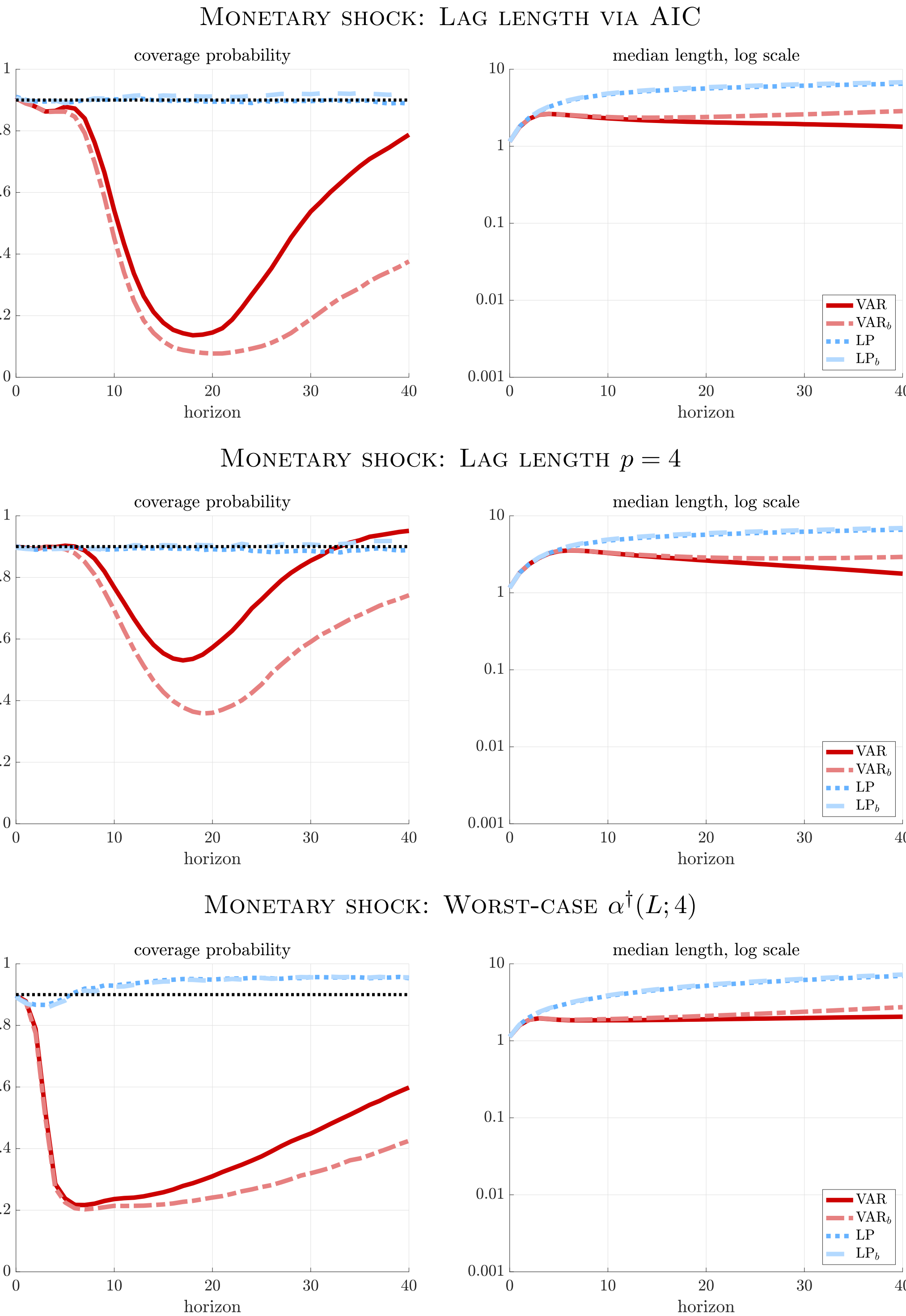

Figure D.3: Coverage probabilities (left) and median confidence interval length (right) for VAR (red) and LP (blue) confidence intervals computed via the delta method or bootstrap (the latter are indicated with subscript "b" in the figure legends).

Cost-push shock: Lag length $p = 2$, larger sample

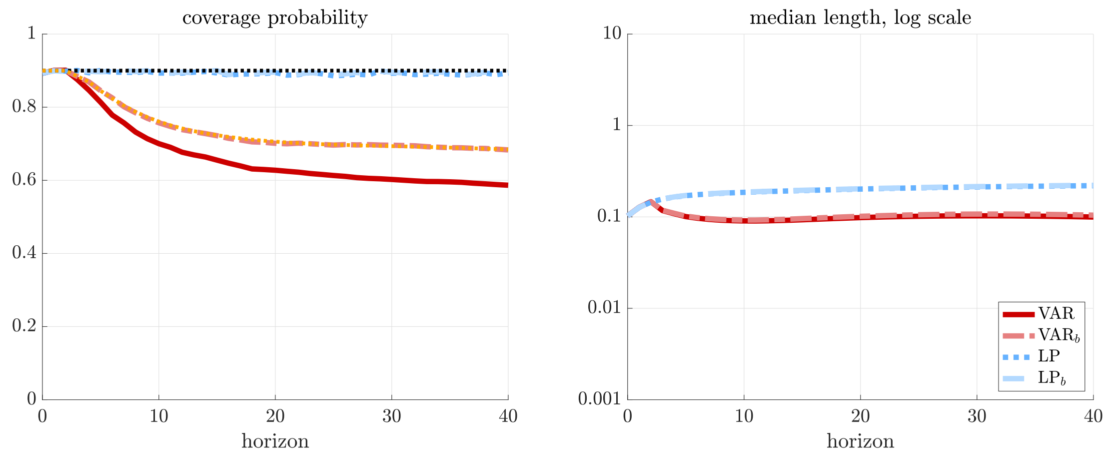

Figure D.4: See caption for Figure D.3. The orange dashed line indicates the asymptotic VAR coverage probability predicted by Corollary 3.1.

Appendix E Further proofs

We impose Assumption 2.1 throughout. Let $\|B\|$ denote the Frobenius norm of any matrix B. It is well known that this norm is sub-multiplicative: $\|BC\| \leq \|B\| \cdot \|C\|$. Let I_n denote the $n \times n$ identity matrix, $0_{m\times n}$ the $m \times n$ matrix of zeros, and $e_{i,n}$ the n-dimensional unit vector with a 1 in the i-th position. Recall from Assumption 2.1 the definitions $D \equiv \text{Var}(\varepsilon_t) = \text{diag}(\sigma_1^2, \dots, \sigma_m^2)$, $\tilde{y}_t \equiv (I_n - AL)^{-1} H\varepsilon_t = \sum_{s=0}^{\infty} A^s H \varepsilon_{t-s}$, and $S \equiv \text{Var}(\tilde{y}_t)$.

E.1 Main lemmas

Lemma E.1. *For any $i^* \in \{1, \dots, n\}$ and $j^* \in \{1, \dots, m\}$, we have*

$$y_{i^*,t+h} = \theta_{h,T}\varepsilon_{j^*,t} + \underline{B}'_{h,i^*,j^*}\underline{y}_{j^*,t} + B'_{h,i^*,j^*} y_{t-1} + \xi_{h,i^*,t} + T^{-\zeta}\Theta_h(L)\varepsilon_t,$$

where

$$\begin{aligned}
\theta_{h,T} &\equiv e'_{i^*,n}(A^h H + T^{-\zeta}\sum_{\ell=1}^{h} A^{h-\ell} H \alpha_\ell) e_{j^*,m}, \\
\underline{B}'_{h,i^*,j^*} &\equiv e'_{i^*,n} A^h \underline{H}_{j^*} H_{11}^{-1}, \\
B'_{h,i^*,j^*} &\equiv e'_{i^*,n}\left[A^{h+1} - A^h \underline{H}_{j^*} H_{11}^{-1} \underline{I}_{j^*} A\right],
\end{aligned}$$

$$\xi_{h,i^*,t} \equiv e'_{i^*,n} A^h \overline{H}_{j^*} \overline{\varepsilon}_{j^*,t} + \sum_{\ell=1}^{h} e'_{i^*,n} A^{h-\ell} H \varepsilon_{t+\ell},$$

and $\Theta_h(L) = \sum_{\ell=-\infty}^{\infty} \Theta_{h,\ell} L^\ell$ *is an absolutely summable,* $1 \times n$ *two-sided lag polynomial with the* j^**-th element of* $\Theta_{h,0}$ *equal to zero. Moreover,*

$$T^{-1} \sum_{t=1}^{T-h} (\Theta_h(L)\varepsilon_t)\varepsilon_{j^*,t} = O_p(T^{-1/2}).$$

Proof. Iteration on the model in Equation (2.1) yields

$$y_{t+h} = A^{h+1} y_{t-1} + \sum_{\ell=0}^{h} A^{h-\ell}(H\varepsilon_{t+\ell} + T^{-\zeta} H \alpha(L)\varepsilon_{t+\ell}). \tag{E.1}$$

As in Section 2.2, let $\underline{y}_{j^*,t} \equiv (y_{1,t}, \ldots, y_{j^*-1,t})'$ denote the variables ordered before $y_{j^*,t}$ (if any). Analogously, let $\overline{y}_{j^*,t} \equiv (y_{j^*+1,t}, \ldots y_{n,t})'$ denote the variables ordered after $y_{j^*,t}$.

Using Assumption 2.1(iii), partition

$$H = (\underline{H}_{j^*}, H_{\bullet,j^*}, \overline{H}_{j^*}) = \begin{pmatrix} H_{11} & 0 & 0 \\ H_{21} & H_{22} & 0 \\ H_{31} & H_{32} & H_{33} \end{pmatrix}$$

conformably with the vector $y_t = (\underline{y}'_{j^*,t}, y_{j^*,t}, \overline{y}'_{j^*,t})'$. Let $\underline{I}_{j^*}$ denote the first $j^* - 1$ rows of the $n \times n$ identity matrix. Using the definition of y_t in Equation (2.1),

$$\underline{y}_{j^*,t} = \underline{I}_{j^*} A y_{t-1} + H_{11} \underline{\varepsilon}_{j^*,t} + T^{-\zeta} H_{11} \underline{I}_{j^*} \alpha(L) \varepsilon_t,$$

where $\underline{\varepsilon}_{j^*,t} = \underline{I}_{j^*} \varepsilon_t$. Using the previous equation to solve for $\underline{\varepsilon}_{j^*,t}$ we get

$$\underline{\varepsilon}_{j^*,t} = H_{11}^{-1}(\underline{y}_{j^*,t} - \underline{I}_{j^*} A y_{t-1} - T^{-\zeta} H_{11} \underline{I}_{j^*} \alpha(L)\varepsilon_t). \tag{E.2}$$

Expanding the terms in (E.1) we get:

$$\begin{aligned} y_{t+h} &= A^{h+1} y_{t-1} + A^h H \varepsilon_t + T^{-\zeta} A^h H \alpha(L) \varepsilon_t + \sum_{\ell=1}^{h} A^{h-\ell}(H\varepsilon_{t+\ell} + T^{-\zeta} H \alpha(L) \varepsilon_{t+\ell}) \\ &= A^{h+1} y_{t-1} + \left(A^h \underline{H}_{j^*} \underline{\varepsilon}_{j^*,t} + A^h H_{\bullet,j^*} \varepsilon_{j^*,t} + A^h \overline{H}_{j^*} \overline{\varepsilon}_{j^*,t}\right) \\ &\quad + T^{-\zeta} A^h H \alpha(L) \varepsilon_t + \sum_{\ell=1}^{h} A^{h-\ell}(H\varepsilon_{t+\ell} + T^{-\zeta} H \alpha(L)\varepsilon_{t+\ell}) \end{aligned}$$

$$= A^{h+1}y_{t-1} + A^h \underline{H}_{j^*} H_{11}^{-1}(\underline{y}_{j^*,t} - \underline{I}_{j^*} A y_{t-1} - T^{-\zeta} H_{11} \underline{I}_{j^*} \alpha(L)\varepsilon_t) + A^h H_{\bullet,j^*}\varepsilon_{j^*,t} + A^h \overline{H}_{j^*}\overline{\varepsilon}_{j^*,t}$$
$$+ T^{-\zeta} A^h H \alpha(L)\varepsilon_t + \sum_{\ell=1}^{h} A^{h-\ell}(H\varepsilon_{t+\ell} + T^{-\zeta} H\alpha(L)\varepsilon_{t+\ell}),$$

where the last equality follows from substituting (E.2). Re-arranging terms we get

$$y_{i^*,t+h} = \left(e'_{i^*,n} A^h H_{\bullet,j^*}\right)\varepsilon_{j^*,t} + \underbrace{\left(e'_{i^*,n} A^h \underline{H}_{j^*} H_{11}^{-1}\right)}_{\equiv \underline{B}'_{h,i^*,j^*}} \underline{y}_{j^*,t} + \underbrace{\left(e'_{i^*,n}\left[A^{h+1} - A^h \underline{H}_{j^*} H_{11}^{-1} \underline{I}_{j^*} A\right]\right)}_{\equiv B'_{h,i^*,j^*}} y_{t-1}$$
$$+ \underbrace{e'_{i^*,n}\left(A^h \overline{H}_{j^*}\overline{\varepsilon}_{j^*,t} + \sum_{\ell=1}^{h} A^{h-\ell} H\varepsilon_{t+\ell}\right)}_{=\xi_{h,i^*,t}}$$
$$+ T^{-\zeta} e'_{i^*,n}\left(-A^h \underline{H}_{j^*} H_{11}^{-1} H_{11} \underline{I}_{j^*}\alpha(L)\varepsilon_t + \sum_{\ell=0}^{h} A^{h-\ell} H\alpha(L)\varepsilon_{t+\ell}\right), \tag{E.3}$$

Using the definition of $\theta_{h,T} \equiv e'_{i^*,n}(A^h H + T^{-\zeta}\sum_{\ell=1}^{h} A^{h-\ell} H\alpha_\ell)e_{j^*,m}$ and adding and subtracting $e'_{i^*,n}\left(T^{-\zeta}\sum_{\ell=1}^{h} A^{h-\ell} H\alpha_\ell\right) e_{j^*,m}\varepsilon_{j^*,t}$ in (E.3), we obtain a representation of the form

$$y_{i^*,t+h} = \theta_{h,T}\varepsilon_{j^*,t} + \underline{B}'_{h,i^*,j^*}\underline{y}_{j^*,t} + B'_{h,i^*,j^*} y_{t-1} + \xi_{h,i^*,t} + T^{-\zeta}\tilde{u}_t, \tag{E.4}$$

where

$$\tilde{u}_t \equiv e'_{i^*,n}\left(-A^h \underline{H}_{j^*}\underline{I}_{j^*}\alpha(L)\varepsilon_t + \sum_{\ell=0}^{h} A^{h-\ell} H\alpha(L)\varepsilon_{t+\ell} - \left(\sum_{\ell=1}^{h} A^{h-\ell} H\alpha_\ell e_{j^*,m} e'_{j^*,m}\right)\varepsilon_t\right). \tag{E.5}$$

Algebra shows that $\tilde{u}_t$ can be written as a two-sided lag polynomial, $\Theta_h(L) = \sum_{\ell=-\infty}^{\infty}\Theta_{h,\ell}L^\ell$, with coefficients of dimension $1 \times n$ given by the following formulae:

1. For $\ell \geq 1$:
$$\Theta_{h,\ell} = -e'_{i^*,n} A^h \underline{H}_{j^*}\underline{I}_{j^*}\alpha_\ell + \sum_{s=0}^{h} e'_{i^*,n} A^{h-s} H\alpha_{\ell+s}.$$

2. For $\ell = 0$:
$$\Theta_{h,0} = \sum_{s=1}^{h} e'_{i^*,n} A^{h-s} H\alpha_s - \sum_{s=1}^{h} e'_{i^*,n} A^{h-s} H\alpha_s e_{j^*,m} e'_{j^*,m},$$
and, consequently, $\Theta_{h,0,j^*} \equiv \Theta_{h,0} e_{j^*,m} = 0$.

3. For $\ell \in \{-(h-1), \ldots, -1\}$:
$$\sum_{s=1}^{h+\ell} e'_{i^*,n} A^{h-s+\ell} H \alpha_s.$$

4. For $\ell \leq -h$, $\Theta_{h,\ell} = 0_{1\times n}$.

We next show that $\Theta_h(L)$ is absolutely summable, that is $\sum_{\ell=-\infty}^{\infty} \|\Theta_{h,l}\| < \infty$. To do this, it suffices to show that $\sum_{\ell=1}^{\infty} \|\Theta_{h,l}\| < \infty$, since all the coefficients with index $\ell \leq -h$ are 0. Note that, by definition, for any $\ell \geq 1$:

$$\|\Theta_{h,\ell}\| \leq \|A^h\| \|\underline{H}_{j^*} \underline{I}_{j^*}\| \|\alpha_\ell\| + \sum_{s=0}^{h} \|A^{h-s}\| \|H\| \|\alpha_{\ell+s}\|.$$

Thus,
$$\sum_{\ell=1}^{\infty} \|\Theta_{h,\ell}\| \leq \|A^h\| \|\underline{H}_{j^*} \underline{I}_{j^*}\| \sum_{\ell=1}^{\infty} \|\alpha_\ell\| + \|H\| \sum_{\ell=1}^{\infty} \sum_{s=0}^{h} \|A^{h-s}\| \|\alpha_{\ell+s}\|.$$

Let $\lambda \in [0,1)$ and $C > 0$ be chosen such that $\|A^\ell\| \leq C\lambda^\ell$ for all $\ell \geq 0$ (such constants exists by Assumption 2.1(ii)). Then

$$\begin{aligned}
\sum_{\ell=1}^{\infty} \sum_{s=0}^{h} \|A^{h-s}\| \|\alpha_{\ell+s}\| &\leq C \sum_{\ell=1}^{\infty} \sum_{s=1}^{h} \lambda^{h-s} \|\alpha_{\ell+s}\| \\
&\leq C \sum_{\ell=1}^{\infty} \sum_{s=1}^{h} \|\alpha_{\ell+s}\| \\
&\leq Ch \sum_{\ell=1}^{\infty} \|\alpha_\ell\| \\
&< \infty,
\end{aligned}$$

where the last inequality holds because the coefficients of $\alpha(L)$ are summable. We thus conclude that

$$y_{i^*,t+h} = \theta_{h,T} \varepsilon_{j^*,t} + \underline{B}'_{h,y} \underline{y}_{j^*,t} + B'_{h,y} y_{t-1} + \xi_{h,i^*,t} + T^{-\zeta} \Theta_h(L) \varepsilon_t,$$

where $\Theta_h(L)$ is a two-sided lag-polynomial with summable coefficients.

Finally, we show that

$$T^{-1} \sum_{t=1}^{T-h} (\Theta_h(L)\varepsilon_t)\varepsilon_{j^*,t} = O_p(T^{-1/2}). \tag{E.6}$$

To do this, we write

$$\Theta_h(L)\varepsilon_t = \sum_{\ell=1}^{\infty} \Theta_{h,\ell}\varepsilon_{t-\ell} + \Theta_{h,0}\varepsilon_t + \sum_{\ell=-(h-1)}^{-1} \Theta_{h,\ell}\varepsilon_{t-\ell}.$$

1. Note first that the process

$$\left\{ \left(\sum_{\ell=1}^{\infty} \Theta_{h,\ell}\varepsilon_{t-\ell} \right) \varepsilon_{j^*,t} \right\}_{t=1}^{\infty}$$

is white noise (mean-zero and serially uncorrelated components). The summability of coefficients of $\Theta_h(L)$ further implies that

$$\text{Var}\left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} \left(\sum_{\ell=1}^{\infty} \Theta_{h,\ell}\varepsilon_{t-\ell} \right) \varepsilon_{j^*,t} \right) = \frac{T-h}{T} \text{Var}\left(\left(\sum_{\ell=1}^{\infty} \Theta_{h,\ell}\varepsilon_{t-\ell} \right) \varepsilon_{j^*,t} \right) < \infty.$$

Thus, by Markov's inequality, we have that

$$\frac{1}{T} \sum_{t=1}^{T-h} \left(\sum_{\ell=1}^{\infty} \Theta_{h,\ell}\varepsilon_{t-\ell} \right) \varepsilon_{j^*,t} = O_p(T^{-1/2}).$$

2. Note second that the process

$$\{(\Theta_{h,0}\varepsilon_t)\, \varepsilon_{j^*,t}\}_{t=1}^{\infty}$$

is i.i.d. with mean zero (since ε_t has independent components and $\Theta_{0,\ell,j^*} = 0$). Since the process has finite variance, we conclude that

$$\frac{1}{T} \sum_{t=1}^{T-h} (\Theta_{h,0}\varepsilon_t)\, \varepsilon_{j^*,t} = O_p(T^{-1/2}).$$

3. Finally, note that the process

$$\left\{ \left(\sum_{\ell=-(h-1)}^{-1} \Theta_{h,\ell}\varepsilon_{t-\ell} \right) \varepsilon_{j^*,t} \right\}_{t=1}^{\infty}$$

is white noise (mean-zero and serially uncorrelated components). Therefore,

$$\text{Var}\left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} \left(\sum_{\ell=-(h-1)}^{-1} \Theta_{h,\ell}\varepsilon_{t-\ell} \right) \varepsilon_{j^*,t} \right) = \frac{T-h}{T} \text{Var}\left(\left(\sum_{\ell=-(h-1)}^{-1} \Theta_{h,\ell}\varepsilon_{t-\ell} \right) \varepsilon_{j^*,t} \right) < \infty.$$

We conclude that

$$\frac{1}{T}\sum_{t=1}^{T-h}\left(\sum_{\ell=-(h-1)}^{-1}\Theta_{h,\ell}\varepsilon_{t-\ell}\right)\varepsilon_{j^*,t}=O_p(T^{-1/2}).$$

This completes the verification of (E.6). □

Lemma E.2.

$$\hat{A}-A=T^{-\zeta}H\sum_{\ell=1}^{\infty}\alpha_\ell DH'(A')^{\ell-1}S^{-1}+T^{-1}\sum_{t=1}^{T}H\varepsilon_t\tilde{y}'_{t-1}S^{-1}+o_p(T^{-\zeta}).$$

In particular, $\hat{A}-A=O_p(T^{-\zeta}+T^{-1/2})$.

Proof. Since,

$$\hat{A}-A=\left(T^{-1}\sum_{t=1}^{T-h}u_t y'_{t-1}\right)\left(T^{-1}\sum_{t=1}^{T-h}y_{t-1}y'_{t-1}\right)^{-1},$$

the result follows from Lemmas E.8 and E.9. □

Lemma E.3.

$$\hat{\nu}-H_{\bullet,j^*}=\frac{1}{\sigma^2_{j^*}}T^{-1}\sum_{t=1}^{T}\xi_{0,t}\varepsilon_{j^*,t}+o_p(T^{-1/2}).$$

Proof. By Lemma E.6, $\hat{\nu}=(0_{1\times(j^*-1)},1,\hat{\bar{\nu}}')$, where the j-th element of $\hat{\bar{\nu}}$ equals the on-impact local projection of $y_{i^*+j,t}$ on $y_{j^*,t}$, controlling for $\underline{y}_{j^*,t}$ and y_{t-1}. The statement of the lemma is therefore a direct consequence of Proposition 3.1 and the fact that (by definition) $\xi_{0,i,t}=0$ for $i\leq j^*$. □

Lemma E.4. *Fix* $h\geq 0$. *Consider the regression of* $y_{j^*,t}$ *on* $q_{j^*,t}\equiv(\underline{y}'_{j^*,t},y'_{t-1})'$, *using the observations* $t=1,2,\dots,T-h$:

$$y_{j^*,t}=\hat{\vartheta}'_h q_{j^*,t}+\hat{x}_{h,t}.$$

Note that the residuals $\hat{x}_{h,t}$ *are consistent with the earlier definition in the proof of Proposition 3.1. Let* $\underline{\lambda}'_{j^*}$ *be the row vector containing the first* j^*-1 *elements of the last row of* $-\tilde{H}^{-1}$ *(where* $\tilde{H}$ *is defined in Assumption 2.1(iii)). Let* $\lambda'_{j^*}\equiv(-\underline{\lambda}'_{j^*},1,0_{1\times(n-j^*)})$ *and* $\vartheta\equiv(\underline{\lambda}'_{j^*},(\lambda'_{j^*}A))'$. *Then:*

i) $\hat{\vartheta}_h-\vartheta=O_p(T^{-\zeta}+T^{-1/2})$.

ii) $T^{-1}\sum_{t=1}^{T-h}(\hat{x}_{h,t}-\varepsilon_{j^*,t})\varepsilon_{j^*,t}=o_p(T^{-1/2})$.

iii) *For* $\ell\geq 1$, $T^{-1}\sum_{t=1}^{T-h}(\hat{x}_{h,t}-\varepsilon_{j^*,t})\varepsilon_{t+\ell}=o_p(T^{-1/2})$.

iv) $T^{-1}\sum_{t=1}^{T-h}(\hat{x}_{h,t}-\varepsilon_{j^*,t})\hat{x}_{h,t}=o_p(T^{-1/2})$.

v) $T^{-1}\sum_{t=1}^{T-h}\hat{x}_{h,t}^2\xrightarrow{p}\sigma_{j^*}^2$.

vi) *For any absolutely summable two-sided lag polynomial* $B(L)$, $T^{-1}\sum_{t=1}^{T-h}(\hat{x}_{h,t}-\varepsilon_{j^*,t})B(L)\varepsilon_t=O_p(T^{-\zeta}+T^{-1/2})$.

Proof. By Equation (2.1), the outcome variables in the model satisfy

$$y_t=Ay_{t-1}+H[I_m+T^{-\zeta}\alpha(L)]\varepsilon_t,\ t=1,2,\ldots,T.$$

By Assumption 2.1(iii), the first j^* rows of the matrix H above are of the form $(\tilde{H},0_{j^*\times(j^*-m)})$, where m is the number of shocks and $\tilde{H}$ is a $j^*\times j^*$ lower triangular matrix with 1's on the diagonal.

$\tilde{H}$ is invertible, which means we can premultiply the first j^* equations of (2.1) by $\tilde{H}^{-1}$ to obtain:

$$[\tilde{H}^{-1},0_{j^*\times(n-j^*)}]y_t=[\tilde{H}^{-1},0_{j^*\times(n-j^*)}]Ay_{t-1}+[I_{j^*},0_{j^*\times(m-j^*)}][I_m+T^{-\zeta}\alpha(L)]\varepsilon_t.$$

By definition, $-\underline{\lambda}_{j^*}'$ is the row vector containing the first j^*-1 elements of the last row of $\tilde{H}^{-1}$ and $\lambda_{j^*}'\equiv(-\underline{\lambda}_{j^*}',1,0_{1\times(n-j^*)})$. Thus, we can re-write the j^*-th equation above as

$$[-\underline{\lambda}_{j^*}',1,0_{j^*\times(n-j^*)}]y_t=\lambda_{j^*}'Ay_{t-1}+\varepsilon_{j^*,t}+T^{-\zeta}\alpha_{j^*}(L)\varepsilon_t,$$

where $\alpha_{j^*}(L)$ is the j^*-th row of $\alpha(L)$. Re-arranging terms we get

$$y_{j^*,t}=\vartheta' q_{j^*,t}+\varepsilon_{j^*,t}+T^{-\zeta}\alpha_{j^*}(L)\varepsilon_t,$$

where $\vartheta\equiv(\underline{\lambda}_{j^*}',(\lambda_{j^*}'A))'$ and $q_{j^*,t}\equiv(\underline{y}_{j^*,t}',y_{t-1}')'$. In a slight abuse of notation, and for notational simplicity, we henceforth replace $q_{j^*,t}$ by q_t.

Statement (i) follows from standard OLS algebra if we can show that a) $T^{-1}\sum_{t=1}^{T-h}q_t\varepsilon_{j^*,t}=O_p(T^{-\zeta}+T^{-1/2})$, b) $(T^{-1}\sum_{t=1}^{T-h}q_tq_t')^{-1}=O_p(1)$, and c) $T^{-\zeta-1}\sum_{t=1}^{T-h}q_t(\alpha_{j^*}(L)\varepsilon_t)=O_p(T^{-\zeta})$. Lemma E.10 establishes these results.

The proof of statements (ii) and (iii) are similar, so we focus on the latter. By definition of $\hat{x}_{h,t}$, we have $\hat{x}_{h,t} - \varepsilon_{j^*,t} = (\vartheta - \hat{\vartheta}_h)' q_t + T^{-\zeta} \alpha_{j^*}(L)\varepsilon_t$. As in Lemma E.7 define $\tilde{y}_t = \sum_{s=0}^{\infty} A^s H \varepsilon_{t-s}$. Let $\tilde{q}_t \equiv (\underline{\tilde{y}}'_{j^*,t}, \tilde{y}'_{t-1})'$ and $\Delta_t \equiv q_t - \tilde{q}_t$. Thus,

$$T^{-1} \sum_{t=1}^{T-h} (\hat{x}_{h,t} - \varepsilon_{j^*,t})\varepsilon_{t+\ell} = (\vartheta - \hat{\vartheta}_h)' \left(\frac{1}{T} \sum_{t=1}^{T-h} \Delta_t \varepsilon_{t+\ell} \right) \tag{E.7}$$

$$+ (\vartheta - \hat{\vartheta}_h)' \left(\frac{1}{T} \sum_{t=1}^{T-h} \tilde{q}_t \varepsilon_{t+\ell} \right) \tag{E.8}$$

$$+ \frac{1}{T^\zeta} \left(\frac{1}{T} \sum_{t=1}^{T-h} \left(\alpha_{j^*}(L)\varepsilon_t\right) \varepsilon_{t+\ell} \right). \tag{E.9}$$

By Lemma E.10, $(\vartheta - \hat{\vartheta}_h) = O_p(T^{-\zeta} + T^{-1/2})$. Direct second-moment calculations can be used to show that the terms in (E.8)–(E.9) are of order

$$O_p(T^{-\zeta} + T^{-1/2}) O_p(T^{-1/2}) \text{ and } O_p(T^{-\zeta-1/2}),$$

respectively. This implies that both terms are $o_p(T^{-1/2})$. Finally, note that Lemma E.7 and Assumption 2.1(i) imply that the sum in (E.7) is $O_p\left(T^{-\zeta}\right)$. Thus, (E.7) is of order

$$O_p(T^{-\zeta} + T^{-1/2}) O_p(T^{-\zeta}) = o_p\left(T^{-1/2}\right),$$

using $\zeta > 1/4$. Since we have shown that (E.7)–(E.9) are $o_p(T^{-1/2})$, then for $\ell \geq 1$,

$$T^{-1} \sum_{t=1}^{T-h} (\hat{x}_{h,t} - \varepsilon_{j^*,t})\varepsilon_{t+\ell} = o_p(T^{-1/2}).$$

For statement (iv), note that

$$T^{-1} \sum_{t=1}^{T-h} (\hat{x}_{h,t} - \varepsilon_{j^*,t})\hat{x}_{h,t} = T^{-1} \sum_{t=1}^{T-h} (\hat{x}_{h,t} - \varepsilon_{j^*,t})^2 + T^{-1} \sum_{t=1}^{T-h} (\hat{x}_{h,t} - \varepsilon_{j^*,t})\varepsilon_{j^*,t}.$$

Lemma E.11 shows that $T^{-1} \sum_{t=1}^{T-h} (\hat{x}_{h,t} - \varepsilon_{j^*,t})^2 = o_p(T^{-1/2})$. This result, combined with (ii), implies that statement (iv) holds.

For statement (v), note that

$$T^{-1} \sum_{t=1}^{T-h} (\hat{x}_{h,t})^2 = T^{-1} \sum_{t=1}^{T-h} (\hat{x}_{h,t} - \varepsilon_{j^*,t} + \varepsilon_{j^*,t})^2$$

$$= T^{-1} \sum_{t=1}^{T-h} (\hat{x}_{h,t} - \varepsilon_{j^*,t})^2 - 2T^{-1} \sum_{t=1}^{T-h} (\hat{x}_{h,t} - \varepsilon_{j^*,t})\varepsilon_{j^*,t} + T^{-1} \sum_{t=1}^{T-h} \varepsilon_{j^*,t}^2.$$

Lemma E.11 and statement (ii) imply that the first two terms converge in probability to zero. Since $T^{-1} \sum_{t=1}^{T-h} \varepsilon_{j^*,t}^2 \xrightarrow{p} \sigma_{j^*}^2$ (by the Law of Large Numbers), statement (v) holds.

Finally, statement (vi) obtains by decomposing

$$\begin{aligned} T^{-1} \sum_{t=1}^{T-h} B(L)\varepsilon_t(\hat{x}_{h,t} - \varepsilon_{j^*,t}) &= T^{-1} \sum_{t=1}^{T-h} B(L)\varepsilon_t q_t'(\vartheta - \hat{\vartheta}_h) + T^{-\zeta} T^{-1} \sum_{t=1}^{T-h} B(L)\varepsilon_t [\alpha_{j^*}(L)\varepsilon_t]' \\ &= O_p(1) \times O_p(T^{-\zeta} + T^{-1/2}) + T^{-\zeta} \times O_p(1), \end{aligned}$$

where the last line follows from statement (i), Lemma E.7, and moment calculations. □

Lemma E.5. *Let $\mathcal{X}$ denote the set of sequences $\{x_\ell\}_{\ell=1}^\infty$ of $m \times m$ matrices x_ℓ satisfying $\sum_{\ell=1}^\infty \|x_\ell\|^2 \leq 1$. Let $\{L_\ell\}_{\ell=1}^\infty$ be a sequence of $r \times m^2$ matrices L_ℓ satisfying $\sum_{\ell=1}^\infty \|L_\ell\|^2 < \infty$. Then*

$$\max_{\{x_\ell\}_{\ell=1}^\infty \in \mathcal{X}} \left\| \sum_{\ell=1}^\infty L_\ell \operatorname{vec}(x_\ell) \right\|^2 = \lambda_{\max}\left(\sum_{\ell=1}^\infty L_\ell L_\ell' \right). \tag{E.10}$$

Proof. A short proof using abstract functional analysis is available upon request from the authors. Below we provide a more elementary proof.

The statement of the lemma is obvious if $\sum_{\ell=1}^\infty \|L_\ell\|^2 = 0$, in which case both sides of the above display equal 0. Hence, we may assume that the series $V \equiv \sum_{\ell=1}^\infty L_\ell L_\ell'$ converges to a non-zero matrix. Let v be the unit-length eigenvector corresponding to the largest eigenvalue $\lambda \equiv \lambda_{\max}(V) \in (0, \infty)$ of V.

The purported maximum (E.10) is achieved by the sequence $\{x_\ell^*\}$ given by $\operatorname{vec}(x_\ell^*) = \lambda^{-1/2} L_\ell' v$:

$$\left\| \sum_{\ell=1}^\infty L_\ell \operatorname{vec}(x_\ell^*) \right\|^2 = \left\| \lambda^{-1/2} \sum_{\ell=1}^\infty L_\ell L_\ell' v \right\|^2 = \lambda^{-1} \|Vv\|^2 = \lambda^{-1} \|\lambda v\|^2 = \lambda \|v\|^2 = \lambda,$$

and

$$\sum_{\ell=1}^\infty \|x_\ell^*\|^2 = \sum_{\ell=1}^\infty \operatorname{vec}(x_\ell^*)' \operatorname{vec}(x_\ell^*) = \lambda^{-1} v' \sum_{\ell=1}^\infty L_\ell L_\ell' v = \lambda^{-1} v' V v = \lambda^{-1} \lambda = 1.$$

We complete the proof by showing that the left-hand side of (E.10) is bounded above by the right-hand side. Let K be an arbitrary positive integer. Then

$$\max_{\{x_\ell\}_{\ell=1}^\infty \in \mathcal{X}} \left\| \sum_{\ell=1}^\infty L_\ell \operatorname{vec}(x_\ell) \right\| \leq \max_{\{x_\ell\}_{\ell=1}^\infty \in \mathcal{X}} \left\| \sum_{\ell=1}^K L_\ell \operatorname{vec}(x_\ell) \right\| + \max_{\{x_\ell\}_{\ell=1}^\infty \in \mathcal{X}} \left\| \sum_{\ell=K+1}^\infty L_\ell \operatorname{vec}(x_\ell) \right\|.$$

The second term on the right-hand side is bounded above by $(\sum_{\ell=K+1}^{\infty} \|L_\ell\|^2)^{1/2}$ by Cauchy-Schwarz. As for the first term, standard results for the eigenvalues of finite-dimensional matrices yield

$$\begin{aligned}
\max_{\{x_\ell\}_{\ell=1}^{\infty} \in \mathcal{X}} \left\| \sum_{\ell=1}^{K} L_\ell \operatorname{vec}(x_\ell) \right\|^2 &= \max_{x \in \mathbb{R}^{Km^2} \colon \|x\| \leq 1} \left\| \begin{pmatrix} L_1 & L_2 & \cdots & L_K \end{pmatrix} x \right\|^2 \\
&= \lambda_{\max}\left(\begin{pmatrix} L_1 & L_2 & \cdots & L_K \end{pmatrix}' \begin{pmatrix} L_1 & L_2 & \cdots & L_K \end{pmatrix} \right) \\
&= \lambda_{\max}\left(\begin{pmatrix} L_1 & L_2 & \cdots & L_K \end{pmatrix} \begin{pmatrix} L_1 & L_2 & \cdots & L_K \end{pmatrix}' \right) \\
&= \lambda_{\max}\left(\sum_{\ell=1}^{K} L_\ell L_\ell' \right).
\end{aligned}$$

We have shown

$$\max_{\{x_\ell\}_{\ell=1}^{\infty} \in \mathcal{X}} \left\| \sum_{\ell=1}^{\infty} L_\ell \operatorname{vec}(x_\ell) \right\| \leq \left(\lambda_{\max}\left(\sum_{\ell=1}^{K} L_\ell L_\ell' \right) \right)^{1/2} + \left(\sum_{\ell=K+1}^{\infty} \|L_\ell\|^2 \right)^{1/2}.$$

Now let $K \to \infty$. Since $\sum_{\ell=1}^{\infty} L_\ell L_\ell'$ is a convergent series, the first term on the right-hand side above converges to $\lambda^{1/2}$ by continuity of eigenvalues, while the second term converges to 0. This establishes the required bound. □

E.2 Auxiliary numerical lemma

Lemma E.6. *Define $\overline{y}_{i,t} \equiv (y_{i+1,t}, y_{i+2,t}, \ldots, y_{nt})'$ to be the (possibly empty) vector of variables that are ordered after $y_{i,t}$ in y_t. Partition*

$$\hat{\Sigma} = \begin{pmatrix} \hat{\Sigma}_{11} & \hat{\Sigma}_{12} & \hat{\Sigma}_{13} \\ \hat{\Sigma}_{21} & \hat{\Sigma}_{22} & \hat{\Sigma}_{23} \\ \hat{\Sigma}_{31} & \hat{\Sigma}_{32} & \hat{\Sigma}_{33} \end{pmatrix}, \quad \hat{C} = \begin{pmatrix} \hat{C}_{11} & 0 & 0 \\ \hat{C}_{21} & \hat{C}_{22} & 0 \\ \hat{C}_{31} & \hat{C}_{32} & \hat{C}_{33} \end{pmatrix},$$

conformably with $y_t = (\underline{y}_{j^*,t}', y_{j^*,t}, \overline{y}_{j^*,t}')'$, where $\hat{\Sigma} = \hat{C}\hat{C}'$ (in particular, $\hat{C}_{22} = \hat{C}_{j^*,j^*}$). Then

$$(\hat{\Sigma}_{31}, \hat{\Sigma}_{32}) \begin{pmatrix} \hat{\Sigma}_{11} & \hat{\Sigma}_{12} \\ \hat{\Sigma}_{21} & \hat{\Sigma}_{22} \end{pmatrix}^{-1} e_{j^*,j^*} = \hat{C}_{22}^{-1} \hat{C}_{32}. \tag{E.11}$$

Note that the lemma implies $\hat{\beta}_0 = \hat{\delta}_0$: If $i^* < j^*$ or $i^* = j^*$, then both estimators equal 0 or 1 (by definition), respectively; if $i^* > j^*$, then $\hat{\beta}_0$ is defined as the $i^* - j^*$ element of the left-hand side of (E.11) (by Frisch-Waugh), while $\hat{\delta}_0$ is defined as the $i^* - j^*$ element of the right-hand side of (E.11).

Proof. From the relationship $\hat{\Sigma} = \hat{C}\hat{C}'$, we get

$$\begin{pmatrix} \hat{\Sigma}_{11} & \hat{\Sigma}_{12} \\ \hat{\Sigma}_{21} & \hat{\Sigma}_{22} \\ \hat{\Sigma}_{31} & \hat{\Sigma}_{32} \end{pmatrix} = \begin{pmatrix} \hat{C}_{11}\hat{C}_{11}' & \hat{C}_{11}\hat{C}_{21}' \\ \hat{C}_{21}\hat{C}_{11}' & \hat{C}_{21}\hat{C}_{21}' + \hat{C}_{22}^2 \\ \hat{C}_{31}\hat{C}_{11}' & \hat{C}_{31}\hat{C}_{21}' + \hat{C}_{32}\hat{C}_{22} \end{pmatrix}.$$

The partitioned inverse formula implies

$$\begin{aligned} \begin{pmatrix} \hat{\Sigma}_{11} & \hat{\Sigma}_{12} \\ \hat{\Sigma}_{21} & \hat{\Sigma}_{22} \end{pmatrix}^{-1} e_{j^*,j^*} &= \frac{1}{\hat{C}_{21}\hat{C}_{21}' + \hat{C}_{22}^2 - \hat{C}_{21}\hat{C}_{11}'(\hat{C}_{11}\hat{C}_{11}')^{-1}\hat{C}_{11}\hat{C}_{21}'} \begin{pmatrix} -(\hat{C}_{11}\hat{C}_{11}')^{-1}\hat{C}_{11}\hat{C}_{21}' \\ 1 \end{pmatrix} \\ &= \frac{1}{\hat{C}_{22}^2} \begin{pmatrix} -\hat{C}_{11}^{-1\prime}\hat{C}_{21}' \\ 1 \end{pmatrix}, \end{aligned}$$

so

$$(\hat{\Sigma}_{31}, \hat{\Sigma}_{32}) \begin{pmatrix} \hat{\Sigma}_{11} & \hat{\Sigma}_{12} \\ \hat{\Sigma}_{21} & \hat{\Sigma}_{22} \end{pmatrix}^{-1} e_{j^*,j^*} = \frac{1}{\hat{C}_{22}^2} \left(-\hat{C}_{31}\hat{C}_{11}'\hat{C}_{11}^{-1\prime}\hat{C}_{21}' + \hat{C}_{31}\hat{C}_{21}' + \hat{C}_{32}\hat{C}_{22} \right) = \frac{1}{\hat{C}_{22}}\hat{C}_{32}. \quad \square$$

E.3 Auxiliary asymptotic lemmas

Lemma E.7. *$T^{-1}\sum_{t=1}^T \|y_t - \tilde{y}_t\|^2 = O_p(T^{-2\zeta})$ and $T^{-1}\sum_{t=1}^T u_t(y_{t-1} - \tilde{y}_{t-1})' = O_p(T^{-2\zeta} + T^{-\zeta-1/2})$, where $u_t \equiv y_t - Ay_{t-1}$.*

Proof. Using Equation (2.1), write y_t as

$$\begin{aligned} y_t &= \sum_{s=0}^{\infty} A^s H(I_m + T^{-\zeta}\alpha(L))\varepsilon_{t-s} \\ &= \underbrace{\sum_{s=0}^{\infty} A^s H \varepsilon_{t-s}}_{\equiv \tilde{y}_t} + T^{-\zeta}\sum_{s=0}^{\infty} A^s H\alpha(L)\varepsilon_{t-s}. \end{aligned}$$

Thus, the definition of $\tilde{y}_t$ implies

$$y_t - \tilde{y}_t = T^{-\zeta}\sum_{s=0}^{\infty} A^s H\alpha(L)\varepsilon_{t-s}.$$

Lemma E.12 below shows that, under Assumption 2.1, $T^{-1}\sum_{t=1}^{T} E\left[\|y_t - \tilde{y}_t\|^2\right] = O(T^{-2\zeta})$. Consequently, the first part of Lemma E.7 follows from Markov's inequality.

In order to establish the second part of Lemma E.7, note that

$$u_t\left(y_{t-1} - \tilde{y}_{t-1}\right)' = H[I_m + T^{-\zeta}\alpha(L)]\varepsilon_t\left(y_{t-1} - \tilde{y}_{t-1}\right)'.$$

Lemma E.13 below implies that

$$\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t(y_{t-1} - \tilde{y}_{t-1})' = O_p\left(T^{-\zeta-1/2}\right). \tag{E.12}$$

Finally, Lemma E.14 below implies that

$$\frac{1}{T}\sum_{t=1}^{T}\alpha(L)\varepsilon_t(y_{t-1} - \tilde{y}_{t-1})' = O_p(T^{-\zeta}). \tag{E.13}$$

Equations (E.12) and (E.13) imply

$$\frac{1}{T}\sum_{t=1}^{T}u_t(y_{t-1} - \tilde{y}_{t-1})' = O_p(T^{-2\zeta} + T^{-\zeta-1/2}). \qquad \square$$

Lemma E.8.

$$T^{-1}\sum_{t=1}^{T}u_t y_{t-1}' = T^{-\zeta}H\sum_{\ell=1}^{\infty}\alpha_\ell D H'(A')^{\ell-1} + T^{-1}\sum_{t=1}^{T}H\varepsilon_t\tilde{y}_{t-1}' + o_p(T^{-\zeta}).$$

Proof.

$$\begin{aligned}
T^{-1}\sum_{t=1}^{T}u_t y_{t-1}' &= T^{-1}\sum_{t=1}^{T}u_t\tilde{y}_{t-1}' + \underbrace{T^{-1}\sum_{t=1}^{T}u_t(y_{t-1} - \tilde{y}_{t-1})'}_{=o_p(T^{-\zeta}) \text{ by Lemma E.7}} \\
&= T^{-1}\sum_{t=1}^{T}H\varepsilon_t\tilde{y}_{t-1}' + T^{-\zeta-1}\sum_{t=1}^{T}H\alpha(L)\varepsilon_t\tilde{y}_{t-1}' + o_p(T^{-\zeta}) \\
&= T^{-1}\sum_{t=1}^{T}H\varepsilon_t\tilde{y}_{t-1}' + T^{-\zeta}H\left(T^{-1}\sum_{t=1}^{T}E[\alpha(L)\varepsilon_t\tilde{y}_{t-1}'] + o_p(1)\right) + o_p(T^{-\zeta}),
\end{aligned}$$

where the last equality follows from Lemma E.15 below. Finally, note that

$$E[\alpha(L)\varepsilon_t \tilde{y}'_{t-1}] = \sum_{\ell=1}^{\infty}\sum_{s=0}^{\infty} \alpha_\ell E[\varepsilon_{t-\ell}\varepsilon'_{t-s-1}]H'(A')^s = \sum_{\ell=1}^{\infty} \alpha_\ell D H'(A')^{\ell-1}. \qquad \square$$

Lemma E.9. $T^{-1}\sum_{t=1}^{T} y_{t-1}y'_{t-1} \xrightarrow{p} S$.

Proof. By Lemma E.7 and Cauchy-Schwarz, $T^{-1}\sum_{t=1}^{T} y_{t-1}y'_{t-1} = T^{-1}\sum_{t=1}^{T} \tilde{y}_{t-1}\tilde{y}'_{t-1} + o_p(1)$. The rest of the proof is standard. $\square$

Lemma E.10. *Fix $h \geq 0$ and $j^* \in \{1, \ldots, n\}$. In a slight abuse of notation, let $q_t \equiv (\underline{y}'_{j^*,t}, y'_{t-1})'$. Then*

i) $T^{-1}\sum_{t=1}^{T-h} q_t \varepsilon_{j^*,t} = O_p(T^{-\zeta} + T^{-1/2})$,

ii) $(T^{-1}\sum_{t=1}^{T-h} q_t q'_t)^{-1} = O_p(1)$,

iii) $T^{-1}\sum_{t=1}^{T-h} q_t(\alpha_{j^*}(L)\varepsilon_t) = O_p(1)$,

where $\alpha_{j^*}(L)$ is the j^*-th row of $\alpha(L)$.

Proof. Let $\tilde{q}_t \equiv (\underline{\tilde{y}}'_{j^*,t}, \tilde{y}'_{t-1})'$ and $\Delta_t \equiv q_t - \tilde{q}_t$. Note that

$$T^{-1}\sum_{t=1}^{T-h} q_t \varepsilon_{j^*,t} = T^{-1}\sum_{t=1}^{T-h} \Delta_t \varepsilon_{j^*,t} + T^{-1}\sum_{t=1}^{T-h} \tilde{q}_t \varepsilon_{j^*,t}. \tag{E.14}$$

Cauchy-Schwarz implies

$$\left\| T^{-1}\sum_{t=1}^{T-h} \Delta_t \varepsilon_{j^*,t} \right\| \leq \left(\frac{1}{T}\sum_{t=1}^{T-h} \|\Delta_t\|^2\right)^{1/2} \left(\frac{1}{T}\sum_{t=1}^{T-h} \varepsilon_{j^*,t}^2\right)^{1/2}.$$

Lemma E.7 implies the first term to the right of the inequality is $O_p(T^{-\zeta})$. Assumption 2.1(i) implies that the second term to the right of the inequality is $O_p(1)$. Thus, from (E.14) we have

$$T^{-1}\sum_{t=1}^{T-h} q_t \varepsilon_{j^*,t} = O_p\left(T^{-\zeta}\right) + T^{-1}\sum_{t=1}^{T-h} \tilde{q}_t \varepsilon_{j^*,t}.$$

Direct second-moment calculations imply that the last term is $O_p\left(T^{-1/2}\right)$. This establishes part (i) of the lemma.

For part (ii) of the lemma, note that

$$\frac{1}{T}\sum_{t=1}^{T-h} q_t q_t' = \frac{1}{T}\sum_{t=1}^{T-h} \Delta_t \Delta_t' + \frac{1}{T}\sum_{t=1}^{T-h} \tilde{q}_t \Delta_t' + \frac{1}{T}\sum_{t=1}^{T-h} \Delta_t \tilde{q}_t' + \frac{1}{T}\sum_{t=1}^{T-h} \tilde{q}_t \tilde{q}_t'. \tag{E.15}$$

Lemma E.7 implies that the first term is $O_p\left(T^{-2\zeta}\right)$. Cauchy-Schwarz, along with Assumption 2.1 and Lemma E.7, imply that the second and third terms are $O_p(T^{-\zeta})$. The last term converges in probability to $\operatorname{Var}(\tilde{q}_t)$. This matrix is non-singular, since $\tilde{q}_t = (\underline{\tilde{y}}_{j^*,t}', \tilde{y}_{t-1}')'$, where $\operatorname{Var}(\tilde{y}_{t-1}) = S$ is non-singular by Assumption 2.1(iv), and Assumption 2.1(iii) implies that $\underline{\tilde{y}}_{j^*,t}$ equals a linear transformation of $\tilde{y}_{t-1}$ plus a non-singular independent noise term.

For part (iii) of the lemma, note that

$$\frac{1}{T}\sum_{t=1}^{T-h} q_t(\alpha_{j^*}(L)\varepsilon_t) = \frac{1}{T}\sum_{t=1}^{T-h} \Delta_t(\alpha_{j^*}(L)\varepsilon_t) + \frac{1}{T}\sum_{t=1}^{T-h} \tilde{q}_t(\alpha_{j^*}(L)\varepsilon_t). \tag{E.16}$$

Assumption 2.1(i) and (v) and Lemma E.7 imply that the first term is $O_p(T^{-\zeta})$. Markov's inequality and a moment calculation imply that the last term is $O_p(1)$. $\square$

Lemma E.11. *Fix $h \geq 0$ and $j^* \in \{1, \dots, n\}$. In a slight abuse of notation, let $q_t \equiv (\underline{y}_{j^*,t}', y_{t-1}')'$ and*

$$\hat{x}_{h,t} \equiv (\vartheta - \hat{\vartheta}_h)' q_t + \varepsilon_{j^*,t} + T^{-\zeta}\alpha_{j^*}(L)\varepsilon_t,$$

where $\alpha_{j^*}(L)$ is the j^*-th row of $\alpha(L)$. Then

$$T^{-1}\sum_{t=1}^{T-h} (\hat{x}_{h,t} - \varepsilon_{j^*,t})^2 = o_p(T^{-1/2}). \tag{E.17}$$

Proof. Let $\tilde{q}_t \equiv (\underline{\tilde{y}}_{j^*,t}', \tilde{y}_{t-1}')'$ and $\Delta_t \equiv q_t - \tilde{q}_t$. Then

$$T^{-1}\sum_{t=1}^{T-h} (\hat{x}_{h,t} - \varepsilon_{j^*,t})^2 = T^{-1}\sum_{t=1}^{T-h} \left((\vartheta - \hat{\vartheta}_h)'\Delta_t + (\vartheta - \hat{\vartheta}_h)'\tilde{q}_t + T^{-\zeta}\alpha_{j^*}(L)\varepsilon_t\right)^2.$$

To establish (E.17), it suffices by the c_r-inequality to show that

a) $T^{-1}\sum_{t=1}^{T-h}\left((\vartheta - \hat{\vartheta}_h)'\Delta_t\right)^2 = o_p\left(T^{-1/2}\right)$,

b) $T^{-1}\sum_{t=1}^{T-h}\left((\vartheta - \hat{\vartheta}_h)'\tilde{q}_t\right)^2 = o_p\left(T^{-1/2}\right)$,

c) $T^{-1}\sum_{t=1}^{T-h}\left(\alpha_{j^*}(L)\varepsilon_t\right)^2 = O_p\left(1\right)$.

To establish (a), note first that Cauchy-Schwarz implies

$$\frac{1}{T}\sum_{t=1}^{T-h}\left((\vartheta-\hat{\vartheta}_h)'\Delta_t\right)^2 \leq \left\|\vartheta-\hat{\vartheta}_h\right\|^2\left(\frac{1}{T}\sum_{t=1}^{T-h}\|\Delta_t\|^2\right).$$

Lemma E.7 implies that the term inside the parenthesis is $O_p(T^{-2\zeta})$. Lemma E.10 implies $(\vartheta-\hat{\vartheta}_h)=O_p\left(T^{-\zeta}+T^{-1/2}\right)$. Since $\zeta>1/4$, statement (a) follows.

To establish (b), we apply Cauchy-Schwarz to obtain

$$\frac{1}{T}\sum_{t=1}^{T-h}\left((\vartheta-\hat{\vartheta}_h)'\tilde{q}_t\right)^2 \leq \left\|\vartheta-\hat{\vartheta}_h\right\|^2\left(\frac{1}{T}\sum_{t=1}^{T-h}\tilde{q}_t^2\right).$$

Assumption 2.1 implies that the term inside the parenthesis is $O_p(1)$. As in the previous paragraph, $\|\vartheta-\hat{\vartheta}_h\|^2=O_p\left((T^{-\zeta}+T^{-1/2})^2\right)$. Since $\zeta>1/4$, statement (b) follows.

Finally, statement (c) follows from Assumption 2.1(i) and (v). □

E.4 Auxiliary lemmas to the auxiliary lemmas

Lemma E.12. *There exists a constant* $\tilde{C}\in(0,\infty)$ *such that*

$$E\left[\|y_t-\tilde{y}_t\|^2\right]\leq \tilde{C}T^{-2\zeta}. \tag{E.18}$$

Proof. The definition of $\tilde{y}_t$ implies

$$y_t-\tilde{y}_t=T^{-\zeta}\sum_{s=0}^{\infty}A^sH\alpha(L)\varepsilon_{t-s}.$$

Expanding $\alpha(L)=\sum_{\ell=1}^{\infty}\alpha_\ell L^\ell$, we obtain

$$y_t-\tilde{y}_t=T^{-\zeta}\sum_{s=1}^{\infty}B_s\varepsilon_{t-s}, \quad \text{where} \quad B_s\equiv\sum_{\ell=1}^{s}A^{s-\ell}H\alpha_\ell. \tag{E.19}$$

By the independence assumption on ε_t in Assumption 2.1(i),

$$E\left[\|y_t-\tilde{y}_t\|^2\right]=T^{-2\zeta}\sum_{s=1}^{\infty}\operatorname{trace}\left(B_sDB_s'\right).$$

Expanding B_s and changing the summation indices shows that $E\left[\|y_t - \tilde{y}_t\|^2\right]$ equals

$$T^{-2\zeta} \sum_{s=1}^{\infty} \sum_{\ell_1=1}^{s} \sum_{\ell_2=1}^{s} \operatorname{trace}\left(A^{s-\ell_1} H \alpha_{\ell_1} D \alpha'_{\ell_2} H' (A')^{s-\ell_2} \right).$$

Moreover, since for any two matrices M_1, M_2 of conformable dimensions $\operatorname{trace}(M_1 M_2) \leq \|M_1\| \|M_2\|$, then

$$\operatorname{trace}\left(A^{s-\ell_1} H \alpha_{\ell_1} D \alpha'_{\ell_2} H' (A')^{s-\ell_2} \right) \leq \|H\|^2 \cdot \|D\| \cdot \|A^{s-\ell_1}\| \cdot \|(A')^{s-\ell_2}\| \cdot \|\alpha_{\ell_1}\| \cdot \|\alpha_{\ell_2}\|.$$

Let $\lambda \in [0, 1)$ and $C > 0$ be chosen such that $\|A^\ell\| \leq C\lambda^\ell$ for all $\ell \geq 0$ (such constants exists by Assumption 2.1(ii)). Then

$$\begin{aligned} E\left[\|y_t - \tilde{y}_t\|^2\right] &\leq T^{-2\zeta} C^2 \|H\|^2 \|D\| \left(\sum_{\tau=0}^{\infty} \lambda^{2\tau} \right) \left(\sum_{\ell_1=1}^{\infty} \|\alpha_{\ell_1}\| \right) \left(\sum_{\ell_2=1}^{\infty} \|\alpha_{\ell_2}\| \right), \\ &\leq T^{-2\zeta} C^2 \|H\|^2 \|D\| \left(\sum_{\tau=0}^{\infty} \lambda^{2\tau} \right) \left(\sum_{\ell=1}^{\infty} \|\alpha_\ell\| \right)^2, \\ &= T^{-2\zeta} \frac{C^2 \|H\|^2 \|D\|}{1 - \lambda^2} \left(\sum_{\ell=1}^{\infty} \|\alpha_\ell\| \right)^2. \end{aligned}$$

□

Lemma E.13.

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \varepsilon_t \left(y_{t-1} - \tilde{y}_{t-1} \right)' = O_p\left(T^{-\zeta} \right).$$

Proof. By Markov's inequality, we need to show that the following expression is bounded:

$$T^{2\zeta} E\left[\left\| \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \varepsilon_t \left(y_{t-1} - \tilde{y}_{t-1} \right)' \right\|^2 \right].$$

Equation (E.19) in the proof of Lemma E.12 and Assumption 2.1(i) imply that the summands are serially uncorrelated, so the above expression equals

$$\begin{aligned} &T^{2\zeta} \frac{1}{T} \sum_{t=1}^{T} E\left[\|\varepsilon_t \left(y_{t-1} - \tilde{y}_{t-1} \right)'\|^2 \right] \\ &\leq T^{2\zeta} \frac{1}{T} \sum_{t=1}^{T} E\left[\|\varepsilon_t\|^2 \|y_{t-1} - \tilde{y}_{t-1}\|^2 \right], \\ &= T^{2\zeta} \frac{1}{T} \sum_{t=1}^{T} E\left[\|\varepsilon_t\|^2 \right] E\left[\|y_{t-1} - \tilde{y}_{t-1}\|^2 \right], \end{aligned}$$

$$= T^{2\zeta}\, \text{trace}\,(D)\, E\left[\|y_{t-1} - \tilde{y}_{t-1}\|^2\right].$$

The third line follows from Assumption 2.1(i), while the last line follows from stationarity. Lemma E.12 implies that the final expression is bounded. □

Lemma E.14.
$$\frac{1}{T}\sum_{t=1}^{T} \alpha(L)\varepsilon_t(y_{t-1} - \tilde{y}_{t-1})' = O_p(T^{-\zeta}).$$

Proof. By Markov's inequality, we need to show that

$$T^{\zeta} E\left[\left\|\frac{1}{T}\sum_{t=1}^{T} \alpha(L)\varepsilon_t\left(y_{t-1} - \tilde{y}_{t-1}\right)'\right\|\right]$$

is bounded. By stationarity and Cauchy-Schwarz, the expression is bounded above by

$$\begin{aligned} &T^{\zeta} E\left[\|\alpha(L)\varepsilon_t\|\, \|y_{t-1} - \tilde{y}_{t-1}\|\right] \\ &\le T^{\zeta}\left(E\left[\|\alpha(L)\varepsilon_t\|^2\right]\right)^{1/2}\left(E\left[\|y_{t-1} - \tilde{y}_{t-1}\|^2\right]\right)^{1/2}. \end{aligned}$$

The first expectation on the right-hand side is bounded due to Assumption 2.1(v). Hence, Lemma E.12 implies that the entire final expression is bounded. □

Lemma E.15.
$$T^{-1}\sum_{t=1}^{T}\left(\alpha(L)\varepsilon_t\tilde{y}_{t-1}' - E[\alpha(L)\varepsilon_t\tilde{y}_{t-1}']\right) = o_p(1).$$

Proof. For an arbitrary $i \in \{1, \dots, n\}$ and $s \ge 1$, define

$$\begin{aligned} \Gamma_s &\equiv \text{Cov}(\alpha(L)\varepsilon_t\tilde{y}_{i,t-1}, \alpha(L)\varepsilon_{t-s}\tilde{y}_{i,t-s-1}) \\ &= \text{Cov}\left(\sum_{\ell_1=1}^{\infty}\alpha_{\ell_1}\varepsilon_{t-\ell_1}\tilde{y}_{i,t-1}, \sum_{\ell_2=1}^{\infty}\alpha_{\ell_2}\varepsilon_{t-s-\ell_2}\tilde{y}_{i,t-s-1}\right) \\ &= \sum_{\ell_1=1}^{\infty}\sum_{\ell_2=1}^{\infty}\alpha_{\ell_1}\,\text{Cov}(\varepsilon_{t-\ell_1}\tilde{y}_{i,t-1}, \varepsilon_{t-s-\ell_2}\tilde{y}_{i,t-s-1})\alpha_{\ell_2}'. \end{aligned}$$

By Theorem 7.1.1 in Brockwell and Davis (1991), the statement of the lemma follows if we can show that $\Gamma_s \to 0$ as $s \to \infty$.

Decompose

$$\tilde{y}_{i,t-1} = \underbrace{E[\tilde{y}_{i,t-1} \mid \{\varepsilon_{t-s}\}_{s=1}^{\ell_1-1}]}_{\equiv\tilde{y}_{i,t-1}^{(-)}} + \underbrace{E[\tilde{y}_{i,t-1} \mid \varepsilon_{t-\ell_1}]}_{\equiv\tilde{y}_{i,t-1}^{(0)}} + \underbrace{E[\tilde{y}_{i,t-1} \mid \{\varepsilon_{t-s}\}_{s=\ell_1+1}^{\infty}]}_{\equiv\tilde{y}_{i,t-1}^{(+)}}.$$

For $\ell_1 \leq s$, the serial independence of ε_t implies that

$$\text{Cov}(\varepsilon_{t-\ell_1}\tilde{y}^{(-)}_{i,t-1}, \varepsilon_{t-s-\ell_2}\tilde{y}_{i,t-s-1}) = E[\tilde{y}^{(-)}_{i,t-1}]E[\varepsilon_{t-\ell_1}\varepsilon'_{t-s-\ell_2}\tilde{y}_{i,t-s-1}] = 0,$$

$$\text{Cov}(\varepsilon_{t-\ell_1}\tilde{y}^{(0)}_{i,t-1}, \varepsilon_{t-s-\ell_2}\tilde{y}_{i,t-s-1}) = 0,$$

$$\text{Cov}(\varepsilon_{t-\ell_1}\tilde{y}^{(+)}_{i,t-1}, \varepsilon_{t-s-\ell_2}\tilde{y}_{i,t-s-1}) = E[\varepsilon_{t-\ell_1}]E[\tilde{y}^{(+)}_{i,t-1}\varepsilon'_{t-s-\ell_2}\tilde{y}_{i,t-s-1}] = 0,$$

and therefore

$$\text{Cov}(\varepsilon_{t-\ell_1}\tilde{y}_{i,t-1}, \varepsilon_{t-s-\ell_2}\tilde{y}_{i,t-s-1}) = 0.$$

Inserting this result back into the earlier expression for Γ_s, we get

$$\begin{aligned}
|\Gamma_s| &= \left| \sum_{\ell_1=s+1}^{\infty} \sum_{\ell_2=1}^{\infty} \alpha_{\ell_1} \text{Cov}(\varepsilon_{t-\ell_1}\tilde{y}_{i,t-1}, \varepsilon_{t-s-\ell_2}\tilde{y}_{i,t-s-1})\alpha'_{\ell_2} \right| \\
&\leq \sum_{\ell_1=s+1}^{\infty} \sum_{\ell_2=1}^{\infty} \|\alpha_{\ell_1}\| \cdot \|\alpha_{\ell_2}\| \cdot \| \text{Cov}(\varepsilon_{t-\ell_1}\tilde{y}_{i,t-1}, \varepsilon_{t-s-\ell_2}\tilde{y}_{i,t-s-1})\| \\
&\leq \sum_{\ell_1=s+1}^{\infty} \sum_{\ell_2=1}^{\infty} \|\alpha_{\ell_1}\| \cdot \|\alpha_{\ell_2}\| \cdot \sup_{\ell \geq 1} \| \text{Var}(\varepsilon_{t-\ell}\tilde{y}_{i,t-1})\| \\
&\leq \underbrace{\left(E[\|\varepsilon_t^4\|] \cdot E[\tilde{y}^4_{i,t}]\right)^{1/2} \left(\sum_{\ell_2=1}^{\infty} \|\alpha_{\ell_2}\|\right)}_{<\infty} \left(\sum_{\ell_1=s+1}^{\infty} \|\alpha_{\ell_1}\|\right) \\
&\to 0 \quad \text{as } s \to \infty,
\end{aligned}$$

where the last line uses absolute summability of $\alpha(L)$. □

References

Brockwell, P. J., and R. A. Davis (1991): *Time Series: Theory and Methods*, Springer Series in Statistics. Springer, 2nd edn.

Francis, N., M. T. Owyang, J. E. Roush, and R. DiCecio (2014): "A flexible finite-horizon alternative to long-run restrictions with an application to technology shocks," *Review of Economics and Statistics*, 96(4), 638–647.

Gertler, M., and P. Karadi (2015): "Monetary Policy Surprises, Credit Costs, and Economic Activity," *American Economic Journal: Macroeconomics*, 7(1), 44–76.

Plagborg-Møller, M., and C. K. Wolf (2021): "Local Projections and VARs Estimate the Same Impulse Responses," *Econometrica*, 89(2), 955–980.

Ramey, V. A. (2011): "Identifying Government Spending Shocks: It's all in the Timing," *Quarterly Journal of Economics*, 126(1), 1–50.

Ramey, V. A. (2016): "Macroeconomic Shocks and Their Propagation," in *Handbook of Macroeconomics*, ed. by J. B. Taylor, and H. Uhlig, vol. 2, chap. 2, pp. 71–162. Elsevier.

Romer, C. D., and D. H. Romer (2010): "The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks," *American Economic Review*, 100(3), 763–801.

Smets, F., and R. Wouters (2007): "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American Economic Review*, 97(3), 586–606.