

Online Appendix

This Online Appendix contains supplementary theoretical, computational, and empirical material for “Granular Instrumental Variables: Estimation and Inference.” Section A provides implementation details for the proposed GIV procedures. Sections B – E contain proofs of the theoretical results and additional technical derivations. Section F reports supplementary Monte Carlo evidence, and G presents additional empirical results. References cited only in the Online Appendix are collected at the end of this document.

A Implementation Details

ALGORITHM 1 (GIV ESTIMATION WITH BALANCED DATA)

Step 0. If exogenous regressors x_t are present, compute the LS estimator

$$\hat{\beta} \equiv \left(\sum_{t \leq T} \tilde{x}_t^\top \tilde{x}_t \right)^{-1} \left(\sum_{t \leq T} \tilde{x}_t^\top \tilde{y}_t \right),$$

and define $\hat{y}_t \equiv y_t - x_t \hat{\beta}$. Otherwise, set $\hat{y}_t = y_t$. If $\bar{r}$ is unknown, estimate it using the BIC criterion in (47)–(48) and replace $\bar{r}$ by the resulting estimator $\hat{r}$ throughout the algorithm.

Step 1. Compute $y_{S,t} \equiv S_t^\top y_t$ and $y_{e,t} \equiv e^\top \hat{y}_t$, where $\hat{y}_t \equiv (\hat{y}_{i,t})_{i \leq n}$, $S_t = (s_{i,t})_{i \leq n}$, and $e \equiv n^{-1} \mathbf{1}_n$. Compute

$$Q_{-1}^\top \left(T^{-1} \sum_{t \leq T} \hat{y}_t \hat{y}_t^\top \right) Q_{-1},$$

and obtain ordered eigenvalues $\{\hat{\mu}_j\}_{j \leq n-1}$ such that $\hat{\mu}_j \leq \hat{\mu}_{j+1}$. Define

$$\hat{\Upsilon} \equiv Q_{-1} \hat{A}_{0,\perp} (\hat{\sigma}_u^2 \mathbf{I}_{\bar{r}-1} - \hat{\Lambda}_\perp)^{-1} \hat{A}_{0,\perp}^\top Q_{-1}^\top,$$

where $\hat{\sigma}_u^2 \equiv \hat{\mu}_1$, $\hat{\Lambda}_\perp \equiv \text{diag}((\hat{\mu}_j)_{n-\bar{r}+1 \leq j \leq n-1})$, and $\hat{A}_0$ and $\hat{A}_{0,\perp}$ collect the eigenvectors associated with the smallest $n - \bar{r}$ and largest $\bar{r} - 1$ eigenvalues, respectively.

Step 2. Set $\hat{A} = Q_{-1} \hat{A}_0$ and compute the preliminary GIV estimator $\hat{\theta}_0(\hat{A})$ using (39) together with the appropriate moment specification:

$$\begin{aligned} \text{Without } w_t: & \quad \text{use (40) and } W_{0,T}(\hat{A}) = \mathbf{I}_{2(n-\bar{r})} \\ \text{With } w_t: & \quad \text{use (66) and } W_{0,T}(\hat{A}) = \mathbf{I}_{2(n-\bar{r})+d_w} \end{aligned}.$$

Step 3. Construct residuals

$$\hat{v}_t \equiv y_{e,t} - \hat{\phi}_0(\hat{A}) p_t,$$

and

$$\hat{\varepsilon}_t \equiv \begin{cases} p_t - \hat{\psi}_0(\hat{A}) y_{S,t}, & \text{without } w_t, \\ p_t - \hat{\psi}_0(\hat{A}) y_{S,t} - w_t^\top \hat{\gamma}_0(\hat{A}), & \text{with } w_t. \end{cases}$$

For $b \in \{v, \varepsilon\}$, define

$$\hat{\xi}_{b,t} \equiv \hat{y}_t \hat{b}_t - T^{-1} \sum_{t \leq T} \hat{y}_t \hat{b}_t + \left(\hat{y}_t \hat{y}_t^\top - T^{-1} \sum_{t \leq T} \hat{y}_t \hat{y}_t^\top \right) \hat{\Upsilon} \left(T^{-1} \sum_{t \leq T} \hat{y}_t \hat{b}_t \right). \quad (\text{A.1})$$

Step 4. Define

$$\hat{\xi}_t \equiv \begin{cases} (\hat{\xi}_{v,t}^\top, \hat{\xi}_{\varepsilon,t}^\top)^\top, & \text{without } w_t, \\ (\hat{\xi}_{v,t}^\top, \hat{\xi}_{\varepsilon,t}^\top, w_t^\top \hat{\varepsilon}_t)^\top, & \text{with } w_t. \end{cases}$$

Compute

$$\hat{V} \equiv T^{-1} \sum_{t \leq T} \hat{\xi}_t \hat{\xi}_t^\top - \left(T^{-1} \sum_{t \leq T} \hat{\xi}_t \right) \left(T^{-1} \sum_{t \leq T} \hat{\xi}_t^\top \right). \quad (\text{A.2})$$

Use a HAC version if ξ_t is serially correlated.

Step 5. Compute the optimal GIV estimator $\hat{\theta}^*(\hat{A})$ using (39) with weight matrix

$$W_{*,T}(\hat{A}) \equiv \begin{cases} ((\mathbf{I}_2 \otimes \hat{A}^\top) \hat{V} (\mathbf{I}_2 \otimes \hat{A}))^{-1}, & \text{without } w_t, \\ (\text{diag}(\hat{A}^\top, \hat{A}^\top, \mathbf{I}_{d_w}) \hat{V} \text{diag}(\hat{A}, \hat{A}, \mathbf{I}_{d_w}))^{-1}, & \text{with } w_t. \end{cases}$$

Step 6. Report standard errors

$$\left(\text{diag} \left((T D_{1,T}(\hat{A})^\top W_{*,T}(\hat{A}) D_{1,T}(\hat{A}))^{-1} \right) \right)^{1/2},$$

and the p -value based on the J -statistic and the CDF of $\chi^2(2(n - \bar{r} - 1))$.

ALGORITHM 2 (GIV ESTIMATION WITH UNBALANCED DATA)

Step 1. If $\bar{r}$ is unknown, estimate it using the BIC criterion in (47)–(48); otherwise treat $\bar{r}$ as known. Compute $y_{S,t} \equiv S_t^\top y_t$ and $y_{e,t} \equiv e_t^\top y_t$, where $y_t \equiv (y_{i,t})_{i \leq n_t}$, $S_t = (s_{i,t})_{i \leq n_t}$ and $e_t \equiv n_t^{-1} \mathbf{1}_{n_t}$.

Step 2. Let y_t^0 denote the $n_0 \times 1$ subvector of y_t observed for all t . Let $Q_{0,-1} \equiv (q_{0,2}, \dots, q_{0,n_0})$ be defined analogously to (13) with n replaced by n_0 . Compute

$$Q_{0,-1}^\top \left(T^{-1} \sum_{t \leq T} y_t^0 y_t^{0\top} \right) Q_{0,-1},$$

and obtain its ordered eigenvalues $\{\hat{\mu}_{0,j}\}_{j \leq n_0-1}$. Obtain

$$\hat{\Upsilon}_0 \equiv Q_{0,-1} \hat{A}_{0,\perp} (\hat{\sigma}_u^2 \mathbf{I}_{\bar{r}_0-1} - \hat{\Lambda}_{0,\perp})^{-1} \hat{A}_{0,\perp}^\top Q_{0,-1}^\top,$$

where $\hat{\sigma}_u^2 \equiv \hat{\mu}_{0,1}$, $\hat{\Lambda}_{0,\perp} \equiv \text{diag}((\hat{\mu}_{0,j})_{n_0-\bar{r}_0+1 \leq j \leq n_0-1})$, and $\hat{A}_0$ and $\hat{A}_{0,\perp}$ collect the eigenvectors associated with the smallest $n_0 - \bar{r}_0$ eigenvalues and the largest $\bar{r}_0 - 1$ eigenvalues, respectively.

Step 3. Set $\hat{A} = Q_{0,-1} \hat{A}_0$ and $W_{0,T}(\hat{A}) = \mathbf{I}_{2(n_0-\bar{r}_0)}$. Compute the preliminary GIV estimator $\hat{\theta}_0(\hat{A}) \equiv (\hat{\phi}_0(\hat{A}), \hat{\psi}_0(\hat{A}))^\top$ using (68) and (69). Construct the residuals

$$\hat{v}_t \equiv y_{e,t} - \hat{\phi}_0(\hat{A}) p_t, \quad \hat{\varepsilon}_t \equiv p_t - \hat{\psi}_0(\hat{A}) y_{S,t},$$

and define, for $b \in \{v, \varepsilon\}$,

$$\hat{\xi}_{b,t}^0 \equiv y_t^0 \hat{b}_t - T^{-1} \sum_{t \leq T} y_t^0 \hat{b}_t + \left(y_t^0 y_t^{0\top} - T^{-1} \sum_{t \leq T} y_t^0 y_t^{0\top} \right) \hat{\Upsilon}_0 \left(T^{-1} \sum_{t \leq T} y_t^0 \hat{b}_t \right).$$

Step 4. Let $\hat{\xi}_t^0 \equiv (\hat{\xi}_{v,t}^{0\top}, \hat{\xi}_{\varepsilon,t}^{0\top})^\top$. Define

$$\hat{V}^0 \equiv T^{-1} \sum_{t \leq T} \hat{\xi}_t^0 \hat{\xi}_t^{0\top} - \left(T^{-1} \sum_{t \leq T} \hat{\xi}_t^0 \right) \left(T^{-1} \sum_{t \leq T} \hat{\xi}_t^0 \right)^\top.$$

Use a HAC version of $\hat{V}^0$ if ξ_t^0 is serially correlated.

Step 5. Compute the optimal GIV estimator $\hat{\theta}^*(\hat{A})$ using (68) and (69) with $W_{0,T}(\hat{A})$ replaced by

$$W_{*,T}(\hat{A}) \equiv ((\mathbf{I}_2 \otimes \hat{A}^\top) \hat{V}^0 (\mathbf{I}_2 \otimes \hat{A}))^{-1}.$$

Step 6. Report the standard errors

$$\sqrt{\text{Diag} \left((T D_{1,T}^0(\hat{A})^\top W_{*,T}(\hat{A}) D_{1,T}^0(\hat{A}))^{-1} \right)},$$

and the p -value based on $T \bar{g}_T^0(\hat{\theta}^*(\hat{A}); \hat{A})^\top W_{*,T}(\hat{A}) \bar{g}_T^0(\hat{\theta}^*(\hat{A}); \hat{A})$ and the CDF of $\chi^2(2(n_0 - \bar{r}_0 - 1))$.

B Proof of the Main Results

PROOF OF LEMMA 1. Multiplying both sides of (11) by Q yields

$$\mathbb{E}[Q^\top (y_t - \phi p_t \mathbf{1}_n)(y_t - \phi p_t \mathbf{1}_n)^\top Q] = (\mathbb{E}[\eta_t^2] + 2\sigma_{\eta u}) Q^\top \mathbf{1}_n \mathbf{1}_n^\top Q + \sigma_u^2 \mathbf{I}_n. \quad (\text{B.1})$$

Since Q is orthonormal, this can be written as

$$\mathbb{E} \left[\begin{pmatrix} n(y_{e,t} - \phi p_t)^2 & n^{1/2}(y_{e,t} - \phi p_t) y_t^\top Q_{-1} \\ n^{1/2} Q_{-1}^\top y_t (y_{e,t} - \phi p_t) & Q_{-1}^\top y_t y_t^\top Q_{-1} \end{pmatrix} \right] = \begin{pmatrix} n(\mathbb{E}[\eta_t^2] + 2\sigma_{\eta u}) + \sigma_u^2 & \mathbf{0}_{1 \times (n-1)} \\ \mathbf{0}_{(n-1) \times 1} & \sigma_u^2 \mathbf{I}_{n-1} \end{pmatrix}. \quad (\text{B.2})$$

Matching the upper-right block, the lower-right block, and the $(1, 1)$ entry of the matrices on both sides of (B.2) yields the moment conditions in (14)–(16). The moment conditions in (17)–(18) can be obtained similarly by multiplying $Q^\top$ on the left-hand side of both terms in (12). $Q.E.D.$

PROOF OF LEMMA 2. Multiplying $(\bar{\lambda}, \bar{\lambda}_\perp)^\top$ from the left and $(\bar{\lambda}, \bar{\lambda}_\perp)$ from the right on both sides of (21) yields

$$\mathbb{E} \left[\bar{\lambda}^\top (y_t - \phi p_t \mathbf{1}_n) y_t^\top \bar{\lambda}_\perp \right] = \mathbf{0}_{\bar{r} \times (n-\bar{r})}, \quad (\text{B.3})$$

$$\mathbb{E} \left[\bar{\lambda}_\perp^\top y_t y_t^\top \bar{\lambda}_\perp \right] - \sigma_{u,t}^2 \mathbf{I}_{n-\bar{r}} = \mathbf{0}_{(n-\bar{r}) \times (n-\bar{r})}, \quad (\text{B.4})$$

$$\mathbb{E} \left[\bar{\lambda}^\top (y_t - \phi p_t \mathbf{1}_n)(y_t - \phi p_t \mathbf{1}_n)^\top \bar{\lambda} \right] = \bar{\lambda}^\top \left(\lambda \mathbb{E}[\eta_t \eta_t^\top] \lambda^\top + \lambda \Gamma_{\eta u, t} \mathbf{1}_n^\top + \mathbf{1}_n \Gamma_{\eta u, t}^\top \lambda^\top + \sigma_{u,t}^2 \mathbf{I}_n \right) \bar{\lambda}. \quad (\text{B.5})$$

By the definition of $\bar{\lambda}$,

$$\bar{\lambda}^\top (y_t - \phi p_t \mathbf{1}_n) y_t^\top \bar{\lambda}_\perp = \begin{pmatrix} n^{-1/2} \mathbf{1}_n^\top (y_t - \phi p_t \mathbf{1}_n) y_t^\top \bar{\lambda}_\perp \\ \bar{\lambda}_{-1}^\top (y_t - \phi p_t \mathbf{1}_n) y_t^\top \bar{\lambda}_\perp \end{pmatrix} = \begin{pmatrix} n^{1/2} (y_{e,t} - \phi p_t) y_t^\top \bar{\lambda}_\perp \\ \bar{\lambda}_{-1}^\top y_t y_t^\top \bar{\lambda}_\perp \end{pmatrix}.$$

Therefore, (23) and (24) follow directly from (B.3). Likewise, (25) and (26) follow from (B.4) and (B.5), respectively.

Next, multiplying $(\bar{\lambda}, \bar{\lambda}_\perp)^\top$ from the left on both sides of (22) yields

$$\mathbb{E} \left[\bar{\lambda}^\top (p_t - \psi y_{S,t}) (y_t - \phi p_t \mathbf{1}_n) \right] = \bar{\lambda}^\top \lambda \mathbb{E}[\varepsilon_t \eta_t] + \sigma_{\varepsilon u, t} \bar{\lambda}^\top \mathbf{1}_n, \quad (\text{B.6})$$

$$\mathbb{E} \left[\bar{\lambda}_\perp^\top (p_t - \psi y_{S,t}) (y_t - \phi p_t \mathbf{1}_n) \right] = \mathbf{0}_{n-\bar{r}}. \quad (\text{B.7})$$

Since

$$\bar{\lambda}_\perp^\top (p_t - \psi y_{S,t}) (y_t - \phi p_t \mathbf{1}_n) = (p_t - \psi y_{S,t}) \bar{\lambda}_\perp^\top y_t,$$

(27) follows from (B.7), while (28) follows from (B.6). *Q.E.D.*

PROOF OF LEMMA 3. Let $\tilde{r} \equiv \text{rank}(\tilde{\lambda})$. Since $\text{rank}((\mathbf{1}_n, M_{\mathbf{1}_n} \lambda)) = \text{rank}((\mathbf{1}_n, \lambda))$ and $\lambda^\top M_{\mathbf{1}_n} \mathbf{1}_n = \mathbf{0}_r$, we have

$$\tilde{r} = \text{rank}((\mathbf{1}_n, \lambda)) - 1 = \bar{r} - 1. \quad (\text{B.8})$$

This together with the assumption that $T^{-1} \sum_{t \leq T} \mathbb{E}[\eta_t \eta_t^\top]$ is nonsingular implies that the matrix $\tilde{\lambda} \left(T^{-1} \sum_{t \leq T} \mathbb{E}[\eta_t \eta_t^\top] \right) \tilde{\lambda}^\top$ has rank $\bar{r} - 1$. Hence, for eigenvectors in $\mathcal{B}_{\mathbf{1}_n}$, it has $\bar{r} - 1$ strictly positive eigenvalues, denoted by $\tilde{\mu}_j > 0$ for $j = 1, \dots, \bar{r} - 1$, and $n - \bar{r}$ zero eigenvalues. For any $x \in \mathcal{B}_{\mathbf{1}_n}$, we have $x^\top M_{\mathbf{1}_n} x = 1$, so restricting (32) to $\mathcal{B}_{\mathbf{1}_n}$ yields

$$x^\top \bar{\Sigma}_{\tilde{y}} x = x^\top \tilde{\lambda} \left(T^{-1} \sum_{t \leq T} \mathbb{E}[\eta_t \eta_t^\top] \right) \tilde{\lambda}^\top x + \bar{\sigma}_u^2 x^\top M_{\mathbf{1}_n} x = x^\top \tilde{\lambda} \left(T^{-1} \sum_{t \leq T} \mathbb{E}[\eta_t \eta_t^\top] \right) \tilde{\lambda}^\top x + \bar{\sigma}_u^2.$$

Therefore, the restricted eigenvalues of $\bar{\Sigma}_{\tilde{y}}$ on $\mathcal{B}_{\mathbf{1}_n}$ are given by $\tilde{\mu}_j + T^{-1} \sum_{t \leq T} \sigma_{u,t}^2$, for $j = 1, \dots, \bar{r} - 1$, and $\bar{\sigma}_u^2$ with multiplicity $n - \bar{r}$. Since $n > \bar{r}$, the minimum eigenvalue of $\bar{\Sigma}_{\tilde{y}}$ on $\mathcal{B}_{\mathbf{1}_n}$ is $\bar{\sigma}_u^2$ with multiplicity $n - \bar{r}$. This establishes (33).

Finally, the minimizers of (33) are the unit eigenvectors associated with the eigenvalue $\bar{\sigma}_u^2$, which are precisely the vectors in $\mathcal{B}_{\mathbf{1}_n}$ orthogonal to the column space of $\tilde{\lambda}$. Equivalently, they are orthogonal to both $\mathbf{1}_n$ and $\text{col}(\tilde{\lambda})$. Since $\tilde{\lambda} = M_{\mathbf{1}_n} \lambda$, we have

$$\text{col}((\mathbf{1}_n, \tilde{\lambda})) = \text{col}((\mathbf{1}_n, \lambda)).$$

Hence these eigenvectors are exactly the orthogonal complement of $\text{col}((\mathbf{1}_n, \lambda))$, and therefore they span $\bar{\lambda}_\perp$. *Q.E.D.*

PROOF OF LEMMA 4. Recall that $Q = (n^{-1/2} \mathbf{1}_n, Q_{-1})$ is an orthonormal matrix, so the columns of Q_{-1} form an orthonormal basis for the subspace orthogonal to $\mathbf{1}_n$. Therefore, the mapping

$\tilde{a} \mapsto a = Q_{-1}\tilde{a}$ is a bijection from B_{n-1} onto $\mathcal{B}_{1_n}$. In particular, $a \in \mathcal{B}_{1_n}$ if and only if $a = Q_{-1}\tilde{a}$ for some $\tilde{a} \in B_{n-1}$.

Hence, the minimization problem in (33) is equivalent to

$$\min_{a \in \mathcal{B}_{1_n}} a^\top \bar{\Sigma}_{\tilde{y}} a = \min_{\tilde{a} \in B_{n-1}} \tilde{a}^\top Q_{-1}^\top \bar{\Sigma}_{\tilde{y}} Q_{-1} \tilde{a}.$$

Next, using $\tilde{y}_t = M_{1_n} y_t$ and the fact that $M_{1_n} Q_{-1} = Q_{-1}$ (equivalently, $Q_{-1}^\top M_{1_n} = Q_{-1}^\top$), we obtain

$$Q_{-1}^\top \mathbb{E}[\tilde{y}_t \tilde{y}_t^\top] Q_{-1} = Q_{-1}^\top M_{1_n} \mathbb{E}[y_t y_t^\top] M_{1_n} Q_{-1} = Q_{-1}^\top \mathbb{E}[y_t y_t^\top] Q_{-1}.$$

Substituting this into the objective function yields (34).

Therefore, $a \in \mathcal{B}_{1_n}$ is a minimizer of (33) if and only if $a = Q_{-1}\tilde{a}$ for some minimizer $\tilde{a} \in B_{n-1}$ of (34). This completes the proof. Q.E.D.

PROOF OF THEOREM 1. We first establish the claims of the theorem with $\hat{A}$ replaced by $\tilde{A}$ where $\tilde{A} \equiv \hat{A} \hat{H}_0^\top$ and $\hat{H}_0$ is a rotation matrix defined in (C.35) of Lemma C.6. We then show that

$$\hat{\theta}(\tilde{A}) = \hat{\theta}(\hat{A}) \tag{B.9}$$

and

$$\Gamma(D_{1,T}, W_{0,T}, \hat{A}) \hat{V}(\hat{A}) \Gamma(D_{1,T}, W_{0,T}, \hat{A})^\top = \Gamma(D_{1,T}, W_{0,T}, \tilde{A}) \hat{V}(\tilde{A}) \Gamma(D_{1,T}, W_{0,T}, \tilde{A})^\top \tag{B.10}$$

wpa1, where $\hat{V}(\tilde{A}) \equiv (\mathbf{I}_2 \otimes \tilde{A}^\top) \hat{V}(\mathbf{I}_2 \otimes \tilde{A})$. The claims of the theorem therefore follow.

By Assumption 3(ii), $\max_{t \leq T} \|S_t\| \leq K$. Therefore, from Assumptions 2(i, iv), (C.25) and (C.26) in the proof of Lemma C.4, it follows that

$$\max_{t \leq T} \left(\|\mathbb{E}[y_t p_t]\| + \|\mathbb{E}[y_t y_t^\top]\| \right) \leq K. \tag{B.11}$$

This along with Lemmas C.4 and C.5 implies that

$$\left\| T^{-1} \sum_{t \leq T} y_t p_t \right\| + \left\| T^{-1} \sum_{t \leq T} y_t y_{S,t} \right\| \leq K, \tag{B.12}$$

wpa1. From Lemmas C.4 and C.5, (B.12) and (C.71) in the proof of Lemma C.11, it follows that

$$D_{1,T}(\tilde{A}) = (\mathbf{I}_2 \otimes \tilde{A}^\top) T^{-1} \sum_{t \leq T} \begin{pmatrix} y_t p_t & \mathbf{0}_n \\ \mathbf{0}_n & y_t y_{S,t} \end{pmatrix} = D_1(A) + O_p(T^{-1/2}) = O_p(1). \tag{B.13}$$

Similarly, by Assumption 4(ii) and (C.71),

$$W_{0,T}(\tilde{A}) = ((\mathbf{I}_2 \otimes \tilde{A}^\top) W_{0,T}(\mathbf{I}_2 \otimes \tilde{A}))^{-1} = W_0(A) + o_p(1) = O_p(1), \tag{B.14}$$

where $W_0(A) \equiv ((\mathbf{I}_2 \otimes A^\top)W_0(\mathbf{I}_2 \otimes A))^{-1}$. Combining (B.13) and (B.14) yields

$$D_{1,T}(\tilde{A})^\top W_{0,T}(\tilde{A})D_{1,T}(\tilde{A}) = D_1(A)^\top W_0(A)D_1(A) + o_p(1). \quad (\text{B.15})$$

Since $A^\top A = A_0^\top Q_{-1}^\top Q_{-1}A_0 = \mathbf{I}_{n-\bar{r}}$, by Assumption 4(ii),

$$\rho_{\min}(W_0(A)) \geq (\rho_{\max}(W_0))^{-1}(\rho_{\max}(\mathbf{I}_2 \otimes A^\top A))^{-1} \geq K^{-1}.$$

This together with Assumption 4(iii) implies

$$\begin{aligned} \rho_{\min}\left(D_1(A)^\top W_0(A)D_1(A)\right) &\geq \rho_{\min}(W_0(A))\rho_{\min}\left(D_1(A)^\top D_1(A)\right) \\ &\geq K^{-1} \min\left\{\left\|T^{-1} \sum_{t \leq T} A^\top \mathbb{E}[y_t p_t]\right\|^2, \left\|T^{-1} \sum_{t \leq T} A^\top \mathbb{E}[y_t y_{S,t}]\right\|^2\right\} \\ &\geq K^{-1}. \end{aligned} \quad (\text{B.16})$$

From (B.15) and (B.16), it follows that

$$\rho_{\min}\left(D_{1,T}(\tilde{A})^\top W_{0,T}(\tilde{A})D_{1,T}(\tilde{A})\right) \geq (2K)^{-1}, \quad \text{wpa1}. \quad (\text{B.17})$$

By the definition of $D_{2,T}(\tilde{A})$, we can write

$$\begin{aligned} D_{2,T}(\tilde{A}) &= (\mathbf{I}_2 \otimes \tilde{A}^\top)T^{-1} \sum_{t \leq T} \begin{pmatrix} \phi y_t p_t + y_t(y_{e,t} - \phi p_t) \\ \psi y_t y_{S,t} + y_t(p_t - \psi y_{S,t}) \end{pmatrix} \\ &= (\mathbf{I}_2 \otimes \tilde{A}^\top)T^{-1} \sum_{t \leq T} \begin{pmatrix} \phi y_t p_t \\ \psi y_t y_{S,t} \end{pmatrix} + (\mathbf{I}_2 \otimes \tilde{A}^\top)T^{-1} \sum_{t \leq T} \begin{pmatrix} y_t(y_{e,t} - \phi p_t) \\ y_t(p_t - \psi y_{S,t}) \end{pmatrix} \\ &= D_{1,T}(\tilde{A})\theta + T^{-1} \sum_{t \leq T} \begin{pmatrix} \tilde{A}^\top y_t v_t \\ \tilde{A}^\top y_t \varepsilon_t \end{pmatrix}. \end{aligned} \quad (\text{B.18})$$

By Assumption 4(i) and Lemma C.11,

$$T^{-1/2} \sum_{t \leq T} \begin{pmatrix} \tilde{A}^\top y_t v_t \\ \tilde{A}^\top y_t \varepsilon_t \end{pmatrix} = T^{-1/2} \sum_{t \leq T} \begin{pmatrix} A^\top \xi_{v,t} \\ A^\top \xi_{\varepsilon,t} \end{pmatrix} + O_p(T^{-1/2}) = O_p(1). \quad (\text{B.19})$$

From the definition of $\hat{\theta}(\tilde{A})$, (B.17) and (B.18), it follows that

$$\hat{\theta}(\tilde{A}) = \theta + (D_{1,T}(\tilde{A})^\top W_{0,T}(\tilde{A})D_{1,T}(\tilde{A}))^{-1}D_{1,T}(\tilde{A})^\top W_{0,T}(\tilde{A})T^{-1} \sum_{t \leq T} \begin{pmatrix} \tilde{A}^\top y_t v_t \\ \tilde{A}^\top y_t \varepsilon_t \end{pmatrix}. \quad (\text{B.20})$$

This along with (B.13), (B.14), (B.15) and (B.19), and (C.71) implies that

$$T^{1/2}(\hat{\theta}(\tilde{A}) - \theta) = (D_{1,T}(\tilde{A})^\top W_{0,T}(\tilde{A})D_{1,T}(\tilde{A}))^{-1}D_{1,T}(\tilde{A})^\top W_{0,T}(\tilde{A})T^{-1/2} \sum_{t \leq T} \begin{pmatrix} \tilde{A}^\top y_t v_t \\ \tilde{A}^\top y_t \varepsilon_t \end{pmatrix}$$

$$= \Gamma(D_1, W_0, A)(\mathbf{I}_2 \otimes A^\top)T^{-1/2} \sum_{t \leq T} \xi_t + o_p(1). \quad (\text{B.21})$$

For any vector $a \in \mathbb{R}^2$ with $a^\top a = 1$, define

$$a_T \equiv V^{1/2}(\mathbf{I}_2 \otimes A)\Gamma(D_1, W_0, A)^\top (\Gamma(D_1, W_0, A)V(A)\Gamma(D_1, W_0, A)^\top)^{-1/2}a.$$

Then by the definition of $\Gamma(D_1, W_0, A)$, we have $a_T^\top a_T = 1$. From (B.21) it follows that

$$a^\top (\Gamma(D_1, W_0, A)V(A)\Gamma(D_1, W_0, A)^\top)^{-1/2}T^{1/2}(\hat{\theta}(\tilde{A}) - \theta) = a_T^\top V^{-1/2}T^{-1/2} \sum_{t \leq T} \xi_t + o_p(1). \quad (\text{B.22})$$

Under Assumption 4(i),

$$V^{-1/2}T^{-1/2} \sum_{t \leq T} \xi_t \rightarrow_d N(0, \mathbf{I}_{2n}). \quad (\text{B.23})$$

Since the set $\{a \in \mathbb{R}^{2n} : a^\top a = 1\}$ is compact, for any subsequence of $\{T\}$, we can extract a further subsequence $\{T_k\}$ such that

$$a_{T_k} \rightarrow a_\infty \in \left\{a \in \mathbb{R}^{2n} : a^\top a = 1\right\}.$$

From (B.23), it is clear that $V^{-1/2}T_k^{-1/2} \sum_{t \leq T_k} \xi_t = O_p(1)$. Therefore,

$$a_{T_k}^\top V^{-1/2}T_k^{-1/2} \sum_{t \leq T_k} \xi_t = a_\infty^\top V^{-1/2}T_k^{-1/2} \sum_{t \leq T_k} \xi_t + (a_{T_k} - a_\infty)^\top V^{-1/2}T_k^{-1/2} \sum_{t \leq T_k} \xi_t \rightarrow_d N(0, 1),$$

from which, we can conclude that

$$a_T^\top V^{-1/2}T^{-1/2} \sum_{t \leq T} \xi_t \rightarrow_d N(0, 1). \quad (\text{B.24})$$

Combining (B.22) and (B.24), and applying Slutsky's theorem and the Cramer-Wold device, we deduce that

$$(\Gamma(D_1, W_0, A)V(A)\Gamma(D_1, W_0, A)^\top)^{-1/2}T^{1/2}(\hat{\theta}(\tilde{A}) - \theta) \rightarrow_d N(0, \mathbf{I}_2). \quad (\text{B.25})$$

This establishes (42) with $\hat{A}$ replaced by $\tilde{A}$. By (B.13), (B.14), (B.15) and (B.16)

$$\begin{aligned} \Gamma(D_{1,T}, W_{0,T}, \tilde{A}) &= (D_{1,T}(\tilde{A})^\top W_{0,T}(\tilde{A})D_{1,T}(\tilde{A}))^{-1}D_{1,T}(\tilde{A})^\top W_{0,T}(\tilde{A}) \\ &= (D_1(A)^\top W_0(A)D_1(A))^{-1}D_1(A)^\top W_0(A) + o_p(1) \\ &= \Gamma(D_1, W_0, A) + o_p(1) = O_p(1). \end{aligned} \quad (\text{B.26})$$

By similar arguments for deriving (B.14), we can show that

$$\hat{V}(\tilde{A}) = (\mathbf{I}_2 \otimes \tilde{A}^\top)\hat{V}(\mathbf{I}_2 \otimes \tilde{A}) = V(A) + o_p(1) = O_p(1), \quad (\text{B.27})$$

which along with (B.26) establishes (43) with $\hat{A}$ replaced by $\tilde{A}$.

Using the definition of $J_T(\theta; \cdot, \cdot)$ in (C.22), the GMM problem in (37) can be written as

$$\hat{\theta}(\hat{A}) \equiv \arg \min_{\theta \in \Theta} J_T(\theta; \hat{A}, W_{0,T}). \quad (\text{B.28})$$

The sufficient conditions of Lemma C.3 hold for $A_1 = \hat{A}$, $W_1 = W_{0,T}$ and $C_1 = \hat{H}_0^\top$ because: (i) $\hat{A}^\top \hat{A} = \hat{A}_0^\top Q_{-1}^\top Q_{-1} \hat{A}_0 = \mathbf{I}_{n-\bar{r}}$ and thus $\hat{A}$ has full rank by construction; (ii) $W_{0,T}$ is positive definite wpa1 by Assumption 4(ii); and (iii) $\hat{H}_0^\top$ is nonsingular wpa1 by Lemma C.6. Therefore, applying Lemma C.3 leads to

$$J_T(\theta; \hat{A}, W_{0,T}) = J_T(\theta; \hat{A} \hat{H}_0^\top, W_{0,T}) = J_T(\theta; \tilde{A}, W_{0,T}) \quad (\text{B.29})$$

wpa1, which together with (B.28) shows that $\hat{\theta}(\hat{A}) = \hat{\theta}(\tilde{A})$ wpa1.

To verify the claim in (B.10), we first observe that

$$D_{1,T}(\tilde{A}) = D_{1,T}(\hat{A} \hat{H}_0^\top) = (\mathbf{I}_2 \otimes \hat{H}_0) D_{1,T}(\hat{A}) \quad (\text{B.30})$$

and

$$W_{0,T}(\tilde{A}) = W_{0,T}(\hat{A} \hat{H}_0^\top) = (\mathbf{I}_2 \otimes \hat{H}_0^\top)^{-1} W_{0,T}(\hat{A}) (\mathbf{I}_2 \otimes \hat{H}_0)^{-1}, \quad (\text{B.31})$$

where $\hat{H}_0$ is nonsingular wpa1 by Lemma C.6. Therefore,

$$\begin{aligned} & D_{1,T}(\tilde{A})^\top W_{0,T}(\tilde{A}) D_{1,T}(\tilde{A}) \\ &= D_{1,T}(\hat{A})^\top (\mathbf{I}_2 \otimes \hat{H}_0^\top) (\mathbf{I}_2 \otimes \hat{H}_0^\top)^{-1} W_{0,T}(\hat{A}) (\mathbf{I}_2 \otimes \hat{H}_0)^{-1} (\mathbf{I}_2 \otimes \hat{H}_0) D_{1,T}(\hat{A}) \\ &= D_{1,T}(\hat{A})^\top W_{0,T}(\hat{A}) D_{1,T}(\hat{A}), \quad \text{wpa1.} \end{aligned} \quad (\text{B.32})$$

Next note that

$$\hat{V}(\tilde{A}) = \hat{V}(\hat{A} \hat{H}_0^\top) = (\mathbf{I}_2 \otimes \hat{H}_0) \hat{V}(\hat{A}) (\mathbf{I}_2 \otimes \hat{H}_0^\top),$$

which along with (B.30) and (B.31) implies that

$$\begin{aligned} & D_{1,T}(\tilde{A})^\top W_{0,T}(\tilde{A}) \hat{V}(\tilde{A}) W_{0,T}(\tilde{A}) D_{1,T}(\tilde{A}) \\ &= D_{1,T}(\hat{A})^\top (\mathbf{I}_2 \otimes \hat{H}_0^\top) (\mathbf{I}_2 \otimes \hat{H}_0^\top)^{-1} W_{0,T}(\hat{A}) (\mathbf{I}_2 \otimes \hat{H}_0)^{-1} \\ &\quad \times (\mathbf{I}_2 \otimes \hat{H}_0) \hat{V}(\hat{A}) (\mathbf{I}_2 \otimes \hat{H}_0^\top) \\ &\quad \times (\mathbf{I}_2 \otimes \hat{H}_0^\top)^{-1} W_{0,T}(\hat{A}) (\mathbf{I}_2 \otimes \hat{H}_0)^{-1} (\mathbf{I}_2 \otimes \hat{H}_0) D_{1,T}(\hat{A}) \\ &= D_{1,T}(\hat{A})^\top W_{0,T}(\hat{A}) \hat{V}(\hat{A}) W_{0,T}(\hat{A}) D_{1,T}(\hat{A}), \quad \text{wpa1.} \end{aligned} \quad (\text{B.33})$$

The claim in (B.10) now follows from (B.17), (B.32), (B.33) and the definitions of $\Gamma(D_{1,T}, W_{0,T}, \hat{A})$, $\Gamma(D_{1,T}, W_{0,T}, \tilde{A})$, $\hat{V}(\hat{A})$ and $\hat{V}(\tilde{A})$. *Q.E.D.*

PROOF OF THEOREM 2. By similar arguments in deriving (B.29), we can verify the sufficient conditions of Lemma C.3 for $A_1 = \hat{A}$, $W_1 = \hat{V}$ and $C_1 = \hat{H}_0^\top$. Hence, we can apply Lemma C.3 to show that

$$J_T(\theta; \hat{A}, \hat{V}) = J_T(\theta; \hat{A}\hat{H}_0^\top, \hat{V}) = J_T(\theta; \tilde{A}, \hat{V})$$

wpa1, from which the claim of the theorem follows if

$$J_T(\hat{\theta}^*(\tilde{A}); \tilde{A}, \hat{V}) \rightarrow_d \chi^2(2(n - \bar{r} - 1)). \quad (\text{B.34})$$

Note that $\hat{\theta}^*(\tilde{A})^\top = (\hat{\phi}^*(\tilde{A}), \hat{\psi}^*(\tilde{A}))$. To prove (B.34), we first observe that

$$\begin{aligned} T^{1/2}\bar{g}_T(\hat{\theta}^*(\tilde{A}); \tilde{A}) &= T^{-1/2} \sum_{t \leq T} \begin{pmatrix} \tilde{A}^\top y_t (y_{e,t} - \hat{\phi}^*(\tilde{A})p_t) \\ \tilde{A}^\top y_t (p_t - \hat{\psi}^*(\tilde{A})y_{S,t}) \end{pmatrix} \\ &= T^{-1/2} \sum_{t \leq T} \begin{pmatrix} \tilde{A}^\top y_t v_t \\ \tilde{A}^\top y_t \varepsilon_t \end{pmatrix} - D_{1,T}(\tilde{A})T^{1/2}(\hat{\theta}^*(\tilde{A}) - \theta). \end{aligned} \quad (\text{B.35})$$

By the same arguments for showing (B.21) but replacing $W_{0,T}(\tilde{A})$ by $W_{*,T}(\tilde{A})$, we obtain

$$\begin{aligned} T^{1/2}(\hat{\theta}^*(\tilde{A}) - \theta) &= (D_1(A)^\top V(A)^{-1} D_1(A))^{-1} D_1(A)^\top V(A)^{-1} (\mathbf{I}_2 \otimes A^\top) T^{-1/2} \sum_{t \leq T} \xi_t + o_p(1) \\ &= \Gamma(D_1, V, A) V(A)^{1/2} T^{-1/2} \sum_{t \leq T} \xi_t^* + o_p(1), \end{aligned} \quad (\text{B.36})$$

where $\xi_t^* \equiv V(A)^{-1/2}(\mathbf{I}_2 \otimes A^\top)\xi_t$. This along with Assumptions 4(i, iii) shows that

$$T^{1/2}(\hat{\theta}^*(\tilde{A}) - \theta) = O_p(1). \quad (\text{B.37})$$

From (B.13) and (B.37), it follows that

$$D_{1,T}(\tilde{A})T^{1/2}(\hat{\theta}^*(\tilde{A}) - \theta) = D_1(A)\Gamma(D_1, V, A)V(A)^{1/2}T^{-1/2} \sum_{t \leq T} \xi_t^* + o_p(1). \quad (\text{B.38})$$

Combining (B.19), (B.35) and (B.38) yields

$$\begin{aligned} T^{1/2}\bar{g}_T(\hat{\theta}^*(\tilde{A}); \tilde{A}) &= (\mathbf{I}_2 \otimes A^\top)T^{-1/2} \sum_{t \leq T} \xi_t \\ &\quad - D_1(A)\Gamma(D_1, V, A)V(A)^{1/2}T^{-1/2} \sum_{t \leq T} \xi_t^* + o_p(1) \\ &= V(A)^{1/2}T^{-1/2} \sum_{t \leq T} \xi_t^* - D_1(A)\Gamma(D_1, V, A)V(A)^{1/2}T^{-1/2} \sum_{t \leq T} \xi_t^* + o_p(1) \\ &= V(A)^{1/2}\Pi(A)T^{-1/2} \sum_{t \leq T} \xi_t^* + o_p(1) = O_p(1), \end{aligned} \quad (\text{B.39})$$

where

$$\Pi(A) \equiv \mathbf{I}_{2(n-\bar{r})} - V(A)^{-1/2} D_1(A) (D_1(A)^\top V(A)^{-1} D_1(A))^{-1} D_1(A)^\top V(A)^{-1/2}.$$

By the definition of $W_{*,T}(\tilde{A})$ and (B.27),

$$(W_{*,T}(\tilde{A}))^{-1} = \hat{V}(\tilde{A}) = V(A) + o_p(1) \quad (\text{B.40})$$

which together with Assumption 4(i) implies that

$$\rho_{\max}(W_{*,T}(\tilde{A})) \leq 2K, \quad (\text{B.41})$$

wpa1. Using (B.39), (B.40) and (B.41), we conclude that

$$\begin{aligned} J_T(\hat{\theta}^*(\tilde{A}); \tilde{A}, \hat{V}) &= T\bar{g}_T(\hat{\theta}^*(\tilde{A}); \tilde{A})^\top W_{*,T}(\tilde{A})\bar{g}_T(\hat{\theta}^*(\tilde{A}); \tilde{A}) \\ &= \left(T^{-1/2} \sum_{t \leq T} \xi_t^*\right)^\top \Pi(A) \left(T^{-1/2} \sum_{t \leq T} \xi_t^*\right) + o_p(1). \end{aligned} \quad (\text{B.42})$$

It is clear that $\Pi(A)$ is an idempotent matrix with

$$\text{rank}(\Pi(A)) = 2(n - \bar{r}) - 2. \quad (\text{B.43})$$

By Assumption 4(i) and similar arguments for showing (B.24), we have

$$T^{-1/2} \sum_{t \leq T} \xi_t^* \rightarrow_d N(0, \mathbf{I}_{2(n-\bar{r})}). \quad (\text{B.44})$$

The claim in (B.34) follows from (B.42)-(B.44) and Slutsky's Theorem. *Q.E.D.*

PROOF OF THEOREM 3. Under Assumptions 2(i, ii, iii) and 3, we can apply Lemma C.4 to obtain

$$\|\hat{S}_y - S_y\| = O_p(T^{-1/2}), \quad (\text{B.45})$$

which implies that

$$\max_{j \in \mathcal{J}} |\hat{\mu}_j - \mu_j| = O_p(T^{-1/2}). \quad (\text{B.46})$$

We first consider the case $j > \bar{r}$. For any $j \in \{\bar{r} + 1, \dots, n - 1\}$,

$$\begin{aligned} \frac{\text{BIC}_T(j) - \text{BIC}_T(\bar{r})}{\log(T)} &= j - \bar{r} + (n - j)^{-1} \sum_{s=1}^{n-j} \frac{T(\hat{\mu}_s - \hat{\mu}_1)^2}{2\hat{\mu}_s^2 \log(T)} - (n - \bar{r})^{-1} \sum_{s=1}^{n-\bar{r}} \frac{T(\hat{\mu}_s - \hat{\mu}_1)^2}{2\hat{\mu}_s^2 \log(T)} \\ &\geq j - \bar{r} - (n - \bar{r})^{-1} \sum_{s=n-j+1}^{n-\bar{r}} \frac{T(\hat{\mu}_s - \hat{\mu}_1)^2}{\hat{\mu}_s^2 \log(T)} \\ &= j - \bar{r} - (n - \bar{r})^{-1} \sum_{s=n-j+1}^{n-\bar{r}} \frac{(T^{1/2}(\hat{\mu}_s - \mu_s) - T^{1/2}(\hat{\mu}_1 - \bar{\sigma}_u^2))^2}{(\bar{\sigma}_u^2 + \hat{\mu}_s - \bar{\sigma}_u^2)^2 \log(T)} \\ &= j - \bar{r} + O_p(\log(T)^{-1}), \end{aligned}$$

where the equality in the third line follows from the fact that $\mu_j = \bar{\sigma}_u^2$ for $j \in \{1, \dots, n - \bar{r}\}$, and the last equality follows from Assumption 2(iii) together with (B.46). Consequently,

$$\min_{j \in \{\bar{r}+1, \dots, n-1\}} \frac{\text{BIC}_T(j) - \text{BIC}_T(\bar{r})}{j - \bar{r}} > \frac{\log(T)}{2}, \quad \text{wpa1.} \quad (\text{B.47})$$

Next, consider the case $j < \bar{r}$. For any $j \in \{1, \dots, \bar{r} - 1\}$,

$$\begin{aligned} \frac{\text{BIC}_T(j) - \text{BIC}_T(\bar{r})}{T} &= (n-j)^{-1} \sum_{s=1}^{n-j} \frac{(\hat{\mu}_s - \hat{\mu}_1)^2}{2\hat{\mu}_s^2} - (n-\bar{r})^{-1} \sum_{s=1}^{n-\bar{r}} \frac{(\hat{\mu}_s - \hat{\mu}_1)^2}{2\hat{\mu}_s^2} + \frac{(j-\bar{r})\log(T)}{T} \\ &= (n-j)^{-1} \sum_{s=n-\bar{r}+1}^{n-j} \frac{(\hat{\mu}_s - \hat{\mu}_1)^2}{2\hat{\mu}_s^2} \\ &\quad + \frac{(j-\bar{r})}{(n-j)(n-\bar{r})} \sum_{s=1}^{n-\bar{r}} \frac{(\hat{\mu}_s - \hat{\mu}_1)^2}{2\hat{\mu}_s^2} + \frac{(j-\bar{r})\log(T)}{T}. \end{aligned} \quad (\text{B.48})$$

From Assumption 2(iii) and (B.46), we have

$$\frac{(j-\bar{r})}{(n-j)(n-\bar{r})} \sum_{s=1}^{n-\bar{r}} \frac{(\hat{\mu}_s - \hat{\mu}_1)^2}{2\hat{\mu}_s^2} = O_p(T^{-1}). \quad (\text{B.49})$$

Similarly,

$$(n-j)^{-1} \sum_{s=n-\bar{r}+1}^{n-j} \frac{(\hat{\mu}_s - \hat{\mu}_1)^2}{2\hat{\mu}_s^2} \geq \frac{\bar{r}-j}{n-j} \frac{(\hat{\mu}_{n-\bar{r}+1} - \hat{\mu}_1)^2}{2\hat{\mu}_{n-j}^2} = \frac{\bar{r}-j}{n-j} \frac{(\mu_{n-\bar{r}+1} - \bar{\sigma}_u^2)^2}{2\mu_{n-j}^2} + O_p(T^{-1/2}).$$

Combining this result with (B.48) and (B.49) yields

$$\frac{\text{BIC}_T(j) - \text{BIC}_T(\bar{r})}{T} \geq \frac{\bar{r}-j}{n-j} \frac{(\mu_{n-\bar{r}+1} - \bar{\sigma}_u^2)^2}{2\mu_{n-j}^2} + O_p(T^{-1/2}). \quad (\text{B.50})$$

By (B.11), $\mu_{n-j}^2 \leq K$, while Assumption 2(iii) implies that $\mu_{n-\bar{r}+1} - \bar{\sigma}_u^2 > K^{-1}$. Therefore, (B.50) implies that

$$\min_{j \in \{1, \dots, \bar{r}-1\}} \frac{\text{BIC}_T(j) - \text{BIC}_T(\bar{r})}{\bar{r}-j} > K^{-1}T, \quad \text{wpa1.} \quad (\text{B.51})$$

The claim of the theorem now follows directly from (B.47) and (B.51). *Q.E.D.*

PROOF OF LEMMA 5. Since $M_{\mathbf{1}_n} \lambda_{-1} = \lambda_{-1}$, the moment condition (59) is equivalent to

$$M_{\mathbf{1}_n} \mathbb{E}[y_t y_t^\top] M_{\mathbf{1}_n} \lambda_{-1} = \lambda_{-1} (\lambda_{-1}^\top \lambda_{-1})^{-1} \lambda_{-1}^\top \mathbb{E}[y_t y_t^\top] \lambda_{-1}. \quad (\text{B.52})$$

Since $\mathbb{E}[y_t y_t^\top]$ is nonsingular while $M_{\mathbf{1}_n}$ has rank $n-1$, $M_{\mathbf{1}_n} \mathbb{E}[y_t y_t^\top] M_{\mathbf{1}_n}$ has rank $n-1$. Hence its null space is spanned by $\mathbf{1}_n$, and its range is contained in

$$\mathcal{H}_{\mathbf{1}_n} \equiv \{a \in \mathbb{R}^n : \mathbf{1}_n^\top a = 0\}.$$

Therefore, there exists an orthonormal basis $D_M \equiv (d_1, \dots, d_{n-1})$ of $\mathcal{H}_{1_n}$ such that

$$M_{1_n} \mathbb{E}[y_t y_t^\top] M_{1_n} D_M = D_M \text{diag}(\rho_1, \dots, \rho_{n-1}), \quad (\text{B.53})$$

where $\rho_1, \dots, \rho_{n-1}$ are the nonzero eigenvalues of $M_{1_n} \mathbb{E}[y_t y_t^\top] M_{1_n}$.

Now fix any subset $J \subset \{1, \dots, n-1\}$ with $|J| = r-1$, and let D_J collect the corresponding columns of D_M . Then

$$\mathbf{1}_n^\top D_J = 0, \quad D_J^\top D_J = \mathbf{I}_{r-1}, \quad (\text{B.54})$$

and

$$M_{1_n} \mathbb{E}[y_t y_t^\top] M_{1_n} D_J = D_J \text{diag}\{(\rho_j)_{j \in J}\}. \quad (\text{B.55})$$

Because $M_{1_n} D_J = D_J$, we also have

$$D_J^\top \mathbb{E}[y_t y_t^\top] D_J = D_J^\top M_{1_n} \mathbb{E}[y_t y_t^\top] M_{1_n} D_J = \text{diag}\{(\rho_j)_{j \in J}\}. \quad (\text{B.56})$$

Combining (B.55) and (B.56) yields

$$M_{1_n} \mathbb{E}[y_t y_t^\top] M_{1_n} D_J = D_J (D_J^\top D_J)^{-1} D_J^\top \mathbb{E}[y_t y_t^\top] D_J.$$

This is precisely (B.52) with λ_{-1} replaced by D_J . Hence D_J satisfies (59).

Finally, since $D_J^\top D_J = \mathbf{I}_{r-1}$ and $\mathbf{1}_n^\top D_J = 0$, the scaled matrix $n^{1/2} D_J$ satisfies the normalization in (49). The condition (59) is invariant to nonsingular right multiplication of D_J , and therefore it is also satisfied by $n^{1/2} D_J$. This proves the claim. *Q.E.D.*

PROOF OF LEMMA 6. Since C_1 and C_2 are orthogonal, we have

$$\mathbf{1}_n^\top \lambda_{C_1} = \mathbf{1}_n^\top \lambda_{-1} C_1 = \mathbf{0}_{r-1}^\top \quad \text{and} \quad n^{-1} \lambda_{C_1}^\top \lambda_{C_1} = n^{-1} C_1^\top \lambda_{-1}^\top \lambda_{-1} C_1 = \mathbf{I}_{r-1}.$$

Moreover,

$$\lambda_{C_2}^\top (\mathbf{1}_n, \lambda_{C_1}) = C_2^\top \lambda_\perp^\top (\mathbf{1}_n, \lambda_{-1} C_1) = \mathbf{0}_{(n-r) \times r}.$$

This shows that λ_{C_1} and λ_{C_2} satisfy the same restrictions of λ_{-1} and $\lambda_\perp$.

We next verify the moment conditions. Since

$$\lambda_{C_1} (C_1^\top b_y) = \lambda_{-1} C_1 C_1^\top b_y = \lambda_{-1} b_y, \quad \lambda_{C_1}^\top y_t = C_1^\top \lambda_{-1}^\top y_t, \quad \text{and} \quad \lambda_{C_2}^\top y_t = C_2^\top \lambda_\perp^\top y_t,$$

we have

$$\begin{aligned} & \mathbb{E} \left[(y_{e,t} - \phi p_t - y_t^\top \lambda_{C_1} C_1^\top b_y) \begin{pmatrix} \lambda_{C_2}^\top y_t \\ \lambda_{C_1}^\top y_t \end{pmatrix} \right] \\ &= \text{diag}(C_2^\top, C_1^\top) \mathbb{E} \left[(y_{e,t} - \phi p_t - y_t^\top \lambda_{-1} b_y) \begin{pmatrix} \lambda_\perp^\top y_t \\ \lambda_{-1}^\top y_t \end{pmatrix} \right] = \mathbf{0}_{n-1}. \end{aligned}$$

Thus (57) continues to hold after the rotation.

Similarly, since $\lambda_{C_1}(C_1^\top b_p) = \lambda_{-1}b_p$, we obtain

$$\begin{aligned} & \mathbb{E} \left[(p_t - \psi y_{S,t} - y_t^\top \lambda_{C_1} C_1^\top b_p) \begin{pmatrix} \lambda_{C_2}^\top y_t \\ \lambda_{C_1}^\top y_t \end{pmatrix} \right] \\ &= \text{diag}(C_2^\top, C_1^\top) \mathbb{E} \left[(p_t - \psi y_{S,t} - y_t^\top \lambda_{-1} b_p) \begin{pmatrix} \lambda_{-1}^\top y_t \\ \lambda_{-1}^\top y_t \end{pmatrix} \right] = \mathbf{0}_{n-1}. \end{aligned}$$

Therefore, (58) also continues to hold.

Finally, since

$$\lambda_{C_1}(\lambda_{C_1}^\top \lambda_{C_1})^{-1} \lambda_{C_1}^\top = \lambda_{-1}(\lambda_{-1}^\top \lambda_{-1})^{-1} \lambda_{-1}^\top,$$

the projection matrix appearing in (59) is unchanged when λ_{-1} is replaced by λ_{C_1} . Moreover,

$$y_t y_t^\top \lambda_{C_1} = y_t y_t^\top \lambda_{-1} C_1.$$

Hence the left-hand side of (59) under λ_{C_1} equals the original left-hand side multiplied by C_1 , and is therefore zero. Thus (59) holds for λ_{C_1} as well. *Q.E.D.*

C Auxiliary Lemmas

This subsection presents a set of lemmas that verify Assumptions 2(iii) and 4(iii) under primitive conditions, as well as auxiliary results used in the proofs of the main results in Sections 2 and 3. We begin with lemmas that establish Assumptions 2(iii) and 4(iii).

Lemma C.1. *Suppose that $\rho_{\min}(T^{-1} \sum_{t \leq T} \mathbb{E}[\eta_t \eta_t^\top]) \geq K^{-1}$, and either (i) $\rho_{\min}((\mathbf{1}_n, \lambda)^\top (\mathbf{1}_n, \lambda)) \geq K^{-1}$; or (ii) $\mathbf{1}_n \in \text{col}(\lambda)$ and $K^{-1} \leq \rho_{\min}(\lambda^\top \lambda / n) \leq \rho_{\max}(\lambda^\top \lambda / n) \leq K$. Then Assumption 2(iii) holds.*

PROOF OF LEMMA C.1. We begin by noting that since $Q_{-1}^\top Q_{-1} = \mathbf{I}_{n-1}$ and $Q_{-1}^\top \mathbf{1}_n = \mathbf{0}_{n-1}$,

$$Q_{-1} Q_{-1}^\top = Q_{-1} (Q_{-1}^\top Q_{-1})^{-1} Q_{-1}^\top$$

is the projection matrix onto the subspace orthogonal to $\mathbf{1}_n$. Therefore, $\mathbf{I}_n - Q_{-1} Q_{-1}^\top$ is the projection matrix onto $\text{col}(\mathbf{1}_n)$, which implies

$$\mathbf{I}_n - Q_{-1} Q_{-1}^\top = n^{-1} \mathbf{1}_n \mathbf{1}_n^\top.$$

Hence,

$$Q_{-1} Q_{-1}^\top = M_{\mathbf{1}_n}. \tag{C.1}$$

Let $\Sigma_\eta \equiv T^{-1} \sum_{t \leq T} \mathbb{E}[\eta_t \eta_t^\top]$. The decomposition in (32) can be written as

$$\bar{\Sigma}_{\tilde{y}} = \tilde{\lambda} \Sigma_\eta \tilde{\lambda}^\top + \bar{\sigma}_u^2 M_{\mathbf{1}_n}, \tag{C.2}$$

where $\tilde{\lambda} \equiv M_{\mathbf{1}_n} \lambda$. Since $\bar{\Sigma}_{\tilde{y}} = M_{\mathbf{1}_n} \bar{\Sigma}_y M_{\mathbf{1}_n}$ and $Q_{-1}^\top M_{\mathbf{1}_n} = Q_{-1}^\top$, it follows that

$$Q_{-1}^\top \bar{\Sigma}_{\tilde{y}} Q_{-1} = Q_{-1}^\top \bar{\Sigma}_y Q_{-1}, \quad Q_{-1}^\top \tilde{\lambda} \Sigma_\eta \tilde{\lambda}^\top Q_{-1} = Q_{-1}^\top \lambda \Sigma_\eta \lambda^\top Q_{-1}.$$

Therefore, multiplying $Q_{-1}^\top$ and Q_{-1} from the left and right of (C.2) yields

$$Q_{-1}^\top \bar{\Sigma}_y Q_{-1} = Q_{-1}^\top \lambda \Sigma_\eta \lambda^\top Q_{-1} + \bar{\sigma}_u^2 \mathbf{I}_{n-1}. \quad (\text{C.3})$$

Since $\mathbf{1}_n^\top \tilde{\lambda} = 0$, we have $\tilde{\lambda} \in \text{col}(Q_{-1})$, and hence

$$\text{rank}(Q_{-1}^\top \tilde{\lambda}) = \text{rank}(\tilde{\lambda}).$$

Noting that $Q_{-1}^\top \tilde{\lambda} = Q_{-1}^\top M_{\mathbf{1}_n} \lambda = Q_{-1}^\top \lambda$, it follows that

$$\text{rank}(Q_{-1}^\top \lambda) = \text{rank}(\tilde{\lambda}) = \bar{r} - 1. \quad (\text{C.4})$$

Because $\rho_{\min}(\Sigma_\eta) \geq K^{-1}$, (C.4) implies that $Q_{-1}^\top \lambda \Sigma_\eta \lambda^\top Q_{-1}$ has $n - \bar{r}$ zero eigenvalues and $\bar{r} - 1$ strictly positive eigenvalues. From (C.3), the $(n - \bar{r} + 1)$ th eigenvalue of $Q_{-1}^\top \bar{\Sigma}_y Q_{-1}$, denoted by $\mu_{n-\bar{r}+1}$, satisfies

$$\mu_{n-\bar{r}+1} = \rho_{\min}^+ \left(Q_{-1}^\top \lambda \Sigma_\eta \lambda^\top Q_{-1} \right) + \bar{\sigma}_u^2,$$

where $\rho_{\min}^+(Q_{-1}^\top \lambda \Sigma_\eta \lambda^\top Q_{-1})$ denotes the smallest positive eigenvalue of $Q_{-1}^\top \lambda \Sigma_\eta \lambda^\top Q_{-1}$. Therefore, Assumption 2(iii) holds provided that

$$\rho_{\min}^+ \left(Q_{-1}^\top \lambda \Sigma_\eta \lambda^\top Q_{-1} \right) \geq K^{-1}. \quad (\text{C.5})$$

For any $x \in \mathbb{R}^{n-1}$,

$$x^\top Q_{-1}^\top \lambda \Sigma_\eta \lambda^\top Q_{-1} x \geq \rho_{\min}(\Sigma_\eta) x^\top Q_{-1}^\top \lambda \lambda^\top Q_{-1} x.$$

Taking the minimum over all unit vectors $x \in \text{col}(Q_{-1}^\top \lambda)$ yields

$$\rho_{\min}^+ \left(Q_{-1}^\top \lambda \Sigma_\eta \lambda^\top Q_{-1} \right) \geq \rho_{\min}(\Sigma_\eta) \rho_{\min}^+ \left(Q_{-1}^\top \lambda \lambda^\top Q_{-1} \right), \quad (\text{C.6})$$

where $\rho_{\min}^+(Q_{-1}^\top \lambda \lambda^\top Q_{-1})$ denotes the smallest positive eigenvalue of $Q_{-1}^\top \lambda \lambda^\top Q_{-1}$. Since $\rho_{\min}(\Sigma_\eta) \geq K^{-1}$, it suffices to show that

$$\rho_{\min}^+ \left(Q_{-1}^\top \lambda \lambda^\top Q_{-1} \right) \geq K^{-1}. \quad (\text{C.7})$$

We next verify (C.7) in two separate cases.

Case (i). Suppose $\rho_{\min}((\mathbf{1}_n, \lambda)^\top (\mathbf{1}_n, \lambda)) \geq K^{-1}$. By (C.1),

$$\lambda^\top Q_{-1} Q_{-1}^\top \lambda = \lambda^\top M_{\mathbf{1}_n} \lambda.$$

By the block inverse formula, $(\lambda^\top M_{\mathbf{1}_n} \lambda)^{-1}$ is the lower-right $r \times r$ block of $((\mathbf{1}_n, \lambda)^\top (\mathbf{1}_n, \lambda))^{-1}$. Therefore,

$$\rho_{\min}(\lambda^\top M_{\mathbf{1}_n} \lambda) \geq \rho_{\min}((\mathbf{1}_n, \lambda)^\top (\mathbf{1}_n, \lambda)) \geq K^{-1},$$

which implies (C.7).

Case (ii). Suppose $\mathbf{1}_n \in \text{col}(\lambda)$ and $\rho_{\min}(\lambda^\top \lambda) \geq K^{-1}$. Define $\alpha_1 \equiv (\lambda^\top \lambda)^{-1} \lambda^\top \mathbf{1}_n$ and $\mathcal{A}_1 \equiv \{x \in \mathbb{R}^r : \mathbf{1}_n^\top \lambda x = 0\}$. Since $\mathbf{1}_n \in \text{col}(\lambda)$, we have $\lambda(\lambda^\top \lambda)^{-1} \lambda^\top \mathbf{1}_n = \mathbf{1}_n$ and

$$(\mathbf{1}_n^\top \lambda) \alpha_1 = \mathbf{1}_n^\top \lambda (\lambda^\top \lambda)^{-1} \lambda^\top \mathbf{1}_n = \mathbf{1}_n^\top \mathbf{1}_n = n, \quad (\text{C.8})$$

which implies that $\mathbf{1}_n^\top \lambda \neq 0$ and $\alpha_1 \notin \mathcal{A}_1$. Therefore, $\mathcal{A}_1$ is a linear subspace of dimension $r - 1$. Let $U \in \mathbb{R}^{r \times (r-1)}$ be a matrix whose columns form an orthonormal basis of $\mathcal{A}_1$. Then the matrix $C_\lambda \equiv (\alpha_1, U)$ is invertible and it satisfies

$$\lambda C_\lambda = (\mathbf{1}_n, \lambda_\perp), \quad \mathbf{1}_n^\top \lambda_\perp = 0, \quad (\text{C.9})$$

where $\lambda_\perp \equiv \lambda U$. Since the columns of U are orthonormal,

$$C_\lambda^\top C_\lambda = \begin{pmatrix} \|\alpha_1\|^2 & \alpha_1^\top U \\ U^\top \alpha_1 & \mathbf{I}_{r-1} \end{pmatrix}, \quad \text{where } c_1 \equiv U^\top \alpha_1. \quad (\text{C.10})$$

Using the block inverse formula, we obtain

$$(C_\lambda^\top C_\lambda)^{-1} = \begin{pmatrix} \delta_1^{-2} & -\delta_1^{-2} c_1^\top \\ -\delta_1^{-2} c_1 & \mathbf{I}_{r-1} + \delta_1^{-2} c_1 c_1^\top \end{pmatrix},$$

where $\delta_1 \equiv (\|\alpha_1\|^2 - \|c_1\|^2)^{1/2}$. Since $U^\top U = \mathbf{I}_{r-1}$,

$$\|c_1\|^2 = \|U^\top \alpha_1\|^2 = \|UU^\top \alpha_1\|^2.$$

Here $UU^\top \alpha_1 = U(U^\top U)^{-1} U^\top \alpha_1$ is the projection of α_1 onto $\mathcal{A}_1$, which is the subspace orthogonal to $\lambda^\top \mathbf{1}_n$. By the Pythagorean theorem,

$$\delta_1^2 = \|\alpha_1\|^2 - \|UU^\top \alpha_1\|^2$$

is the squared length of the projection of α_1 onto the subspace spanned by $\lambda^\top \mathbf{1}_n$. Therefore,

$$\delta_1^2 = \alpha_1^\top \lambda^\top \mathbf{1}_n (\mathbf{1}_n^\top \lambda \lambda^\top \mathbf{1}_n)^{-1} \mathbf{1}_n^\top \lambda \alpha_1 = \frac{(\mathbf{1}_n^\top \lambda (\lambda^\top \lambda)^{-1} \lambda^\top \mathbf{1}_n)^2}{\mathbf{1}_n^\top \lambda \lambda^\top \mathbf{1}_n} = \frac{n^2}{\mathbf{1}_n^\top \lambda \lambda^\top \mathbf{1}_n}. \quad (\text{C.11})$$

Here the last equality is by (C.8). Using (C.9) and (C.10), we have

$$\begin{aligned} Q_{-1}^\top \lambda \lambda^\top Q_{-1} &= Q_{-1}^\top \lambda C_\lambda (C_\lambda^\top C_\lambda)^{-1} C_\lambda^\top \lambda^\top Q_{-1} \\ &= (\mathbf{0}_n, Q_{-1}^\top \lambda_\perp) \begin{pmatrix} \delta_1^{-2} & -\delta_1^{-2} c_1^\top \\ -\delta_1^{-2} c_1 & \mathbf{I}_{r-1} + \delta_1^{-2} c_1 c_1^\top \end{pmatrix} (\mathbf{0}_n, Q_{-1}^\top \lambda_\perp)^\top \\ &= Q_{-1}^\top \lambda_\perp (\mathbf{I}_{r-1} + \delta_1^{-2} c_1 c_1^\top) \lambda_\perp^\top Q_{-1}. \end{aligned}$$

It then follows that

$$\rho_{\min}^+(Q_{-1}^\top \lambda \lambda^\top Q_{-1}) \geq \rho_{\min}^+(Q_{-1}^\top \lambda_\perp \lambda_\perp^\top Q_{-1}) \rho_{\min}(\mathbf{I}_{r-1} + \delta_1^{-2} c_1 c_1^\top) \geq \rho_{\min}^+(Q_{-1}^\top \lambda_\perp \lambda_\perp^\top Q_{-1}). \quad (\text{C.12})$$

Since the positive eigenvalues of $Q_{-1}^\top \lambda_\perp \lambda_\perp^\top Q_{-1}$ are the same as those of $\lambda_\perp^\top Q_{-1} Q_{-1}^\top \lambda_\perp$, and the latter satisfies

$$\lambda_\perp^\top Q_{-1} Q_{-1}^\top \lambda_\perp = \lambda_\perp^\top M_{\mathbf{1}_n} \lambda_\perp = \lambda_\perp^\top \lambda_\perp$$

by (C.1), it follows that the (C.7) holds if

$$\rho_{\min}(\lambda_\perp^\top \lambda_\perp) \geq K^{-1}. \quad (\text{C.13})$$

We next verify (C.13).

From the first equation in (C.9), we have

$$C_\lambda^\top \lambda^\top \lambda C_\lambda = (\lambda C_\lambda)^\top (\lambda C_\lambda) = \begin{pmatrix} n & 0 \\ 0 & \lambda_\perp^\top \lambda_\perp \end{pmatrix},$$

which implies that

$$\rho_{\min}(\lambda_\perp^\top \lambda_\perp) \geq \rho_{\min}(C_\lambda^\top \lambda^\top \lambda C_\lambda) \geq \rho_{\min}(C_\lambda^\top C_\lambda) \rho_{\min}(\lambda^\top \lambda) \geq K^{-1} \rho_{\min}(C_\lambda^\top C_\lambda). \quad (\text{C.14})$$

Let $x = (x_1, x_2^\top)^\top \in \mathbb{R}^r$, where $x_1 \in \mathbb{R}$ and $x_2 \in \mathbb{R}^{r-1}$. Then

$$\begin{aligned} x^\top (C_\lambda^\top C_\lambda)^{-1} x &= x_1^2 \delta_1^{-2} - 2\delta_1^{-2} x_1 x_2^\top c_1 + x_2^\top x_2 + \delta_1^{-2} x_2^\top c_1 c_1^\top x_2 \\ &\leq \delta_1^{-2} (|x_1| + \|x_2\| \|c_1\|)^2 + \|x_2\|^2 \\ &\leq (1 + \delta_1^{-2} (1 + \|c_1\|)^2) \|x\|^2, \end{aligned}$$

which together with $\|c_1\| \leq \|\alpha_1\|$ implies that

$$\rho_{\min}(C_\lambda^\top C_\lambda) = (\rho_{\max}((C_\lambda^\top C_\lambda)^{-1}))^{-1} \geq \frac{1}{1 + \delta_1^{-2} (1 + \|\alpha_1\|)^2}. \quad (\text{C.15})$$

From (C.11) and $\|\alpha_1\|^2 = \mathbf{1}_n^\top \lambda (\lambda^\top \lambda)^{-2} \lambda^\top \mathbf{1}_n$, it follows that

$$\begin{aligned} \delta_1^{-2} (1 + \|\alpha_1\|)^2 &\leq 2\delta_1^{-2} (1 + \|\alpha_1\|^2) = \frac{2(\mathbf{1}_n^\top \lambda \lambda^\top \mathbf{1}_n)(1 + \mathbf{1}_n^\top \lambda (\lambda^\top \lambda)^{-2} \lambda^\top \mathbf{1}_n)}{n^2} \\ &\leq 2\rho_{\max}(\lambda^\top \lambda) \frac{\mathbf{1}_n^\top \lambda (\lambda^\top \lambda)^{-1} \lambda^\top \mathbf{1}_n}{n^2} \left(1 + \frac{\mathbf{1}_n^\top \lambda (\lambda^\top \lambda)^{-1} \lambda^\top \mathbf{1}_n}{\rho_{\min}(\lambda^\top \lambda)}\right) \\ &= \frac{2\rho_{\max}(\lambda^\top \lambda)}{\rho_{\min}(\lambda^\top \lambda)} \left(1 + \frac{\rho_{\min}(\lambda^\top \lambda)}{n}\right) \leq K, \end{aligned} \quad (\text{C.16})$$

where the second inequality follows from

$$\mathbf{1}_n^\top \lambda \lambda^\top \mathbf{1}_n = \mathbf{1}_n^\top \lambda (\lambda^\top \lambda)^{-1/2} (\lambda^\top \lambda) (\lambda^\top \lambda)^{-1/2} \lambda^\top \mathbf{1}_n \leq \rho_{\max}(\lambda^\top \lambda) \mathbf{1}_n^\top \lambda (\lambda^\top \lambda)^{-1} \lambda^\top \mathbf{1}_n$$

and

$$\begin{aligned}\mathbf{1}_n^\top \lambda (\lambda^\top \lambda)^{-2} \lambda^\top \mathbf{1}_n &= \mathbf{1}_n^\top \lambda (\lambda^\top \lambda)^{-1/2} (\lambda^\top \lambda)^{-1} (\lambda^\top \lambda)^{-1/2} \lambda^\top \mathbf{1}_n \\ &\leq \rho_{\max}((\lambda^\top \lambda)^{-1}) \mathbf{1}_n^\top \lambda (\lambda^\top \lambda)^{-1} \lambda^\top \mathbf{1}_n = \frac{\mathbf{1}_n^\top \lambda (\lambda^\top \lambda)^{-1} \lambda^\top \mathbf{1}_n}{\rho_{\min}(\lambda^\top \lambda)},\end{aligned}$$

and the second equality follows from (C.8). The desired result in (C.13) then follows from (C.14), (C.15) and (C.16). *Q.E.D.*

Lemma C.2. *Suppose that $|\psi| \geq K^{-1}$, $0 < |1 - \phi\psi| \leq K$, S_t is independent of u_t and η_t , and either (i) $\rho_{\min}((\mathbf{1}_n, \lambda)^\top (\mathbf{1}_n, \lambda)) > 0$ and $S_u^\top M_{(\mathbf{1}_n, \lambda)} S_u \geq K^{-1}$; or (ii) $\mathbf{1}_n \in \text{col}(\lambda)$, $\rho_{\min}(\lambda^\top \lambda) > 0$ and $S_u^\top M_\lambda S_u \geq K^{-1}$, where $S_u \equiv T^{-1} \sum_{t \leq T} \sigma_{u,t}^2 \mathbb{E}[S_t]$. Then Assumption 4(iii) holds.*

PROOF OF LEMMA C.2. Since $A \equiv Q_{-1} A_0$ spans the same subspace as $\bar{\lambda}_\perp$, which is orthogonal to $(\mathbf{1}_n, \lambda)$, we have

$$A^\top \mathbf{1}_n = \mathbf{0}_{n-\bar{r}} \quad \text{and} \quad A^\top \lambda = \mathbf{0}_{(n-\bar{r}) \times r}. \quad (\text{C.17})$$

This, together with the reduced form expression of y_t in (C.25), implies that

$$A^\top y_t = A^\top (u_t + \lambda \eta_t) = A^\top u_t. \quad (\text{C.18})$$

Using (C.17), the reduced-form expression of p_t in (C.26), the expression for $A^\top y_t$ in (C.18), and Assumption 1, we obtain

$$\mathbb{E}[A^\top y_t p_t] = \frac{\psi}{1 - \phi\psi} A^\top \mathbb{E}[u_t (u_t + \lambda \eta_t)^\top S_t] + \frac{1}{1 - \phi\psi} A^\top \mathbb{E}[u_t \varepsilon_t] = \frac{\psi \sigma_{u,t}^2}{1 - \phi\psi} A^\top \mathbb{E}[S_t].$$

Therefore,

$$\left\| T^{-1} \sum_{t \leq T} \mathbb{E}[A^\top y_t p_t] \right\|^2 = \frac{\psi^2}{(1 - \phi\psi)^2} \|A^\top S_u\|^2. \quad (\text{C.19})$$

Next by the reduced form expression of y_t in (C.25)

$$y_{S,t} \equiv S_t^\top y_t = \frac{1}{1 - \phi\psi} S_t^\top (u_t + \lambda \eta_t) + \frac{\phi}{1 - \phi\psi} \varepsilon_t,$$

which, together with Assumption 1, (C.17) and (C.18), implies that

$$\mathbb{E}[A^\top y_t y_{S,t}] = \frac{1}{1 - \phi\psi} A^\top \mathbb{E}[u_t (u_t + \lambda \eta_t)^\top S_t] + \frac{\phi}{1 - \phi\psi} A^\top \mathbb{E}[u_t \varepsilon_t] = \frac{\sigma_{u,t}^2}{1 - \phi\psi} A^\top \mathbb{E}[S_t].$$

Therefore,

$$\left\| T^{-1} \sum_{t \leq T} \mathbb{E}[A^\top y_t y_{S,t}] \right\|^2 = \frac{1}{(1 - \phi\psi)^2} \|A^\top S_u\|^2. \quad (\text{C.20})$$

Since $A^\top A = \mathbf{I}_{n-\bar{r}}$, we have $\|A^\top S_u\| = \|AA^\top S_u\|$, so $\|A^\top S_u\|$ is the length of the projection of S_u onto $\text{col}(A)$. Because $\text{col}(A)$ is orthogonal to $\text{col}(\mathbf{1}_n, \lambda)$, it follows that

$$\|A^\top S_u\|^2 = \begin{cases} S_u^\top M_{(\mathbf{1}_n, \lambda)} S_u, & \text{if } \rho_{\min}((\mathbf{1}_n, \lambda)^\top (\mathbf{1}_n, \lambda)) > 0 \\ S_u^\top M_\lambda S_u, & \text{if } \mathbf{1}_n \in \text{col}(\lambda) \text{ and } \rho_{\min}(\lambda^\top \lambda) > 0 \end{cases}. \quad (\text{C.21})$$

Combining (C.19), (C.20) and (C.21), as well as the maintained conditions of the lemma gives

$$\left\| T^{-1} \sum_{t \leq T} \mathbb{E}[A^\top y_t p_t] \right\| \geq K^{-1} \quad \text{and} \quad \left\| T^{-1} \sum_{t \leq T} \mathbb{E}[A^\top y_t y_{S,t}] \right\| \geq K^{-1}$$

This verifies Assumption 4(iii).

Q.E.D.

Lemma C.3. *For any $n \times (n - \bar{r})$ full rank matrix A_1 , any $2n \times 2n$ positive definite matrix W_1 , and any nonsingular $(n - \bar{r}) \times (n - \bar{r})$ matrix C_1 , we have*

$$J_T(\theta; A_1, W_1) = J_T(\theta; A_1 C_1, W_1),$$

where

$$J_T(\theta; A_1, W_1) \equiv \bar{g}_T(\theta; A_1)^\top ((\mathbf{I}_2 \otimes A_1^\top) W_1 (\mathbf{I}_2 \otimes A_1))^{-1} \bar{g}_T(\theta; A_1), \quad (\text{C.22})$$

and $J_T(\theta; A_1 C_1, W_1)$ is defined analogously to $J_T(\theta; A_1, W_1)$ with A_1 replaced by $A_1 C_1$.

PROOF OF LEMMA C.3. By definition of the GMM criterion,

$$J_T(\theta; A_1 C_1, W_1) = \bar{g}_{0,T}(\theta)^\top (\mathbf{I}_2 \otimes A_1 C_1) ((\mathbf{I}_2 \otimes C_1^\top A_1^\top) W_1 (\mathbf{I}_2 \otimes A_1 C_1))^{-1} (\mathbf{I}_2 \otimes C_1^\top A_1^\top) \bar{g}_{0,T}(\theta) \quad (\text{C.23})$$

where

$$\bar{g}_{0,T}(\theta) \equiv \begin{pmatrix} T^{-1} \sum_{t \leq T} y_t (y_{e,t} - \phi p_t) \\ T^{-1} \sum_{t \leq T} y_t (p_t - \psi y_{S,t}) \end{pmatrix}.$$

Using the identity

$$(\mathbf{I}_2 \otimes C_1^\top A_1^\top) = (\mathbf{I}_2 \otimes C_1^\top) (\mathbf{I}_2 \otimes A_1^\top) \text{ and } (\mathbf{I}_2 \otimes A_1 C_1) = (\mathbf{I}_2 \otimes A_1) (\mathbf{I}_2 \otimes C_1),$$

we obtain

$$((\mathbf{I}_2 \otimes C_1^\top A_1^\top) W_1 (\mathbf{I}_2 \otimes A_1 C_1))^{-1} = (\mathbf{I}_2 \otimes C_1)^{-1} ((\mathbf{I}_2 \otimes A_1^\top) W_1 (\mathbf{I}_2 \otimes A_1))^{-1} (\mathbf{I}_2 \otimes C_1^\top)^{-1}. \quad (\text{C.24})$$

Substituting (C.24) into (C.23), we get

$$\begin{aligned} J_T(\theta; A_1 C_1, W_1) &= \bar{g}_{0,T}(\theta)^\top (\mathbf{I}_2 \otimes A_1) (\mathbf{I}_2 \otimes C_1) \\ &\quad \times (\mathbf{I}_2 \otimes C_1)^{-1} ((\mathbf{I}_2 \otimes A_1^\top) W_1 (\mathbf{I}_2 \otimes A_1))^{-1} (\mathbf{I}_2 \otimes C_1^\top)^{-1} \\ &\quad \times (\mathbf{I}_2 \otimes C_1^\top) (\mathbf{I}_2 \otimes A_1^\top) \bar{g}_{0,T}(\theta) \\ &= \bar{g}_{0,T}(\theta)^\top (\mathbf{I}_2 \otimes A_1) ((\mathbf{I}_2 \otimes A_1^\top) W_1 (\mathbf{I}_2 \otimes A_1))^{-1} (\mathbf{I}_2 \otimes A_1^\top) \bar{g}_{0,T}(\theta) \\ &= J_T(\theta; A_1, W_1), \end{aligned}$$

which establishes the claim of the lemma.

Q.E.D.

Lemma C.4. *Under Assumptions 2(i, ii) and 3, we have:*

$$T^{-1/2} \sum_{t \leq T} (y_t y_t^\top - \mathbb{E}[y_t y_t^\top]) = O_p(1) \quad \text{and} \quad T^{-1/2} \sum_{t \leq T} (p_t y_t - \mathbb{E}[p_t y_t]) = O_p(1).$$

PROOF OF LEMMA C.4. From (2) and (19), we obtain

$$y_t = \left(\mathbf{I}_n + \frac{\phi\psi}{1-\phi\psi} \mathbf{1}_n S_t^\top \right) (u_t + \lambda\eta_t) + \frac{\phi}{1-\phi\psi} \mathbf{1}_n \varepsilon_t, \quad (\text{C.25})$$

$$p_t = \frac{\psi}{1-\phi\psi} S_t^\top (u_t + \lambda\eta_t) + \frac{\varepsilon_t}{1-\phi\psi}. \quad (\text{C.26})$$

Therefore,

$$\begin{aligned} T^{-1} \sum_{t \leq T} y_t y_t^\top &= T^{-1} \sum_{t \leq T} \left(\mathbf{I}_n + \frac{\phi\psi}{1-\phi\psi} \mathbf{1}_n S_t^\top \right) (u_t + \lambda\eta_t) (u_t + \lambda\eta_t)^\top \left(\mathbf{I}_n + \frac{\phi\psi}{1-\phi\psi} S_t \mathbf{1}_n^\top \right) \\ &\quad + \frac{\phi}{1-\phi\psi} T^{-1} \sum_{t \leq T} \left(\mathbf{I}_n + \frac{\phi\psi}{1-\phi\psi} \mathbf{1}_n S_t^\top \right) (u_t + \lambda\eta_t) \varepsilon_t \mathbf{1}_n^\top \\ &\quad + \frac{\phi}{1-\phi\psi} \mathbf{1}_n T^{-1} \sum_{t \leq T} \varepsilon_t (u_t + \lambda\eta_t)^\top \left(\mathbf{I}_n + \frac{\phi\psi}{1-\phi\psi} S_t \mathbf{1}_n^\top \right) \\ &\quad + \frac{\phi^2}{(1-\phi\psi)^2} \mathbf{1}_n \mathbf{1}_n^\top T^{-1} \sum_{t \leq T} \varepsilon_t^2, \end{aligned} \quad (\text{C.27})$$

and

$$\begin{aligned} T^{-1} \sum_{t \leq T} p_t y_t &= \frac{\psi}{1-\phi\psi} T^{-1} \sum_{t \leq T} \left(\mathbf{I}_n + \frac{\phi\psi}{1-\phi\psi} \mathbf{1}_n S_t^\top \right) (u_t + \lambda\eta_t) (u_t + \lambda\eta_t)^\top S_t \\ &\quad + \frac{1}{1-\phi\psi} T^{-1} \sum_{t \leq T} \left(\mathbf{I}_n + \frac{\phi\psi}{1-\phi\psi} \mathbf{1}_n S_t^\top \right) (u_t + \lambda\eta_t) \varepsilon_t \\ &\quad + \frac{\phi\psi}{(1-\phi\psi)^2} \mathbf{1}_n T^{-1} \sum_{t \leq T} \varepsilon_t (u_t + \lambda\eta_t)^\top S_t + \frac{\phi}{(1-\phi\psi)^2} \mathbf{1}_n T^{-1} \sum_{t \leq T} \varepsilon_t^2. \end{aligned} \quad (\text{C.28})$$

From Assumptions 2(i, ii), we have

$$T^{-1} \sum_{t \leq T} (u_t + \lambda\eta_t) (u_t + \lambda\eta_t)^\top = T^{-1} \sum_{t \leq T} \mathbb{E}[(u_t + \lambda\eta_t) (u_t + \lambda\eta_t)^\top] + O_p(T^{-1/2}), \quad (\text{C.29})$$

$$T^{-1} \sum_{t \leq T} (u_t + \lambda\eta_t) \varepsilon_t = T^{-1} \sum_{t \leq T} \mathbb{E}[(u_t + \lambda\eta_t) \varepsilon_t] + O_p(T^{-1/2}), \quad (\text{C.30})$$

$$T^{-1} \sum_{t \leq T} \varepsilon_t^2 = T^{-1} \sum_{t \leq T} \mathbb{E}[\varepsilon_t^2] + O_p(T^{-1/2}). \quad (\text{C.31})$$

Similarly, from Assumptions 2(i) and 3(i), it follows that

$$T^{-1} \sum_{t \leq T} S_t^\top (u_t + \lambda\eta_t) (u_t + \lambda\eta_t)^\top S_t = T^{-1} \sum_{t \leq T} \mathbb{E}[S_t^\top (u_t + \lambda\eta_t) (u_t + \lambda\eta_t)^\top S_t] + O_p(T^{-1/2}), \quad (\text{C.32})$$

$$T^{-1} \sum_{t \leq T} S_t^\top (u_t + \lambda \eta_t)(u_t + \lambda \eta_t)^\top = T^{-1} \sum_{t \leq T} \mathbb{E}[S_t^\top (u_t + \lambda \eta_t)(u_t + \lambda \eta_t)^\top] + O_p(T^{-1/2}), \quad (\text{C.33})$$

$$T^{-1} \sum_{t \leq T} S_t^\top (u_t + \lambda \eta_t) \varepsilon_t = T^{-1} \sum_{t \leq T} \mathbb{E}[S_t^\top (u_t + \lambda \eta_t) \varepsilon_t] + O_p(T^{-1/2}). \quad (\text{C.34})$$

The claim of the lemma follows from Assumption 2(i), and (C.27)-(C.34). *Q.E.D.*

Lemma C.5. *Under Assumptions 2(i, ii) and 3, we have:*

$$T^{-1/2} \sum_{t \leq T} (y_{S,t} y_t - \mathbb{E}[y_{S,t} y_t]) = O_p(1).$$

PROOF OF LEMMA C.5. Since $\mathbf{1}_n^\top S_t = 1$, from (C.25) we have

$$y_{S,t} = \frac{1}{1 - \phi\psi} S_t^\top (u_t + \lambda \eta_t) + \frac{\phi}{1 - \phi\psi} \varepsilon_t.$$

Therefore,

$$\begin{aligned} y_{S,t} y_t &= \frac{1}{1 - \phi\psi} \left(\mathbf{I}_n + \frac{\phi\psi}{1 - \phi\psi} \mathbf{1}_n S_t^\top \right) (u_t + \lambda \eta_t)(u_t + \lambda \eta_t)^\top S_t \\ &\quad + \frac{\phi}{1 - \phi\psi} \left(\mathbf{I}_n + \frac{\phi\psi}{1 - \phi\psi} \mathbf{1}_n S_t^\top \right) (u_t + \lambda \eta_t) \varepsilon_t \\ &\quad + \frac{\phi}{(1 - \phi\psi)^2} S_t^\top (u_t + \lambda \eta_t) \varepsilon_t \mathbf{1}_n + \frac{\phi^2}{(1 - \phi\psi)^2} \mathbf{1}_n \varepsilon_t^2. \end{aligned}$$

The claim of the lemma thus follows from (C.30)-(C.34). *Q.E.D.*

The GIV estimator depends on $\hat{A} \equiv Q_{-1} \hat{A}_0$, which serves as an estimator of $Q_{-1} A_0$. Since A_0 collects the eigenvectors associated with the smallest $n - \bar{r}$ eigenvalues of S_y , it is not uniquely identified when $n - \bar{r} > 1$. Therefore, it is generally unrealistic to expect that $\hat{A}_0$ (and hence $\hat{A}$) converges to A_0 as $T \rightarrow \infty$. Instead, the lemma below shows that a suitably rotated version of $\hat{A}_0$ is a consistent estimator of A_0 .

Lemma C.6. *Define $K_0 \equiv (A_{0,\perp}^\top \hat{A}_0)(A_0^\top \hat{A}_0)^{-1}$ and*

$$\hat{H}_0 \equiv \left(\mathbf{I}_{n-\bar{r}} + K_0^\top K_0 \right)^{-1/2} (A_0^\top \hat{A}_0 + K_0^\top A_{0,\perp}^\top \hat{A}_0). \quad (\text{C.35})$$

Under Assumptions 2 and 3, we have

$$\hat{A}_0 \hat{H}_0^\top - A_0 = A_{0,\perp} (\sigma_u^2 \mathbf{I}_{\bar{r}-1} - \Lambda_\perp)^{-1} A_{0,\perp}^\top \hat{\Delta}_y A_0 + O_p(T^{-1}), \quad (\text{C.36})$$

where $\hat{\Delta}_y \equiv \hat{S}_y - S_y$. Moreover, $\hat{H}_0$ satisfies $\hat{H}_0^\top \hat{H}_0 = \mathbf{I}_{n-\bar{r}}$ wpa1.

PROOF OF LEMMA C.6. The proof follows directly from Lemmas C.7 - C.10. *Q.E.D.*

Lemma C.7. *Under Assumptions 2 and 3, we have*

$$\|\hat{A}_0^\top A_0\|_o \geq 1 - O_p(T^{-1/2}) \quad \text{and} \quad \|K_0\| = O_p(T^{-1/2}), \quad (\text{C.37})$$

where $\|\cdot\|_o$ denotes the matrix operator norm.

PROOF OF LEMMA C.7. By Lemma C.4,

$$\hat{\Delta}_y \equiv \hat{S}_y - S_y = O_p(T^{-1/2}). \quad (\text{C.38})$$

Recall that $\hat{A}_0$ collects the orthonormal eigenvectors associated with the smallest $n - \bar{r}$ eigenvalues of $\hat{S}_y$. We next apply the Davis–Kahan $\sin \Theta$ theorem by verifying the conditions of Theorem 2 in Yu, Wang, and Samworth (2015). Specifically, we aim to establish that

$$\|A_{0,\perp}^\top \hat{A}_0\| \leq \frac{2 \min((n - \bar{r})^{1/2} \|\hat{\Delta}_y\|_o, \|\hat{\Delta}_y\|)}{\mu_{n-\bar{r}+1} - \bar{\sigma}_u^2}. \quad (\text{C.39})$$

Let $\Theta(A_0, \hat{A}_0)$ be the $(n - \bar{r}) \times (n - \bar{r})$ diagonal matrix of principal angles between the true and estimated eigenspaces, with j -th diagonal entry $\arccos(\sigma_j)$, where $\sigma_1 \geq \dots \geq \sigma_{n-\bar{r}}$ are the singular values of $A_0^\top \hat{A}_0$. Define $\sin(\Theta(A_0, \hat{A}_0))$ componentwise. Set the parameters of Theorem 2 in Yu, Wang, and Samworth (2015) as

$$p = n - 1, \quad r = \bar{r}, \quad s = n - 1 \quad \text{and} \quad d = s - r + 1 = n - \bar{r}.$$

Let $\tilde{\mu}_j \equiv \mu_{n-j}$ for $j = 1, \dots, n - 1$, and $\tilde{\mu}_n = -\infty$, such that $\tilde{\mu}_j \geq \tilde{\mu}_{j+1}$. By Assumption 2(iii),

$$\min\{\tilde{\mu}_{r-1} - \tilde{\mu}_r, \tilde{\mu}_s - \tilde{\mu}_{s+1}\} = \min\{\tilde{\mu}_{\bar{r}-1} - \tilde{\mu}_{\bar{r}}, \tilde{\mu}_{n-1} - \tilde{\mu}_n\} = \mu_{n-\bar{r}+1} - \mu_{n-\bar{r}} = \mu_{n-\bar{r}+1} - \bar{\sigma}_u^2 > K^{-1}. \quad (\text{C.40})$$

Since A_0 and $\hat{A}_0$ consist of the eigenvectors corresponding to the smallest $n - \bar{r}$ eigenvalues of S_y and $\hat{S}_y$, respectively, Theorem 2 of Yu, Wang, and Samworth (2015) yields

$$\|\sin(\Theta(A_0, \hat{A}_0))\| \leq \frac{2 \min\left\{d^{1/2} \|\hat{\Delta}_y\|_o, \|\hat{\Delta}_y\|\right\}}{\mu_{n-\bar{r}+1} - \bar{\sigma}_u^2}. \quad (\text{C.41})$$

Finally, because $\|\sin(\Theta(A_0, \hat{A}_0))\| = \|A_{0,\perp}^\top \hat{A}_0\|$ (see the expression below (A4) in Yu, Wang, and Samworth (2015)), the desired bound in (C.39) follows directly from (C.41).

Using (C.39) along with (C.38) and Assumption 2(iii), we obtain

$$\|A_{0,\perp}^\top \hat{A}_0\| = O_p(T^{-1/2}). \quad (\text{C.42})$$

Let $P_0 \equiv (A_0, A_{0,\perp})$. Since P_0 and $\hat{A}_0$ are orthonormal by construction,

$$(P_0^\top \hat{A}_0)^\top (P_0^\top \hat{A}_0) = \mathbf{I}_{n-\bar{r}}.$$

Expanding the left-hand side,

$$\mathbf{I}_{n-\bar{r}} = ((A_0, A_{0,\perp})^\top \hat{A}_0)^\top (A_0, A_{0,\perp})^\top \hat{A}_0 = \hat{A}_0^\top A_0 A_0^\top \hat{A}_0 + \hat{A}_0^\top A_{0,\perp} A_{0,\perp}^\top \hat{A}_0. \quad (\text{C.43})$$

Taking operator norms and applying (C.42),

$$\|\hat{A}_0^\top A_0\|_o^2 \geq 1 - \|\hat{A}_0^\top A_{0,\perp}\|_o^2 \geq 1 - \|\hat{A}_0^\top A_{0,\perp}\|^2 = 1 - O_p(T^{-1}). \quad (\text{C.44})$$

This proves the first claim of (C.37).

From (C.43) and (C.44), $\hat{A}_0^\top A_0$ is nonsingular wpa1. Hence $K_0 \equiv (A_{0,\perp}^\top \hat{A}_0)(A_0^\top \hat{A}_0)^{-1}$ is well defined wpa1. This, together with (C.42) and (C.44) shows

$$\|K_0\| = O_p(T^{-1/2}),$$

which establishes the second claim. Q.E.D.

Lemma C.8. *Under Assumptions 2 and 3, we have wpa1,*

$$\hat{A}_0 \hat{H}_0^\top = P_0 \begin{pmatrix} \mathbf{I}_{n-\bar{r}} \\ K_0 \end{pmatrix} (\mathbf{I}_{n-\bar{r}} + K_0^\top K_0)^{-1/2} \quad \text{and} \quad \hat{H}_0^\top \hat{H}_0 = \mathbf{I}_{n-\bar{r}}, \quad (\text{C.45})$$

where $P_0 \equiv (A_0, A_{0,\perp})$.

PROOF OF LEMMA C.8. Under Assumption 2, we can use (C.42) and (C.43) in the proof of Lemma C.7 to show that

$$\|\mathbf{I}_{n-\bar{r}} - \hat{A}_0^\top A_0 A_0^\top \hat{A}_0\| = O_p(T^{-1}), \quad (\text{C.46})$$

which implies that $A_0^\top \hat{A}_0$ is invertible wpa1. Conditioning on the event that $A_0^\top \hat{A}_0$ is invertible, we may decompose $P_0^\top \hat{A}_0 = (A_0, A_{0,\perp})^\top \hat{A}_0$ as

$$P_0^\top \hat{A}_0 = \begin{pmatrix} A_0^\top \hat{A}_0 \\ A_{0,\perp}^\top \hat{A}_0 \end{pmatrix} = \begin{pmatrix} \mathbf{I}_{n-\bar{r}} \\ (A_{0,\perp}^\top \hat{A}_0)(A_0^\top \hat{A}_0)^{-1} \end{pmatrix} A_0^\top \hat{A}_0 = \begin{pmatrix} \mathbf{I}_{n-\bar{r}} \\ K_0 \end{pmatrix} A_0^\top \hat{A}_0. \quad (\text{C.47})$$

Because both P_0 and $\hat{A}_0$ have orthonormal columns,

$$(P_0^\top \hat{A}_0)^\top (P_0^\top \hat{A}_0) = \hat{A}_0^\top P_0 P_0^\top \hat{A}_0 = \mathbf{I}_{n-\bar{r}},$$

so $P_0^\top \hat{A}_0$ itself has orthonormal columns. Define

$$U_0 \equiv \begin{pmatrix} \mathbf{I}_{n-\bar{r}} \\ K_0 \end{pmatrix} (\mathbf{I}_{n-\bar{r}} + K_0^\top K_0)^{-1/2}. \quad (\text{C.48})$$

Then $U_0^\top U_0 = \mathbf{I}_{n-\bar{r}}$, so U_0 also has orthonormal columns.

Since $A_0^\top \hat{A}_0$ and $\mathbf{I}_{n-\bar{r}} + K_0^\top K_0$ are nonsingular, (C.47) and (C.48) imply that

$$\text{col}(P_0^\top \hat{A}_0) = \text{col}(U_0), \text{ wpa1.} \quad (\text{C.49})$$

Hence there exists a square matrix R_0 such that

$$P_0^\top \hat{A}_0 = U_0 R_0. \quad (\text{C.50})$$

Because both $P_0^\top \hat{A}_0$ and U_0 have orthonormal columns, and R_0 is a square matrix, we have

$$R_0^\top R_0 = R_0 R_0^\top = \mathbf{I}_{n-\bar{r}}. \quad (\text{C.51})$$

It follows that:

$$(P_0^\top \hat{A}_0)(P_0^\top \hat{A}_0)^\top = (U_0 R_0)(U_0 R_0)^\top = U_0 R_0 R_0^\top U_0^\top = U_0 U_0^\top. \quad (\text{C.52})$$

By definition of $\hat{H}_0$ (see Lemma C.6),

$$\hat{H}_0 = U_0^\top (P_0^\top \hat{A}_0), \quad (\text{C.53})$$

based on which we obtain

$$\hat{A}_0 \hat{H}_0^\top = \hat{A}_0 \hat{A}_0^\top P_0 U_0 = P_0 (P_0^\top \hat{A}_0) (P_0^\top \hat{A}_0)^\top U_0 = P_0 U_0 U_0^\top U_0 = P_0 U_0, \quad (\text{C.54})$$

where we use $P_0 P_0^\top = I_{n-1}$, (C.52) and $U_0^\top U_0 = \mathbf{I}_{n-\bar{r}}$. This proves the first claim of the lemma.

For the second claim,

$$\hat{H}_0^\top \hat{H}_0 = (P_0^\top \hat{A}_0)^\top U_0 U_0^\top (P_0^\top \hat{A}_0) = (U_0 R_0)^\top U_0 U_0^\top (U_0 R_0) = R_0^\top (U_0^\top U_0)^2 R_0 = \mathbf{I}_{n-\bar{r}},$$

using $U_0^\top U_0 = \mathbf{I}_{n-\bar{r}}$, (C.50), (C.51) and (C.53). Both statements in the lemma now follow. *Q.E.D.*

Lemma C.9. *Under Assumptions 2 and 3, we have*

$$\tilde{A}_0 - A_0 = A_{0,\perp} K_0 + O_p(T^{-1}), \quad (\text{C.55})$$

where

$$\tilde{A}_0 \equiv P_0 \begin{pmatrix} \mathbf{I}_{n-\bar{r}} \\ K_0 \end{pmatrix} (\mathbf{I}_{n-\bar{r}} + K_0^\top K_0)^{-1/2}. \quad (\text{C.56})$$

PROOF OF LEMMA C.9. The proof uses $K_0 = (A_{0,\perp}^\top \hat{A}_0)(A_0^\top \hat{A}_0)^{-1}$, which holds wpa1 by (C.46) in the proof of Lemma C.8. From the definition of $\tilde{A}_0$,

$$\begin{aligned} \tilde{A}_0 - A_0 &= (A_0 + A_{0,\perp} K_0)(\mathbf{I}_{n-\bar{r}} + K_0^\top K_0)^{-1/2} - A_0 \\ &= A_{0,\perp} K_0 + (A_0 + A_{0,\perp} K_0) \left((\mathbf{I}_{n-\bar{r}} + K_0^\top K_0)^{-1/2} - \mathbf{I}_{n-\bar{r}} \right). \end{aligned} \quad (\text{C.57})$$

By the perturbation inequality for positive definite matrices (see, e.g., Lemma A.2 in Belloni, Chernozhukov, Chetverikov, and Kato (2015), applied with $A = \mathbf{I}_{n-\bar{r}} + K_0^\top K_0$ and $B = \mathbf{I}_{n-\bar{r}}$ in that lemma),

$$\|(\mathbf{I}_{n-\bar{r}} + K_0^\top K_0)^{1/2} - \mathbf{I}_{n-\bar{r}}\|_o \leq \|K_0^\top K_0\|_o. \quad (\text{C.58})$$

Since $\|K_0^\top K_0\|_o = O_p(T^{-1})$ by Lemma C.7, we obtain

$$\|(\mathbf{I}_{n-\bar{r}} + K_0^\top K_0)^{1/2} - \mathbf{I}_{n-\bar{r}}\|_o \leq \|K_0^\top K_0\|_o = O_p(T^{-1}). \quad (\text{C.59})$$

Next, Lemma C.7 along with the triangle inequality and (C.59) implies

$$\begin{aligned} \|(A_0 + A_{0,\perp} K_0)(\mathbf{I}_{n-\bar{r}} + K_0^\top K_0)^{-1/2}\|_o &\leq (\rho_{\max}((\mathbf{I}_{n-\bar{r}} + K_0^\top K_0)^{-1}))^{1/2} \|A_0 + A_{0,\perp} K_0\|_o \\ &\leq \|A_0\|_o + \|A_{0,\perp}\|_o \|K_0\|_o = 1 + O_p(T^{-1/2}) \leq 2, \end{aligned} \quad (\text{C.60})$$

wpa1. Combining (C.59) and (C.60) yields

$$\begin{aligned} &\|(A_0 + A_{0,\perp} K_0)((\mathbf{I}_{n-\bar{r}} + K_0^\top K_0)^{-1/2} - \mathbf{I}_{n-\bar{r}})\|_o \\ &\leq \|(A_0 + A_{0,\perp} K_0)(\mathbf{I}_{n-\bar{r}} + K_0^\top K_0)^{-1/2}\|_o \|(\mathbf{I}_{n-\bar{r}} + K_0^\top K_0)^{1/2} - \mathbf{I}_{n-\bar{r}}\|_o = O_p(T^{-1}). \end{aligned} \quad (\text{C.61})$$

The statement of the lemma follows from (C.57) and (C.61). Q.E.D.

Lemma C.10. *Under Assumptions 2 and 3, we have*

$$K_0 = (\bar{\sigma}_u^2 \mathbf{I}_{\bar{r}-1} - \Lambda_\perp)^{-1} A_{0,\perp}^\top \hat{\Delta}_y A_0 + O_p(T^{-1}).$$

PROOF OF LEMMA C.10. The proof uses $(A_0^\top \hat{A}_0)^- = (A_0^\top \hat{A}_0)^{-1}$, which holds wpa1 by (C.46) in the proof of Lemma C.8. Let $\hat{\Lambda}$ be the diagonal matrix of the $n - \bar{r}$ smallest eigenvalues of $\hat{S}_y$. By definition of $\hat{A}_0$, $\hat{S}_y \hat{A}_0 = \hat{A}_0 \hat{\Lambda}$. Premultiplying by $P_0^\top$ and using $\hat{S}_y = S_y + \hat{\Delta}_y$

$$P_0^\top \hat{A}_0 \hat{\Lambda} = P_0^\top \hat{S}_y \hat{A}_0 = P_0^\top \hat{S}_y P_0 P_0^\top \hat{A}_0 = (P_0^\top S_y P_0 + P_0^\top \hat{\Delta}_y P_0) P_0^\top \hat{A}_0. \quad (\text{C.62})$$

Since

$$P_0^\top \hat{A}_0 = \begin{pmatrix} A_0^\top \hat{A}_0 \\ A_{0,\perp}^\top \hat{A}_0 \end{pmatrix}, \quad P_0^\top S_y P_0 = \begin{pmatrix} \bar{\sigma}_u^2 \mathbf{I}_{n-\bar{r}} & 0 \\ 0 & \Lambda_\perp \end{pmatrix} \quad \text{and} \quad P_0^\top \hat{\Delta}_y P_0 = \begin{pmatrix} \hat{\Delta}_{y,11} & \hat{\Delta}_{y,12} \\ \hat{\Delta}_{y,21} & \hat{\Delta}_{y,22} \end{pmatrix},$$

where $\hat{\Delta}_{y,11} \equiv A_0^\top \hat{\Delta}_y A_0$, $\hat{\Delta}_{y,12} \equiv A_0^\top \hat{\Delta}_y A_{0,\perp}$, $\hat{\Delta}_{y,22} \equiv A_{0,\perp}^\top \hat{\Delta}_y A_{0,\perp}$ and $\hat{\Delta}_{y,21}^\top = \hat{\Delta}_{y,12}$, equation (C.62) can be written as

$$\begin{pmatrix} A_0^\top \hat{A}_0 \hat{\Lambda} \\ A_{0,\perp}^\top \hat{A}_0 \hat{\Lambda} \end{pmatrix} = \begin{pmatrix} (\bar{\sigma}_u^2 \mathbf{I}_{n-\bar{r}} + \hat{\Delta}_{y,11}) A_0^\top \hat{A}_0 + \hat{\Delta}_{y,12} A_{0,\perp}^\top \hat{A}_0 \\ \hat{\Delta}_{y,21} A_0^\top \hat{A}_0 + (\Lambda_\perp + \hat{\Delta}_{y,22}) A_{0,\perp}^\top \hat{A}_0 \end{pmatrix}. \quad (\text{C.63})$$

The top block of (C.63) implies that

$$A_0^\top \hat{A}_0 \hat{\Lambda} = (\bar{\sigma}_u^2 \mathbf{I}_{n-\bar{r}} + \hat{\Delta}_{y,11}) A_0^\top \hat{A}_0 + \hat{\Delta}_{y,12} A_{0,\perp}^\top \hat{A}_0.$$

Premultiplying by $(A_0^\top \hat{A}_0)^{-1}$, we obtain

$$\hat{\Lambda} = (A_0^\top \hat{A}_0)^{-1}(\bar{\sigma}_u^2 \mathbf{I}_{n-\bar{r}} + \hat{\Delta}_{y,11})A_0^\top \hat{A}_0 + (A_0^\top \hat{A}_0)^{-1}\hat{\Delta}_{y,12}A_{0,\perp}^\top \hat{A}_0. \quad (\text{C.64})$$

The second block of (C.63) implies that

$$A_{0,\perp}^\top \hat{A}_0 \hat{\Lambda} = \hat{\Delta}_{y,21}A_0^\top \hat{A}_0 + (\Lambda_\perp + \hat{\Delta}_{y,22})A_{0,\perp}^\top \hat{A}_0.$$

Using (C.64) and $K_0 = (A_{0,\perp}^\top \hat{A}_0)(A_0^\top \hat{A}_0)^{-1}$, we obtain

$$K_0(\bar{\sigma}_u^2 \mathbf{I}_{n-\bar{r}} + \hat{\Delta}_{y,11}) + K_0 \hat{\Delta}_{y,12} K_0 = \hat{\Delta}_{y,21} + (\Lambda_\perp + \hat{\Delta}_{y,22})K_0.$$

Rearranging,

$$(\bar{\sigma}_u^2 \mathbf{I}_{\bar{r}-1} - \Lambda_\perp)K_0 = \hat{\Delta}_{y,21} - K_0 \hat{\Delta}_{y,11} + \hat{\Delta}_{y,22}K_0 - K_0 \hat{\Delta}_{y,12}K_0. \quad (\text{C.65})$$

From (C.38) and Lemma C.7, we have

$$\|K_0 \hat{\Delta}_{y,11}\| = \|K_0 A_0^\top \hat{\Delta}_y A_0\| \leq \|K_0\| \|A_0\|^2 \|\hat{\Delta}_y\| = O_p(T^{-1}). \quad (\text{C.66})$$

Similarly,

$$\|\hat{\Delta}_{y,22}K_0\| = \|A_{0,\perp}^\top \hat{\Delta}_y A_{0,\perp} K_0\| \leq \|K_0\| \|A_{0,\perp}\|^2 \|\hat{\Delta}_y\| = O_p(T^{-1}), \quad (\text{C.67})$$

and

$$\|K_0 \hat{\Delta}_{y,12} K_0\| = \|K_0 A_0^\top \hat{\Delta}_y A_{0,\perp} K_0\| \leq \|K_0\|^2 \|A_{0,\perp}\| \|A_0\| \|\hat{\Delta}_y\| = O_p(T^{-3/2}). \quad (\text{C.68})$$

Substituting (C.66)-(C.68) into (C.65),

$$(\bar{\sigma}_u^2 \mathbf{I}_{\bar{r}-1} - \Lambda_\perp)K_0 = \hat{\Delta}_{y,21} + O_p(T^{-1}),$$

which together with Assumption 2(iii) establishes the claim. Q.E.D.

Lemma C.11. *Under Assumptions 1, 2 and 3, we have*

$$T^{-1/2} \sum_{t \leq T} \tilde{A}^\top y_t b_t = A^\top T^{-1/2} \sum_{t \leq T} \xi_{b,t} + O_p(T^{-1/2}), \text{ for } b \in \{v, \varepsilon\}, \quad (\text{C.69})$$

where $\tilde{A} \equiv Q_{-1} \tilde{A}_0$ and $\tilde{A}_0 = \hat{A}_0 \hat{H}_0^\top$ by (C.45) and (C.56).

PROOF OF LEMMA C.11. Under Assumptions 2(i, ii) and 3, we can apply Lemmas C.4 and C.5 to obtain

$$T^{-1} \sum_{t \leq T} (y_t v_t - \mathbb{E}[y_t v_t]) = O_p(T^{-1/2}) \quad \text{and} \quad T^{-1} \sum_{t \leq T} (y_t \varepsilon_t - \mathbb{E}[y_t \varepsilon_t]) = O_p(T^{-1/2}). \quad (\text{C.70})$$

From Lemma C.6, we have

$$\begin{aligned}
\tilde{A} - A &= Q_{-1}(\hat{A}_0 \hat{H}_0^\top - A_0) \\
&= Q_{-1} A_{0,\perp} (\bar{\sigma}_u^2 \mathbf{I}_{\bar{r}-1} - \Lambda_\perp)^{-1} A_{0,\perp}^\top \hat{\Delta}_y A_0 + O_p(T^{-1}) \\
&= Q_{-1} A_{0,\perp} (\bar{\sigma}_u^2 \mathbf{I}_{\bar{r}-1} - \Lambda_\perp)^{-1} A_{0,\perp}^\top Q_{-1}^\top \left(T^{-1} \sum_{t \leq T} (y_t y_t^\top - \mathbb{E}[y_t y_t^\top]) \right) Q_{-1} A_0 + O_p(T^{-1}) \\
&= \Upsilon T^{-1} \sum_{t \leq T} (y_t y_t^\top - \mathbb{E}[y_t y_t^\top]) A + O_p(T^{-1}) = O_p(T^{-1/2}), \tag{C.71}
\end{aligned}$$

where the last equality follows from Assumption 2(i, iii, iv) and Lemma C.4. Combining (C.70) and (C.71), we obtain

$$\begin{aligned}
T^{-1/2} \sum_{t \leq T} \tilde{A}^\top y_t v_t &= T^{-1/2} \sum_{t \leq T} A^\top y_t v_t + T^{1/2} (\tilde{A} - A)^\top T^{-1} \sum_{t \leq T} y_t v_t \\
&= T^{-1/2} \sum_{t \leq T} A^\top y_t v_t + T^{1/2} (\tilde{A} - A)^\top T^{-1} \sum_{t \leq T} \mathbb{E}[y_t v_t] + O_p(T^{-1/2}) \\
&= A^\top T^{-1/2} \sum_{t \leq T} y_t v_t \\
&\quad + A^\top T^{-1/2} \sum_{t \leq T} (y_t y_t^\top - \mathbb{E}[y_t y_t^\top]) \Upsilon T^{-1} \sum_{t \leq T} \mathbb{E}[y_t v_t] + O_p(T^{-1/2}). \tag{C.72}
\end{aligned}$$

Next, note that $A \equiv Q_{-1} A_0$ is orthogonal to $(\mathbf{1}_n, \lambda)$. Using (19), we can write

$$A^\top y_t = A_0^\top Q_{-1}^\top (\phi p_t \mathbf{1}_n + \lambda \eta_t + u_t) = A_0^\top Q_{-1}^\top u_t.$$

Together with the definition of v_t , this implies

$$A^\top \mathbb{E}[y_t v_t] = A_0^\top Q_{-1}^\top \mathbb{E}[u_t v_t] = A_0^\top Q_{-1}^\top \mathbb{E}[u_t (y_{e,t} - \phi p_t)]. \tag{C.73}$$

Since $y_{e,t} = e^\top y_t$ and $y_t = \phi p_t \mathbf{1}_n + \lambda \eta_t + u_t$, we have

$$\mathbb{E}[u_t (y_{e,t} - \phi p_t)] = \mathbb{E}[u_t (\eta_t^\top \lambda^\top + u_t^\top)] e.$$

By Assumption 1, it follows that

$$Q_{-1}^\top \mathbb{E}[(u_t \eta_t^\top \lambda^\top + u_t u_t^\top)] e = n^{-1} Q_{-1}^\top (\mathbf{1}_n \Gamma_{\eta u, t}^\top \lambda^\top + \sigma_{u, t}^2 \mathbf{I}_n) \mathbf{1}_n = \mathbf{0}_{n-1}. \tag{C.74}$$

Combining (C.73) and (C.74) yields

$$A^\top \mathbb{E}[y_t v_t] = 0. \tag{C.75}$$

Substituting this into (C.72) and using the definition of $\xi_{v,t}$ establishes (C.69) for $b = v$. The proof for $b = \varepsilon$ follows by analogous arguments. *Q.E.D.*

D Proof of Consistency of the Variance Estimator

In this section, we establish the consistency of the variance matrix estimator $\hat{V}$ proposed in the algorithm under the assumption that ξ_t is uncorrelated across t . The following conditions are useful for proving this result.

Assumption D.1. (i) For any two distinct eigenvalues μ_{j_1} and μ_{j_2} of S_y , $|\mu_{j_1} - \mu_{j_2}| > K^{-1}$; (ii) for any sequence $\{c_t\}_{t \leq T} \subset \mathbb{R}$ with $\max_{t \leq T} |c_t| \leq K$, and $a, b \in \{u, \eta, \varepsilon\}$,

$$T^{-1} \sum_{t \leq T} c_t (a_t b_t^\top - \mathbb{E}[a_t b_t^\top]) = o_p(1);$$

(iii) for every random variable of the form $b_t = \prod_{j=1}^4 z_{j,t}$,

$$T^{-1} \sum_{t \leq T} (b_t - \mathbb{E}[b_t]) = o_p(1),$$

where each $z_{j,t}$ belongs to $\{u_{i,t}, \eta_{k,t}, \varepsilon_t, S_t^\top u_t, s_{i,t} \eta_{k,t} : i \leq n, k \leq r\}$; (iv) $\max_{t \leq T} \mathbb{E}[(u_t^\top u_t)^2 + \varepsilon_t^4 + (\eta_t^\top \eta_t)^2] \leq K$.

Assumption D.1(i) imposes a uniform separation (gap) between distinct eigenvalues of S_y , which is useful for establishing the consistency of $\hat{\Upsilon}$ used in constructing $\hat{\xi}_{b,t}$ for $b \in \{v, \varepsilon\}$. Since $\mu_1 = \dots = \mu_{n-\bar{r}} = \bar{\sigma}_u^2 < \mu_{n-\bar{r}+1}$, it is clear that Assumption 2(iii) is implied by Assumption D.1(i). Assumption D.1(ii) is analogous to Assumption 2(ii) and can be verified using a law of large numbers. Assumption D.1(iii) imposes a law of large numbers for fourth-order products formed from ε_t and components of u_t and η_t . Finally, Assumption D.1(iv) imposes uniform fourth-moment bounds on the shocks in the demand and supply equations.

Theorem D.1. Suppose that ξ_t is uncorrelated across t . Under Assumptions 1, 2, 3, 4(i, ii, iii), and D.1, we have

$$\hat{V} = V + o_p(1),$$

where $\hat{V}$ is defined in (A.2).

PROOF OF THEOREM D.1. Let $\gamma \equiv \Upsilon(T^{-1} \sum_{t \leq T} \mathbb{E}[y_t v_t])$. We then have

$$\begin{aligned} \xi_{v,t} &\equiv y_t v_t - \mathbb{E}[y_t v_t] + (y_t y_t^\top - \mathbb{E}[y_t y_t^\top]) \Upsilon \left(T^{-1} \sum_{t \leq T} \mathbb{E}[y_t v_t] \right) \\ &= y_t v_t - \mathbb{E}[y_t v_t] + (y_t y_t^\top - \mathbb{E}[y_t y_t^\top]) \gamma \\ &= y_t v_t - \mathbb{E}[y_t v_t] + (y_{\gamma,t} y_t - \mathbb{E}[y_{\gamma,t} y_t]), \end{aligned}$$

which implies that

$$\xi_{v,t}^\top \xi_{v,t} \leq K \left(v_t^2 y_t^\top y_t + y_{\gamma,t}^2 y_t^\top y_t + \|\mathbb{E}[v_t y_t]\|^2 + \|\mathbb{E}[y_{\gamma,t} y_t]\|^2 \right). \quad (\text{D.1})$$

Under Assumption 2(i), we can apply (B.11) to show that

$$\max_{t \leq T} \|\mathbb{E}[y_t v_t]\| \leq \max_{t \leq T} \|\mathbb{E}[y_t y_t^\top]\| + |\phi| \max_{t \leq T} \|\mathbb{E}[y_t p_t]\| \leq K. \quad (\text{D.2})$$

This together with (D.19), the triangle inequality and Cauchy-Schwarz inequality implies

$$\|\gamma\| \leq K \quad (\text{D.3})$$

Note that $v_t = y_{e,t} - \phi p_t$. By (C.25) and (C.26), both y_t and p_t are linear functions of u_t , η_t , and ε_t . It then follows from Assumptions 2(i), 3(ii) and D.1(iv) that

$$\max_{t \leq T} \left(\mathbb{E}[v_t^2 y_t^\top y_t] + \mathbb{E}[y_{\gamma,t}^2 y_t^\top y_t] \right) \leq K. \quad (\text{D.4})$$

Combining (D.1) and (D.4) with Markov's inequality yields

$$T^{-1} \sum_{t \leq T} \xi_{v,t}^\top \xi_{v,t} = O_p(1). \quad (\text{D.5})$$

Similarly, we can show that $T^{-1} \sum_{t \leq T} \xi_{\varepsilon,t}^\top \xi_{\varepsilon,t} = O_p(1)$. Together with (D.5), this implies

$$T^{-1} \sum_{t \leq T} \xi_t^\top \xi_t = O_p(1). \quad (\text{D.6})$$

By the triangle inequality, the Cauchy-Schwarz inequality, Lemma D.8, and (D.6), we have

$$\left\| T^{-1} \sum_{t \leq T} (\hat{\xi}_t - \xi_t)(\hat{\xi}_t - \xi_t)^\top \right\| \leq T^{-1} \sum_{t \leq T} \|\hat{\xi}_t - \xi_t\|^2 = o_p(1),$$

and

$$\left\| T^{-1} \sum_{t \leq T} (\hat{\xi}_t - \xi_t) \xi_t^\top \right\| \leq T^{-1} \sum_{t \leq T} \|\hat{\xi}_t - \xi_t\| \|\xi_t\| = o_p(1),$$

which, along with Lemma D.7, implies

$$\begin{aligned} T^{-1} \sum_{t \leq T} \hat{\xi}_t \hat{\xi}_t^\top - T^{-1} \sum_{t \leq T} \mathbb{E}[\xi_t \xi_t^\top] &= T^{-1} \sum_{t \leq T} (\xi_t \xi_t^\top - \mathbb{E}[\xi_t \xi_t^\top]) + T^{-1} \sum_{t \leq T} (\hat{\xi}_t - \xi_t)(\hat{\xi}_t - \xi_t)^\top \\ &\quad + T^{-1} \sum_{t \leq T} (\hat{\xi}_t - \xi_t) \xi_t^\top + T^{-1} \sum_{t \leq T} \xi_t (\hat{\xi}_t - \xi_t)^\top \\ &= o_p(1). \end{aligned} \quad (\text{D.7})$$

Since ξ_t is uncorrelated across t , we have $V = T^{-1} \sum_{t \leq T} \mathbb{E}[\xi_t \xi_t^\top]$. Thus, the result follows from (D.7) once we show that

$$T^{-1} \sum_{t \leq T} \hat{\xi}_t = o_p(1). \quad (\text{D.8})$$

To verify (D.8), note that by the triangle inequality,

$$\left\| T^{-1} \sum_{t \leq T} \hat{\xi}_t \right\| \leq \left\| T^{-1} \sum_{t \leq T} (\hat{\xi}_t - \xi_t) \right\| + \left\| T^{-1} \sum_{t \leq T} \xi_t \right\|$$

$$\leq \left(T^{-1} \sum_{t \leq T} \|\hat{\xi}_t - \xi_t\|^2 \right)^{1/2} + O_p(T^{-1/2}) = o_p(1),$$

where the second inequality follows from the Cauchy–Schwarz inequality and Assumption 4(i), and the last equality follows from Lemma D.8. This completes the proof. Q.E.D.

D.1 Useful Lemmas for consistency of variance estimator

Lemma D.1. *Suppose that the eigenvalues $\{\mu_j\}_{n-\bar{r}+1 \leq j \leq n-1}$ take m distinct values $\{\mu_j^*\}_{1 \leq j \leq m}$, ordered increasingly, with multiplicities $\{s_j\}_{1 \leq j \leq m}$. That is,*

$$\mu_{n-\bar{r}+\bar{s}_{j-1}+1} = \cdots = \mu_{n-\bar{r}+\bar{s}_j} = \mu_j^*, \quad (\text{D.9})$$

for $j \in \{1, \dots, m\}$, where $\bar{s}_{j-1} = \sum_{l=1}^{j-1} s_l$. Let A_j collect the orthonormal eigenvectors of S_y associated with the eigenvalues $\{\mu_{n-\bar{r}+\bar{s}_{j-1}+l}\}_{l=1}^{s_j}$, and let $\hat{A}_j$ collect the orthonormal eigenvectors of $\hat{S}_y$ associated with $\{\hat{\mu}_{n-\bar{r}+\bar{s}_{j-1}+l}\}_{l=1}^{s_j}$. Then under Assumptions 2(i, ii), 3, 4(i, ii), and D.1(i), there exists an orthogonal matrix $\hat{O}_j \in \mathbb{R}^{s_j \times s_j}$ such that for each $j \in \{1, \dots, m\}$,

$$\|\hat{A}_j \hat{O}_j - A_j\| = O_p(T^{-1/2}).$$

PROOF OF LEMMA D.1. Fix $j \in \{1, \dots, m\}$, and define

$$r_j \equiv n - \bar{r} + \bar{s}_{j-1} + 1, \quad s_j^* \equiv n - \bar{r} + \bar{s}_j.$$

Then $r_j \leq s_j^*$, and by (D.9), the columns of A_j and $\hat{A}_j$ are the orthonormal eigenvectors of S_y and $\hat{S}_y$, respectively, associated with the block of eigenvalues

$$\mu_{r_j} = \cdots = \mu_{s_j^*} = \mu_j^*.$$

Since Theorem 2 of Yu, Wang, and Samworth (2015) is stated for eigenvalues ordered in decreasing order, we apply it to $-S_y$ and $-\hat{S}_y$. By construction, $\mu_{r_1-1} = \mu_{n-\bar{r}} = \bar{\sigma}_u^2$. Following Yu, Wang, and Samworth (2015), we set $\mu_n = \infty$, which implies $\mu_{s_j^*+1} = \infty$. It then follows that there exists an orthogonal matrix $\hat{O}_j \in \mathbb{R}^{s_j \times s_j}$ such that

$$\|\hat{A}_j \hat{O}_j - A_j\| \leq \frac{2^{3/2} \min \left\{ s_j^{1/2} \|\hat{S}_y - S_y\|_o, \|\hat{S}_y - S_y\| \right\}}{\min \left\{ \mu_{r_j} - \mu_{r_j-1}, \mu_{s_j^*+1} - \mu_{s_j^*} \right\}}. \quad (\text{D.10})$$

We next show that the denominator in (D.10) is bounded away from zero uniformly in j . For $j \in \{2, \dots, m\}$, the eigenvalues μ_{r_j-1} and μ_{r_j} are distinct, and hence Assumption D.1(i) implies

$$\mu_{r_j} - \mu_{r_j-1} \geq K^{-1}.$$

Similarly, for $j \in \{1, \dots, m-1\}$, the eigenvalues $\mu_{s_j^*}$ and $\mu_{s_j^*+1}$ are distinct, so Assumption D.1(i) implies

$$\mu_{s_j^*+1} - \mu_{s_j^*} \geq K^{-1}.$$

For the boundary block $j = 1$, the left gap is

$$\mu_{r_1} - \mu_{r_1-1} = \mu_{n-\bar{r}+1} - \mu_{n-\bar{r}},$$

where $\mu_{n-\bar{r}} = \bar{\sigma}_u^2$ is not contained in the collection $\{\mu_j\}_{n-\bar{r}+1 \leq j \leq n-1}$. By Assumption D.1(i), this gap is bounded away from zero. Therefore,

$$\min\{\mu_{r_j} - \mu_{r_j-1}, \mu_{s_j^*+1} - \mu_{s_j^*}\} \geq K^{-1} \quad (\text{D.11})$$

for all $j \in \{1, \dots, m\}$.

Finally, by (B.45), $\|\hat{S}_y - S_y\| = O_p(T^{-1/2})$. Combining this with (D.10) and (D.11) establishes the claim of the lemma. *Q.E.D.*

Lemma D.2. *Under Assumptions 2(i, ii, iii, iv), 3, 4(i, ii) and D.1(i), we have*

$$\hat{\Upsilon} = \Upsilon + O_p(T^{-1/2}) = O_p(1), \quad (\text{D.12})$$

where $\hat{\Upsilon} \equiv Q_{-1} \hat{A}_{0,\perp} (\hat{\sigma}_u^2 \mathbf{I}_{\bar{r}-1} - \hat{\Lambda}_\perp)^{-1} \hat{A}_{0,\perp}^\top Q_{-1}^\top$ and $\Upsilon \equiv Q_{-1} A_{0,\perp} (\bar{\sigma}_u^2 \mathbf{I}_{\bar{r}-1} - \Lambda_\perp)^{-1} A_{0,\perp}^\top Q_{-1}^\top$.

PROOF OF LEMMA D.2. Suppose that the eigenvalues $\{\mu_j\}_{n-\bar{r}+1 \leq j \leq n-1}$ take m distinct values $\{\mu_j^*\}_{1 \leq j \leq m}$, with multiplicities $\{s_j\}_{1 \leq j \leq m}$. Then we can write

$$A_{0,\perp} = (A_1, \dots, A_m) \quad \text{and} \quad \hat{A}_{0,\perp} = (\hat{A}_1, \dots, \hat{A}_m),$$

where A_j and $\hat{A}_j$ collect the orthonormal eigenvectors of S_y and $\hat{S}_y$, respectively, associated with the eigenvalue μ_j^* of multiplicity s_j . By Lemma D.1, there exists a block diagonal matrix $\hat{H}_{0,\perp}^\top = \text{diag}(\hat{O}_1, \dots, \hat{O}_m)$, where each $\hat{O}_j \in \mathbb{R}^{s_j \times s_j}$ is orthogonal, such that

$$\hat{A}_{0,\perp} \hat{H}_{0,\perp}^\top = A_{0,\perp} + O_p(T^{-1/2}) = O_p(1). \quad (\text{D.13})$$

From $\|\hat{S}_y - S_y\| = O_p(T^{-1/2})$, we have

$$\hat{\sigma}_u^2 = \bar{\sigma}_u^2 + O_p(T^{-1/2}), \quad \hat{\Lambda}_\perp = \Lambda_\perp + O_p(T^{-1/2}), \quad (\text{D.14})$$

where $\bar{\sigma}_u^2$ and $\Lambda_\perp$ are bounded by (B.11). Since

$$\Lambda_\perp = \text{diag}(\mu_1^* I_{s_1}, \dots, \mu_m^* I_{s_m}),$$

and $\hat{H}_{0,\perp}^\top$ is block diagonal with orthogonal blocks, we have

$$\hat{H}_{0,\perp} \Lambda_\perp \hat{H}_{0,\perp}^\top = \Lambda_\perp. \quad (\text{D.15})$$

Therefore,

$$(\hat{\sigma}_u^2 I_{\bar{r}-1} - \hat{H}_{0,\perp} \hat{\Lambda}_\perp \hat{H}_{0,\perp}^\top) - (\bar{\sigma}_u^2 I_{\bar{r}-1} - \Lambda_\perp) = (\hat{\sigma}_u^2 - \bar{\sigma}_u^2) I_{\bar{r}-1} - \hat{H}_{0,\perp} (\hat{\Lambda}_\perp - \Lambda_\perp) \hat{H}_{0,\perp}^\top = O_p(T^{-1/2}). \quad (\text{D.16})$$

By Assumption D.1(i), the eigenvalues of $\Lambda_\perp - \bar{\sigma}_u^2 I_{\bar{r}-1}$ are bounded below by K^{-1} . Hence, by (D.16),

$$\rho_{\min}(\hat{H}_{0,\perp} \hat{\Lambda}_\perp \hat{H}_{0,\perp}^\top - \hat{\sigma}_u^2 I_{\bar{r}-1}) \geq (2K)^{-1}, \quad \text{wpa1}. \quad (\text{D.17})$$

Since $\hat{H}_{0,\perp}^\top \hat{H}_{0,\perp} = I_{\bar{r}-1}$, we have

$$(\hat{\sigma}_u^2 I_{\bar{r}-1} - \hat{\Lambda}_\perp)^{-1} = \hat{H}_{0,\perp}^\top (\hat{\sigma}_u^2 I_{\bar{r}-1} - \hat{H}_{0,\perp} \hat{\Lambda}_\perp \hat{H}_{0,\perp}^\top)^{-1} \hat{H}_{0,\perp}.$$

Notice that

$$\begin{aligned} \hat{\Upsilon} - \Upsilon &= Q_{-1} \hat{A}_{0,\perp} \hat{H}_{0,\perp}^\top (\hat{\sigma}_u^2 I_{\bar{r}-1} - \hat{H}_{0,\perp} \hat{\Lambda}_\perp \hat{H}_{0,\perp}^\top)^{-1} \hat{H}_{0,\perp} \hat{A}_{0,\perp}^\top Q_{-1}^\top \\ &\quad - Q_{-1} A_{0,\perp} (\bar{\sigma}_u^2 I_{\bar{r}-1} - \Lambda_\perp)^{-1} A_{0,\perp}^\top Q_{-1}^\top \\ &= Q_{-1} (\hat{A}_{0,\perp} \hat{H}_{0,\perp}^\top - A_{0,\perp}) (\hat{\sigma}_u^2 I_{\bar{r}-1} - \hat{H}_{0,\perp} \hat{\Lambda}_\perp \hat{H}_{0,\perp}^\top)^{-1} \hat{H}_{0,\perp} \hat{A}_{0,\perp}^\top Q_{-1}^\top \\ &\quad + Q_{-1} A_{0,\perp} \left[(\hat{\sigma}_u^2 I_{\bar{r}-1} - \hat{H}_{0,\perp} \hat{\Lambda}_\perp \hat{H}_{0,\perp}^\top)^{-1} - (\bar{\sigma}_u^2 I_{\bar{r}-1} - \Lambda_\perp)^{-1} \right] \hat{H}_{0,\perp} \hat{A}_{0,\perp}^\top Q_{-1}^\top \\ &\quad + Q_{-1} A_{0,\perp} (\bar{\sigma}_u^2 I_{\bar{r}-1} - \Lambda_\perp)^{-1} (\hat{H}_{0,\perp} \hat{A}_{0,\perp}^\top - A_{0,\perp}^\top) Q_{-1}^\top. \end{aligned}$$

By (D.13)–(D.17), each term on the right-hand side is $O_p(T^{-1/2})$. Therefore,

$$\hat{\Upsilon} - \Upsilon = O_p(T^{-1/2}), \quad (\text{D.18})$$

which establishes the first equality in (D.12).

Since $Q_{-1}^\top Q_{-1} = \mathbf{I}_{n-1}$, $A_{0,\perp}^\top A_{0,\perp} = \mathbf{I}_{\bar{r}-1}$ and $\Lambda_\perp \equiv \text{diag}((\mu_j)_{n-\bar{r}+1 \leq j \leq n-1})$, from the Cauchy-Schwarz inequality and Assumption 2(iii), it follows that

$$\|\Upsilon\| = \|Q_{-1} A_{0,\perp} (\bar{\sigma}_u^2 \mathbf{I}_{\bar{r}-1} - \Lambda_\perp)^{-1} A_{0,\perp}^\top Q_{-1}^\top\| \leq \frac{\|Q_{-1}\|^2 \|A_{0,\perp}\|^2 (\bar{r} - 1)}{\mu_{n-\bar{r}+1} - \bar{\sigma}_u^2} \leq K, \quad (\text{D.19})$$

which together with (D.18) shows the second equality in (D.12). Q.E.D.

Lemma D.3. *Under Assumptions 1, 2, 3 and 4(i, iii), we have for $b \in \{v, \varepsilon\}$*

$$T^{-1} \sum_{t \leq T} y_t \hat{b}_t = T^{-1} \sum_{t \leq T} \mathbb{E}[y_t b_t] + O_p(T^{-1/2}) = O_p(1).$$

PROOF OF LEMMA D.3. Because $v_t \equiv y_{e,t} - \phi p_t$, we can write

$$\begin{aligned} T^{-1} \sum_{t \leq T} y_t v_t &= T^{-1} \sum_{t \leq T} \mathbb{E}[y_t v_t] + T^{-1} \sum_{t \leq T} (y_t v_t - \mathbb{E}[y_t v_t]) \\ &= T^{-1} \sum_{t \leq T} \mathbb{E}[y_t v_t] + T^{-1} \sum_{t \leq T} (y_t y_{e,t} - \mathbb{E}[y_t y_{e,t}]) - \phi T^{-1} \sum_{t \leq T} (y_t p_t - \mathbb{E}[y_t p_t]) \end{aligned}$$

$$= T^{-1} \sum_{t \leq T} \mathbb{E}[y_t v_t] + O_p(T^{-1/2}) = O_p(1), \quad (\text{D.20})$$

where the third equality follows from Lemma C.4, and the final equality uses (D.2).

Since $\hat{\theta}_0(\hat{A})$ denotes the GIV estimator with weighting matrix $W_{0,T} = \mathbf{I}_{2n}$ (see Step 2 of the Implementation Algorithm 1 in Appendix A), Assumption 4(ii) holds with $W_0 = \mathbf{I}_{2n}$. Therefore, under Assumptions 1, 2, 3, and 4(i, iii), we can apply (42) of Theorem 1 to obtain

$$\hat{\theta}_0(\hat{A}) - \theta = O_p(T^{-1/2}). \quad (\text{D.21})$$

Combining this with (B.12) yields

$$T^{-1} \sum_{t \leq T} y_t \hat{v}_t - T^{-1} \sum_{t \leq T} y_t v_t = -(\hat{\phi}_0(\hat{A}) - \phi) T^{-1} \sum_{t \leq T} y_t p_t = O_p(T^{-1/2}). \quad (\text{D.22})$$

The result for $b = v$ follows from (D.20) and (D.22). The case $b = \varepsilon$ can be established analogously and is therefore omitted. Q.E.D.

Lemma D.4. *Under Assumptions 2(i), 3 and D.1(ii), we have for any $c_t \in \mathbb{R}$ with $\max_{t \leq T} |c_t| \leq K$,*

$$T^{-1} \sum_{t \leq T} c_t (y_t y_t^\top - \mathbb{E}[y_t y_t^\top]) = o_p(1) \quad \text{and} \quad T^{-1} \sum_{t \leq T} c_t (p_t y_t - \mathbb{E}[p_t y_t]) = o_p(1).$$

PROOF OF LEMMA D.4. Using the reduced-form representations of y_t and p_t in (C.25) and (C.26), the stated results follow by the same argument as in the proof of Lemma C.4, with Assumption 2(ii) replaced by Assumption D.1(ii). Q.E.D.

Lemma D.5. *Under Assumptions 2(i), 3(ii) and D.1(iii, iv), we have*

$$T^{-1} \sum_{t \leq T} (y_{a_1,t} y_{a_2,t} y_t^\top - \mathbb{E}[y_{a_1,t} y_{a_2,t} y_t^\top]) = o_p(1),$$

for any $a_1, a_2 \in \{a \in \mathbb{R}^n : \|a\| \leq K\}$.

PROOF OF LEMMA D.5. From the reduced form expression of y_t in (C.25), the triangle inequality, and the Cauchy-Schwarz inequality,

$$\begin{aligned} \|y_t\| &\leq \left(1 + \frac{|\phi\psi|n^{1/2}\|S_t\|}{|1 - \phi\psi|}\right) (\|u_t\| + \|\lambda\| \|\eta_t\|) + \frac{|\phi|n^{1/2}|\varepsilon_t|}{|1 - \phi\psi|} \\ &\leq K(\|u_t\| + \|\eta_t\| + |\varepsilon_t|) \end{aligned}$$

where the second inequality follows by Assumptions 2(i) and 3(ii). Therefore, by Assumption D.1(iv),

$$\mathbb{E}[\|y_t\|^4] \leq K(\mathbb{E}[\|u_t\|^4 + \|\eta_t\|^4 + \varepsilon_t^4]) \leq K.$$

This together with the Cauchy-Schwarz inequality implies that

$$\|\mathbb{E}[y_{a_1,t}y_{a_2,t}y_t y_t^\top]\| \leq K,$$

for any $a_1, a_2 \in \{a \in \mathbb{R}^n : \|a\| \leq K\}$. Since y_t is a linear function of $u_t, \eta_t, \varepsilon_t, S_t^\top u_t$ and $(\eta_t^\top \otimes S_t^\top)\text{vec}(\lambda)$, each entry of $y_t y_t^\top$ is a finite linear combination of random variables of the form

$$b_{1,t} = \prod_{j=1}^2 z_{j,t},$$

where each $z_{j,t}$ belongs to $\{u_{i,t}, \eta_{k,t}, \varepsilon_t, S_t^\top u_t, (\eta_t^\top \otimes S_t^\top)\text{vec}(\lambda) : i \leq n, k \leq r\}$, with coefficients that do not depend on t . Since

$$y_{a_1,t}y_{a_2,t} = a_1^\top y_t y_t^\top a_2,$$

and $\|a_1\|, \|a_2\| \leq K$, the scalar $y_{a_1,t}y_{a_2,t}$ is also a finite linear combination of terms of the form $w_{1,t}$, with coefficients bounded uniformly over admissible a_1, a_2 . Therefore, for each (i_1, i_2) , the (i_1, i_2) th entry of

$$y_{a_1,t}y_{a_2,t}y_t y_t^\top$$

is a finite linear combination of random variables of the form

$$b_t = \prod_{j=1}^4 z_{j,t}$$

with coefficients uniformly bounded over admissible a_1, a_2 , where each $z_{j,t}$ belongs to the same set above. Assumption D.1(iii) then yields

$$T^{-1} \sum_{t \leq T} (b_t - \mathbb{E}[b_t]) = o_p(1)$$

for each such fourth-order polynomial. Since n is fixed, each entry of

$$T^{-1} \sum_{t \leq T} (y_{a_1,t}y_{a_2,t}y_t y_t^\top - \mathbb{E}[y_{a_1,t}y_{a_2,t}y_t y_t^\top])$$

is a finite linear combination of $o_p(1)$ terms, and is therefore $o_p(1)$. This proves the claim. *Q.E.D.*

Lemma D.6. *Under Assumptions 2(i), 3(ii) and D.1(iii, iv), we have for any $a \in \mathbb{R}^n$ with $\|a\| \leq K$*

$$T^{-1} \sum_{t \leq T} (y_{a,t} p_t y_t y_t^\top - \mathbb{E}[y_{a,t} p_t y_t y_t^\top]) = o_p(1),$$

and

$$T^{-1} \sum_{t \leq T} (p_t^2 y_t y_t^\top - \mathbb{E}[p_t^2 y_t y_t^\top]) = o_p(1).$$

PROOF OF LEMMA D.6. Under Assumption 2(i), it follows from (C.26) that p_t is a linear function of $S_t^\top u_t$, $\eta_t \otimes S_t$, and ε_t . The stated results then follow by the same argument as in the proof of Lemma D.5. Q.E.D.

Lemma D.7. *Under Assumptions 2(i), 3(ii) and D.1(ii, iii, iv), we have*

$$T^{-1} \sum_{t \leq T} \xi_t \xi_t^\top = T^{-1} \sum_{t \leq T} \mathbb{E}[\xi_t \xi_t^\top] + o_p(1).$$

PROOF OF LEMMA D.7. By the definition of ξ_t , it suffices to show that

$$T^{-1} \sum_{t \leq T} \xi_{b,t} \xi_{b,t}^\top = T^{-1} \sum_{t \leq T} \mathbb{E}[\xi_{b,t} \xi_{b,t}^\top] + o_p(1), \quad (\text{D.23})$$

for $b \in \{v, \varepsilon\}$, and

$$T^{-1} \sum_{t \leq T} \xi_{v,t} \xi_{\varepsilon,t}^\top = T^{-1} \sum_{t \leq T} \mathbb{E}[\xi_{v,t} \xi_{\varepsilon,t}^\top] + o_p(1). \quad (\text{D.24})$$

We establish (D.23) for $b = v$; the remaining cases follow analogously.

Recall that $\gamma \equiv \Upsilon(T^{-1} \sum_{t \leq T} \mathbb{E}[y_t v_t])$, and

$$\begin{aligned} \xi_{v,t} &\equiv y_t v_t - \mathbb{E}[y_t v_t] + (y_t y_t^\top - \mathbb{E}[y_t y_t^\top]) \Upsilon \left(T^{-1} \sum_{t \leq T} \mathbb{E}[y_t v_t] \right) \\ &= y_t v_t - \mathbb{E}[y_t v_t] + (y_{\gamma,t} y_t - \mathbb{E}[y_{\gamma,t} y_t]). \end{aligned}$$

This implies that

$$\begin{aligned} T^{-1} \sum_{t \leq T} \xi_{v,t} \xi_{v,t}^\top &= T^{-1} \sum_{t \leq T} (v_t y_t - \mathbb{E}[v_t y_t]) (v_t y_t^\top - \mathbb{E}[v_t y_t^\top]) \\ &\quad + T^{-1} \sum_{t \leq T} (v_t y_t - \mathbb{E}[v_t y_t]) (y_{\gamma,t} y_t^\top - \mathbb{E}[y_{\gamma,t} y_t^\top]) \\ &\quad + T^{-1} \sum_{t \leq T} (y_{\gamma,t} y_t - \mathbb{E}[y_{\gamma,t} y_t]) (v_t y_t^\top - \mathbb{E}[v_t y_t^\top]) \\ &\quad + T^{-1} \sum_{t \leq T} (y_{\gamma,t} y_t - \mathbb{E}[y_{\gamma,t} y_t]) (y_{\gamma,t} y_t^\top - \mathbb{E}[y_{\gamma,t} y_t^\top]). \end{aligned} \quad (\text{D.25})$$

We next study the four terms on the right-hand side of (D.25) one by one.

Step 1: The first term. Expanding the first term yields

$$\begin{aligned} &T^{-1} \sum_{t \leq T} (y_t v_t - \mathbb{E}[y_t v_t]) (y_t v_t - \mathbb{E}[y_t v_t])^\top \\ &= T^{-1} \sum_{t \leq T} v_t^2 y_t y_t^\top + T^{-1} \sum_{t \leq T} \mathbb{E}[v_t y_t] \mathbb{E}[v_t y_t^\top] - T^{-1} \sum_{t \leq T} (\mathbb{E}[v_t y_t] v_t y_t^\top + v_t y_t \mathbb{E}[v_t y_t^\top]). \end{aligned} \quad (\text{D.26})$$

Since $v_t = y_{e,t} - \phi p_t$, Lemmas D.5 and D.6 imply that

$$\begin{aligned}
T^{-1} \sum_{t \leq T} (v_t^2 y_t y_t^\top - \mathbb{E}[v_t^2 y_t y_t^\top]) &= T^{-1} \sum_{t \leq T} (y_{e,t}^2 y_t y_t^\top - \mathbb{E}[y_{e,t}^2 y_t y_t^\top]) \\
&\quad + \phi^2 T^{-1} \sum_{t \leq T} (p_t^2 y_t y_t^\top - \mathbb{E}[p_t^2 y_t y_t^\top]) \\
&\quad - 2\phi T^{-1} \sum_{t \leq T} (y_{e,t} p_t y_t y_t^\top - \mathbb{E}[y_{e,t} p_t y_t y_t^\top]) \\
&= o_p(1).
\end{aligned} \tag{D.27}$$

Moreover, since $\max_{t \leq T} \|\mathbb{E}[y_t v_t]\| \leq K$ by (D.2), Lemma D.4 implies that

$$\begin{aligned}
T^{-1} \sum_{t \leq T} \mathbb{E}[y_t v_t] (y_t^\top v_t - \mathbb{E}[y_t^\top v_t]) &= T^{-1} \sum_{t \leq T} \mathbb{E}[y_t v_t] (y_t^\top y_{e,t} - \mathbb{E}[y_t^\top y_{e,t}]) \\
&\quad - \phi T^{-1} \sum_{t \leq T} \mathbb{E}[y_t v_t] (y_t^\top p_t - \mathbb{E}[y_t^\top p_t]) \\
&= o_p(1).
\end{aligned} \tag{D.28}$$

Combining (D.26)–(D.28) yields

$$T^{-1} \sum_{t \leq T} (y_t v_t - \mathbb{E}[y_t v_t]) (y_t v_t - \mathbb{E}[y_t v_t])^\top = T^{-1} \sum_{t \leq T} \mathbb{E}[(y_t v_t - \mathbb{E}[y_t v_t]) (y_t v_t - \mathbb{E}[y_t v_t])^\top] + o_p(1). \tag{D.29}$$

Step 2: The cross terms. For the second term on the right-hand side of (D.25), we have

$$\begin{aligned}
&T^{-1} \sum_{t \leq T} (y_t v_t - \mathbb{E}[y_t v_t]) (y_{\gamma,t} y_t^\top - \mathbb{E}[y_{\gamma,t} y_t^\top]) \\
&= T^{-1} \sum_{t \leq T} y_{\gamma,t} v_t y_t y_t^\top - T^{-1} \sum_{t \leq T} y_t v_t \mathbb{E}[y_{\gamma,t} y_t^\top] \\
&\quad - T^{-1} \sum_{t \leq T} \mathbb{E}[y_t v_t] y_{\gamma,t} y_t^\top + T^{-1} \sum_{t \leq T} \mathbb{E}[y_t v_t] \mathbb{E}[y_{\gamma,t} y_t^\top].
\end{aligned} \tag{D.30}$$

Using (D.3), Lemmas D.5 and D.6, we obtain

$$\begin{aligned}
T^{-1} \sum_{t \leq T} (y_{\gamma,t} v_t y_t y_t^\top - \mathbb{E}[y_{\gamma,t} v_t y_t y_t^\top]) &= T^{-1} \sum_{t \leq T} (y_{\gamma,t} y_{e,t} y_t y_t^\top - \mathbb{E}[y_{\gamma,t} y_{e,t} y_t y_t^\top]) \\
&\quad - \phi T^{-1} \sum_{t \leq T} (y_{\gamma,t} p_t y_t y_t^\top - \mathbb{E}[y_{\gamma,t} p_t y_t y_t^\top]) = o_p(1).
\end{aligned} \tag{D.31}$$

Moreover, from (B.11) and (D.3),

$$\max_{t \leq T} \|\mathbb{E}[y_{\gamma,t} y_t^\top]\| \leq K, \tag{D.32}$$

which, together with Lemma D.4, implies

$$T^{-1} \sum_{t \leq T} (y_t v_t - \mathbb{E}[y_t v_t]) \mathbb{E}[y_{\gamma,t} y_t^\top] = T^{-1} \sum_{t \leq T} (y_t y_{e,t} - \mathbb{E}[y_t y_{e,t}]) \mathbb{E}[y_{\gamma,t} y_t^\top]$$

$$-\phi T^{-1} \sum_{t \leq T} (y_t p_t - \mathbb{E}[y_t p_t]) \mathbb{E}[y_{\gamma, t} y_t^\top] = o_p(1). \quad (\text{D.33})$$

Similarly, we can show that

$$T^{-1} \sum_{t \leq T} \mathbb{E}[y_t v_t] (y_{\gamma, t} y_t^\top - \mathbb{E}[y_{\gamma, t} y_t^\top]) = o_p(1). \quad (\text{D.34})$$

From (D.30)-(D.34), we conclude that

$$\begin{aligned} & T^{-1} \sum_{t \leq T} (y_t v_t - \mathbb{E}[y_t v_t]) (y_{\gamma, t} y_t^\top - \mathbb{E}[y_{\gamma, t} y_t^\top]) \\ &= T^{-1} \sum_{t \leq T} \mathbb{E} \left[(y_t v_t - \mathbb{E}[y_t v_t]) (y_{\gamma, t} y_t^\top - \mathbb{E}[y_{\gamma, t} y_t^\top]) \right] + o_p(1). \end{aligned} \quad (\text{D.35})$$

The third term in (D.25) can be handled analogously.

Step 3: The last term. Finally, for the fourth term in (D.25), we have

$$\begin{aligned} & T^{-1} \sum_{t \leq T} (y_{\gamma, t} y_t - \mathbb{E}[y_{\gamma, t} y_t]) (y_{\gamma, t} y_t^\top - \mathbb{E}[y_{\gamma, t} y_t^\top]) \\ &= T^{-1} \sum_{t \leq T} y_{\gamma, t}^2 y_t y_t^\top - T^{-1} \sum_{t \leq T} \mathbb{E}[y_{\gamma, t} y_t] y_{\gamma, t} y_t^\top \\ &\quad - T^{-1} \sum_{t \leq T} y_{\gamma, t} y_t \mathbb{E}[y_{\gamma, t} y_t^\top] + T^{-1} \sum_{t \leq T} \mathbb{E}[y_{\gamma, t} y_t] \mathbb{E}[y_{\gamma, t} y_t^\top] \\ &= T^{-1} \sum_{t \leq T} \mathbb{E}[y_{\gamma, t}^2 y_t y_t^\top] - T^{-1} \sum_{t \leq T} \mathbb{E}[y_{\gamma, t} y_t] \mathbb{E}[y_{\gamma, t} y_t^\top] + o_p(1) \end{aligned} \quad (\text{D.36})$$

where the second equality holds by (D.32), Lemma D.4 and Lemma D.5.

Combining (D.25), (D.29), (D.35), and (D.36) establishes (D.23) for $b = v$, completing the proof. *Q.E.D.*

Lemma D.8. *Under Assumptions 1, 2, 3, 4(i, iii) and D.1, we have*

$$T^{-1} \sum_{t \leq T} (\hat{\xi}_t - \xi_t)^\top (\hat{\xi}_t - \xi_t) = o_p(1).$$

PROOF OF LEMMA D.8. Let $\hat{\gamma} \equiv \hat{\Upsilon} T^{-1} \sum_{t \leq T} y_t \hat{v}_t$. Then, by Lemma D.2 and Lemma D.3,

$$\hat{\gamma} - \gamma = (\hat{\Upsilon} - \Upsilon) T^{-1} \sum_{t \leq T} y_t \hat{v}_t + \Upsilon \left(T^{-1} \sum_{t \leq T} y_t \hat{v}_t - T^{-1} \sum_{t \leq T} \mathbb{E}[y_t v_t] \right) = o_p(1). \quad (\text{D.37})$$

Because

$$\hat{\xi}_{v, t} \equiv y_t \hat{v}_t - T^{-1} \sum_{t \leq T} y_t \hat{v}_t + \left(y_t y_t^\top - T^{-1} \sum_{t \leq T} y_t y_t^\top \right) \hat{\Upsilon} \left(T^{-1} \sum_{t \leq T} y_t \hat{v}_t \right)$$

$$= y_t \hat{v}_t - T^{-1} \sum_{t \leq T} y_t \hat{v}_t + \left(y_t y_t^\top - T^{-1} \sum_{t \leq T} y_t y_t^\top \right) \hat{\gamma}$$

and

$$\xi_{v,t} = y_t v_t - \mathbb{E}[y_t v_t] + (y_{\gamma,t} y_t - \mathbb{E}[y_{\gamma,t} y_t]),$$

we have

$$\hat{\xi}_{v,t} - \xi_{v,t} = y_t(\hat{v}_t - v_t) + \left(y_t y_t^\top - \mathbb{E}[y_t y_t^\top] \right) (\hat{\gamma} - \gamma) - \Delta_T, \quad (\text{D.38})$$

where

$$\Delta_T = T^{-1} \sum_{t \leq T} (y_t \hat{v}_t - \mathbb{E}[y_t v_t]) + T^{-1} \sum_{t \leq T} (y_t y_t^\top - \mathbb{E}[y_t y_t^\top]) \hat{\gamma}.$$

Therefore by the Cauchy-Schwarz inequality,

$$\begin{aligned} T^{-1} \sum_{t \leq T} (\hat{\xi}_{v,t} - \xi_{v,t})^\top (\hat{\xi}_{v,t} - \xi_{v,t}) &\leq 3T^{-1} \sum_{t \leq T} (\hat{v}_t - v_t)^2 y_t^\top y_t + 3\|\Delta_T\|^2 \\ &\quad + 3(\hat{\gamma} - \gamma)^\top T^{-1} \sum_{t \leq T} \left(y_t y_t^\top - \mathbb{E}[y_t y_t^\top] \right)^2 (\hat{\gamma} - \gamma). \end{aligned} \quad (\text{D.39})$$

Since $\hat{v}_t - v_t = -(\hat{\phi}_0(\hat{A}) - \phi)p_t$, we have

$$T^{-1} \sum_{t \leq T} (\hat{v}_t - v_t)^2 y_t^\top y_t = (\hat{\phi}_0(\hat{A}) - \phi)^2 T^{-1} \sum_{t \leq T} p_t^2 y_t^\top y_t. \quad (\text{D.40})$$

Under Assumptions 2(i) and D.1(iv), together with the reduced-form expressions for y_t and p_t in (C.25) and (C.26), we have

$$\max_{t \leq T} \left(\mathbb{E}[p_t^2 y_t^\top y_t] + \mathbb{E}[(y_t^\top y_t)^2] \right) \leq K. \quad (\text{D.41})$$

Hence, by Markov's inequality,

$$T^{-1} \sum_{t \leq T} p_t^2 y_t^\top y_t = O_p(1), \quad (\text{D.42})$$

which, together with (D.21) and (D.40), implies

$$T^{-1} \sum_{t \leq T} (\hat{v}_t - v_t)^2 y_t^\top y_t = o_p(1). \quad (\text{D.43})$$

From Lemma C.4, Lemma D.3, (D.3), and (D.37), it follows that

$$\|\Delta_T\| \leq \left\| T^{-1} \sum_{t \leq T} (y_t \hat{v}_t - \mathbb{E}[y_t v_t]) \right\| + \left\| T^{-1} \sum_{t \leq T} (y_t y_t^\top - \mathbb{E}[y_t y_t^\top]) \hat{\gamma} \right\| = o_p(1). \quad (\text{D.44})$$

Finally, by (D.41), the Cauchy-Schwarz inequality and Markov's inequality,

$$T^{-1} \sum_{t \leq T} \left(y_t y_t^\top - \mathbb{E}[y_t y_t^\top] \right)^2 = O_p(1),$$

which, together with (D.37), implies

$$(\hat{\gamma} - \gamma)^\top T^{-1} \sum_{t \leq T} \left(y_t y_t^\top - \mathbb{E}[y_t y_t^\top] \right)^2 (\hat{\gamma} - \gamma) = o_p(1). \quad (\text{D.45})$$

Combining (D.39), (D.43), (D.44), and (D.45) yields

$$T^{-1} \sum_{t \leq T} (\hat{\xi}_{v,t} - \xi_{v,t})^\top (\hat{\xi}_{v,t} - \xi_{v,t}) = o_p(1).$$

Similarly, we can show that

$$T^{-1} \sum_{t \leq T} (\hat{\xi}_{\varepsilon,t} - \xi_{\varepsilon,t})^\top (\hat{\xi}_{\varepsilon,t} - \xi_{\varepsilon,t}) = o_p(1).$$

Combining these results establishes the claim of the lemma. Q.E.D.

E Asymptotic Variance with Exogenous Regressors

This section studies the asymptotic variance of the GIV estimator proposed in Subsection 4.2, where both the supply and demand equations include exogenous regressors. The key challenge is to account for the role of the estimation error in β in the asymptotic variance of the GIV estimator. As we show below, this estimation error is asymptotically negligible and can therefore be ignored. Consequently, the standard errors of the GIV estimator, as well as the J -test, can be constructed as if β were known. We begin by introducing a set of conditions that characterize the properties of the additional regressors in the model.

Assumption E.1. (i) For any $t \leq T$ and $i \leq n$, we have $\mathbb{E}[x_{i,t}(u_t^\top, \eta_t^\top)] = \mathbf{0}_{d_x \times (n+r)}$, $\mathbb{E}[x_{i,t}\varepsilon_t] = \mathbf{0}_{d_x}$ and $\mathbb{E}[w_t\varepsilon_t] = \mathbf{0}_{d_w \times 1}$; (ii) let $\tilde{\varepsilon}_t \equiv \tilde{\lambda}\eta_t + \tilde{u}_t$, then

$$T^{1/2}(\hat{\beta} - \beta) = \left(T^{-1} \sum_{t \leq T} \mathbb{E}[\tilde{x}_t^\top \tilde{x}_t] \right)^{-1} \left(T^{-1/2} \sum_{t \leq T} \tilde{x}_t^\top \tilde{\varepsilon}_t \right) + o_p(1),$$

where $\rho_{\min}(T^{-1} \sum_{t \leq T} \mathbb{E}[\tilde{x}_t^\top \tilde{x}_t]) \geq K^{-1}$ and $T^{-1/2} \sum_{t \leq T} \tilde{x}_t^\top \tilde{\varepsilon}_t = O_p(1)$; (iii) for $a \in \{u, \eta, \varepsilon\}$,

$$T^{-1} \sum_{t \leq T} (x_t \otimes a_t - \mathbb{E}[x_t \otimes a_t]) = O_p(T^{-1/2});$$

(iv) $\max_{t \leq T} \mathbb{E}[\|x_t\|^2 + \|w_t\|^2] \leq K$, $\|\beta\| \leq K$ and $\|\gamma\| \leq K$.

Assumption E.1(i) ensures that x_t and w_t are exogenous in the demand and supply equations, respectively. Assumption E.1(ii) provides a linear representation of the estimation error in $\hat{\beta}$, implying that $\hat{\beta}$ is $T^{1/2}$ -consistent. Assumption E.1(iii) imposes a $T^{-1/2}$ law of large numbers for products of x_t and the demand and supply shocks. Assumption E.1(iv) places uniform bounds on the second moments of the exogenous regressors, as well as on the magnitudes of their coefficients.

Since the matrix $\hat{A}$ in the moment conditions $\bar{g}_T(\theta; \hat{A}, \hat{\beta})$ in (62) plays the same role as in $\bar{g}_T(\theta; \hat{A})$ defined in (36), we can apply the same arguments as in the proof of Lemma C.3 to show that both the GIV estimator $\hat{\theta}(\hat{A})$ in (63) and the J -statistic

$$T \bar{g}_T(\hat{\theta}(\hat{A}); \hat{A}, \hat{\beta})^\top W_{0,T}(\hat{A}) \bar{g}_T(\hat{\theta}(\hat{A}); \hat{A}, \hat{\beta})$$

are invariant to nonsingular rotations of $\hat{A}$ of the form $\hat{A}C_1$, for any nonsingular $(n - \bar{r}) \times (n - \bar{r})$ matrix C_1 .

Moreover, Lemmas C.4 (with y_t in that lemma corresponding to y_t^* here) and E.2 imply that

$$Q_{-1}^\top (\hat{\Sigma}_{\hat{y}^*} - \bar{\Sigma}_{y^*}) Q_{-1} = O_p(T^{-1/2}).$$

Therefore, we can use the same arguments as in the proof of Lemma C.6 to obtain

$$\hat{A}_0 \hat{H}_0^\top - A_0 = A_{0,\perp} (\bar{\sigma}_u^2 \mathbf{I}_{\bar{r}-1} - \Lambda_\perp)^{-1} A_{0,\perp}^\top Q_{-1}^\top (\hat{\Sigma}_{\hat{y}^*} - \bar{\Sigma}_{y^*}) Q_{-1} A_0 + O_p(T^{-1}), \quad (\text{E.1})$$

where $\hat{H}_0$ is nonsingular with $\hat{H}_0^\top \hat{H}_0 = \mathbf{I}_{n-\bar{r}}$ wpa1. Here A_0 , $A_{0,\perp}$, $\Lambda_\perp$, and $\hat{H}_0$ are defined by replacing $\bar{\Sigma}_y$ and $\hat{\Sigma}_y$ in their original definitions with $\bar{\Sigma}_{y^*}$ and $\hat{\Sigma}_{\hat{y}^*}$, respectively.

Since the contribution of the estimation error in β to the asymptotic variance of the GIV estimator arises through $\bar{g}_T(\theta; \hat{A}, \hat{\beta})$, and given the invariance of both the GIV estimator $\hat{\theta}(\hat{A})$ in (63) and the J -statistic, it suffices to study the asymptotic variance of $\bar{g}_T(\theta; \tilde{A}, \hat{\beta})$ in order to assess the effect of the estimation error in β . The asymptotic expansion of $\bar{g}_T(\theta; \tilde{A}, \hat{\beta})$ is given in the following lemma.

Lemma E.1. *Under Assumptions 1, 2, 3 and E.1, we have*

$$\bar{g}_T(\theta; \tilde{A}, \hat{\beta}) = \text{diag}(A^\top, A^\top, \mathbf{I}_{d_w}) T^{-1} \sum_{t \leq T} \bar{\xi}_t + O_p(T^{-1}),$$

where $\bar{\xi}_t \equiv (\bar{\xi}_{v,t}^\top, \bar{\xi}_{\varepsilon,t}^\top, w_t^\top \varepsilon_t)^\top$, and for $b \in \{v, \varepsilon\}$,

$$\bar{\xi}_{b,t} \equiv y_t^* b_t - \mathbb{E}[y_t^* b_t] + (y_t^* y_t^{*\top} - \mathbb{E}[y_t^* y_t^{*\top}]) \Upsilon \left(T^{-1} \sum_{s \leq T} \mathbb{E}[y_s^* b_s] \right). \quad (\text{E.2})$$

PROOF OF LEMMA E.1. The result follows immediately from Lemmas E.2, E.3, and E.4, and is therefore omitted. Q.E.D.

Lemma E.1 shows that the asymptotic variance of $\hat{\theta}(\hat{A})$ is determined by that of $T^{-1/2} \sum_{t \leq T} \bar{\xi}_t$. Compared with ξ_t defined in Assumption 4(i), $\bar{\xi}_t$ contains the additional component $w_t^\top \varepsilon_t$, which arises from the moment conditions used to identify and estimate the coefficients on w_t in the supply equation. More importantly, $\bar{\xi}_{b,t}$ in (E.2) has the same form as $\xi_{b,t}$ in (41), with the only difference being that y there is replaced by y^* here. Therefore, the estimation error in β does not affect the asymptotic variance of $\bar{g}_T(\theta; \tilde{A}, \hat{\beta})$, and hence does not affect that of the GIV estimator $\hat{\theta}(\hat{A})$ in (63). As a result, the standard errors of $\hat{\theta}(\hat{A})$, the optimal weight matrix, and the J -statistic can be constructed as if β were known.

Lemma E.2. *Under Assumptions 1, 2, 3 and E.1, we have*

$$Q_{-1}^\top (\hat{\Sigma}_{\hat{y}^*} - \bar{\Sigma}_{y^*}) Q_{-1} = Q_{-1}^\top T^{-1} \sum_{t \leq T} (y_t^* y_t^{*\top} - \mathbb{E}[y_t^* y_t^{*\top}]) Q_{-1} + O_p(T^{-1}).$$

PROOF OF LEMMA E.2. By the definitions of $\hat{y}_t^*$ and $\hat{y}_{e,t}^*$,

$$\hat{y}_t^* - y_t^* = -x_t(\hat{\beta} - \beta), \quad \hat{y}_{e,t}^* - y_{e,t}^* = -x_{e,t}(\hat{\beta} - \beta), \quad (\text{E.3})$$

where $x_{e,t} \equiv e^\top x_t$. It follows that

$$\begin{aligned} \hat{\Sigma}_{\hat{y}^*} - \bar{\Sigma}_{y^*} &= T^{-1} \sum_{t \leq T} \hat{y}_t^* \hat{y}_t^{*\top} - T^{-1} \sum_{t \leq T} \mathbb{E}[y_t^* y_t^{*\top}] \\ &= T^{-1} \sum_{t \leq T} (y_t^* - x_t(\hat{\beta} - \beta))(y_t^* - x_t(\hat{\beta} - \beta))^\top - T^{-1} \sum_{t \leq T} \mathbb{E}[y_t^* y_t^{*\top}] \\ &= T^{-1} \sum_{t \leq T} (y_t^* y_t^{*\top} - \mathbb{E}[y_t^* y_t^{*\top}]) - T^{-1} \sum_{t \leq T} y_t^* (\hat{\beta} - \beta)^\top x_t^\top \\ &\quad - T^{-1} \sum_{t \leq T} x_t (\hat{\beta} - \beta) y_t^{*\top} + T^{-1} \sum_{t \leq T} x_t (\hat{\beta} - \beta) (\hat{\beta} - \beta)^\top x_t^\top. \end{aligned} \quad (\text{E.4})$$

From Assumption E.1(ii),

$$\hat{\beta} - \beta = O_p(T^{-1/2}). \quad (\text{E.5})$$

By the triangle and Cauchy–Schwarz inequalities,

$$\left\| T^{-1} \sum_{t \leq T} x_t (\hat{\beta} - \beta) (\hat{\beta} - \beta)^\top x_t^\top \right\| \leq \|\hat{\beta} - \beta\|^2 T^{-1} \sum_{t \leq T} \|x_t\|^2 = O_p(T^{-1}), \quad (\text{E.6})$$

where the last equality follows from Assumption E.1(iv) and (E.5).

Let λ_i denote the i th row of λ . By Assumptions E.1(i, iii)

$$\begin{aligned} T^{-1} \sum_{t \leq T} y_{i,t}^* x_t &= T^{-1} \sum_{t \leq T} (\phi p_t + \lambda_i \eta_t + u_{i,t}) x_t \\ &= \phi T^{-1} \sum_{t \leq T} p_t x_t + T^{-1} \sum_{t \leq T} \lambda_i \eta_t x_t + T^{-1} \sum_{t \leq T} u_{i,t} x_t \\ &= \phi T^{-1} \sum_{t \leq T} p_t x_t + O_p(T^{-1/2}). \end{aligned} \quad (\text{E.7})$$

Under Assumptions 2(i, iv) and E.1(iv), the demand and supply equations (60)-(61) admit the following reduced form expressions

$$y_t = \left(\mathbf{I}_n + \frac{\phi\psi}{1 - \phi\psi} \mathbf{1}_n S_t^\top \right) (u_t + x_t \beta + \lambda \eta_t) + \frac{\phi \mathbf{1}_n}{1 - \phi\psi} (\varepsilon_t + w_t^\top \gamma), \quad (\text{E.8})$$

$$p_t = \frac{\psi}{1 - \phi\psi} S_t^\top (u_t + x_t \beta + \lambda \eta_t) + \frac{\varepsilon_t + w_t^\top \gamma}{1 - \phi\psi}. \quad (\text{E.9})$$

From Assumptions 2(i, iv) and E.1(iv), it follows that

$$\mathbb{E}[p_t^2] \leq K.$$

Together with Assumption E.1(iv), the Cauchy–Schwarz inequality, and Markov’s inequality, this implies that

$$T^{-1} \sum_{t \leq T} p_t x_t = O_p(1). \quad (\text{E.10})$$

Combining (E.5), (E.7) and (E.10),

$$T^{-1} \sum_{t \leq T} x_t (\hat{\beta} - \beta) y_t^{*\top} = \phi \left(T^{-1} \sum_{t \leq T} p_t x_t \right) (\hat{\beta} - \beta) \mathbf{1}_n^\top + O_p(T^{-1}). \quad (\text{E.11})$$

Since $\mathbf{1}_n^\top Q_{-1} = \mathbf{0}_{n-1}$, it follows that

$$Q_{-1}^\top \left(T^{-1} \sum_{t \leq T} x_t (\hat{\beta} - \beta) y_t^{*\top} \right) Q_{-1} = O_p(T^{-1}).$$

The same bound holds for its transpose. Combining these bounds with the quadratic term gives the desired result. Q.E.D.

Lemma E.3. *Under Assumptions 1, 2, 3 and E.1, we have*

$$\begin{aligned} T^{-1/2} \sum_{t \leq T} \tilde{A}^\top \hat{y}_t^* (\hat{y}_{e,t}^* - \phi p_t) &= A^\top T^{-1/2} \sum_{t \leq T} (y_t^* v_t - \mathbb{E}[y_t^* v_t]) \\ &\quad + A^\top \left(T^{-1/2} \sum_{t \leq T} (y_t^* y_t^{*\top} - \mathbb{E}[y_t^* y_t^{*\top}]) \right) \Upsilon \left(T^{-1} \sum_{t \leq T} \mathbb{E}[y_t^* v_t] \right) \\ &\quad + O_p(T^{-1/2}). \end{aligned}$$

PROOF OF LEMMA E.3. We begin with the decomposition

$$\begin{aligned} T^{-1/2} \sum_{t \leq T} \tilde{A}^\top \hat{y}_t^* (\hat{y}_{e,t}^* - \phi p_t) &= T^{-1/2} \sum_{t \leq T} A^\top y_t^* (y_{e,t}^* - \phi p_t) + T^{-1/2} \sum_{t \leq T} A^\top y_t^* (\hat{y}_{e,t}^* - y_{e,t}^*) \\ &\quad + A^\top T^{-1/2} \sum_{t \leq T} (\hat{y}_t^* - y_t^*) (\hat{y}_{e,t}^* - \phi p_t) + (\tilde{A} - A)^\top T^{-1/2} \sum_{t \leq T} \hat{y}_t^* (\hat{y}_{e,t}^* - \phi p_t). \end{aligned} \quad (\text{E.12})$$

Since $v_t \equiv y_{e,t}^* - \phi p_t$, by (C.75) we have

$$T^{-1/2} \sum_{t \leq T} A^\top y_t^* (y_{e,t}^* - \phi p_t) = T^{-1/2} \sum_{t \leq T} A^\top y_t^* v_t = A^\top T^{-1/2} \sum_{t \leq T} (y_t^* v_t - \mathbb{E}[y_t^* v_t]). \quad (\text{E.13})$$

We next control the last three terms on the right-hand side of (E.12).

Step 1: Second term. Using $\hat{\beta} - \beta = O_p(T^{-1/2})$ in (E.5) and the same arguments as in (E.7), we have

$$\begin{aligned} T^{-1/2} \sum_{t \leq T} y_t^* (\hat{y}_{e,t}^* - y_{e,t}^*) &= -T^{-1/2} \sum_{t \leq T} y_t^* x_{e,t} (\hat{\beta} - \beta) \\ &= -\mathbf{1}_n \left(\phi T^{-1} \sum_{t \leq T} p_t x_{e,t} \right) T^{1/2} (\hat{\beta} - \beta) + O_p(T^{-1/2}). \end{aligned} \quad (\text{E.14})$$

Since $A^\top \mathbf{1}_n = A_0^\top Q_{-1}^\top \mathbf{1}_n = \mathbf{0}_{n-1}$, it follows that

$$T^{-1/2} \sum_{t \leq T} A^\top y_t^* (\hat{y}_{e,t}^* - y_{e,t}^*) = O_p(T^{-1/2}). \quad (\text{E.15})$$

Step 2: Third term. By Assumptions 2(i) and E.1(i, iii),

$$T^{-1/2} \sum_{t \leq T} x_t (y_{e,t}^* - \phi p_t) = O_p(1).$$

Combining this with $\hat{\beta} - \beta = O_p(T^{-1/2})$ yields

$$T^{-1/2} \sum_{t \leq T} (\hat{y}_t^* - y_t^*) (y_{e,t}^* - \phi p_t) = -T^{-1/2} \sum_{t \leq T} x_t (\hat{\beta} - \beta) (y_{e,t}^* - \phi p_t) = O_p(T^{-1/2}). \quad (\text{E.16})$$

Moreover, by the same arguments as in (E.6),

$$T^{-1/2} \sum_{t \leq T} x_t (\hat{\beta} - \beta) x_{e,t} (\hat{\beta} - \beta) = O_p(T^{-1/2}).$$

Therefore,

$$\begin{aligned} T^{-1/2} \sum_{t \leq T} (\hat{y}_t^* - y_t^*) (\hat{y}_{e,t}^* - \phi p_t) &= -T^{-1/2} \sum_{t \leq T} x_t (\hat{\beta} - \beta) (y_{e,t}^* - \phi p_t) \\ &\quad + T^{-1/2} \sum_{t \leq T} x_t (\hat{\beta} - \beta) x_{e,t} (\hat{\beta} - \beta) = O_p(T^{-1/2}). \end{aligned} \quad (\text{E.17})$$

Hence,

$$A^\top T^{-1/2} \sum_{t \leq T} (\hat{y}_t^* - y_t^*) (\hat{y}_{e,t}^* - \phi p_t) = O_p(T^{-1/2}). \quad (\text{E.18})$$

Step 3: Fourth term. To study the last term, we first use (E.17) to obtain

$$\begin{aligned} T^{-1/2} \sum_{t \leq T} \hat{y}_t^* (\hat{y}_{e,t}^* - \phi p_t) &= T^{-1/2} \sum_{t \leq T} y_t^* (\hat{y}_{e,t}^* - \phi p_t) + T^{-1/2} \sum_{t \leq T} (\hat{y}_t^* - y_t^*) (\hat{y}_{e,t}^* - \phi p_t) \\ &= T^{-1/2} \sum_{t \leq T} y_t^* v_t + T^{-1/2} \sum_{t \leq T} y_t^* (\hat{y}_{e,t}^* - y_{e,t}^*) + O_p(T^{-1/2}), \end{aligned} \quad (\text{E.19})$$

where $v_t \equiv y_{e,t}^* - \phi p_t$. From (E.10), we have

$$T^{-1} \sum_{t \leq T} p_t x_{e,t} = O_p(1), \quad (\text{E.20})$$

which, together with (E.5), implies

$$\mathbf{1}_n \left(\phi T^{-1} \sum_{t \leq T} p_t x_{e,t} \right) T^{1/2} (\hat{\beta} - \beta) = O_p(1). \quad (\text{E.21})$$

From Lemmas C.4 and C.6 (with y_t in that lemma corresponding to y_t^* here), and E.2, we have

$$T^{1/2}(\tilde{A} - A) = \Upsilon \left(T^{-1/2} \sum_{t \leq T} (y_t^* y_t^{*\top} - \mathbb{E}[y_t^* y_t^{*\top}]) \right) A + O_p(T^{-1/2}) = O_p(1), \quad (\text{E.22})$$

where the second equality follows from (D.19) and Lemma C.4.

Using $\Upsilon \mathbf{1}_n = 0$, together with (E.14) and (E.21), we obtain

$$(\tilde{A} - A)^\top T^{-1/2} \sum_{t \leq T} y_t^* (\hat{y}_{e,t}^* - y_{e,t}^*) = O_p(T^{-1/2}). \quad (\text{E.23})$$

By (D.2) and (C.70) (with y_t there corresponding to y_t^* here), we have

$$T^{-1} \sum_{t \leq T} \mathbb{E}[y_t^* v_t] = O(1), \quad T^{-1} \sum_{t \leq T} (y_t^* v_t - \mathbb{E}[y_t^* v_t]) = O_p(1). \quad (\text{E.24})$$

From (E.22) and (E.24), it follows that

$$\begin{aligned} (\tilde{A} - A)^\top T^{-1/2} \sum_{t \leq T} y_t^* v_t &= A^\top \left(T^{-1/2} \sum_{t \leq T} (y_t^* y_t^{*\top} - \mathbb{E}[y_t^* y_t^{*\top}]) \right) \Upsilon \left(T^{-1} \sum_{t \leq T} \mathbb{E}[y_t^* v_t] \right) \\ &\quad + O_p(T^{-1/2}). \end{aligned} \quad (\text{E.25})$$

Combining (E.19), (E.23) and (E.25), we obtain

$$\begin{aligned} (\tilde{A} - A)^\top T^{-1/2} \sum_{t \leq T} \hat{y}_{e,t}^* (\hat{y}_{e,t}^* - \phi p_t) &= A^\top \left(T^{-1/2} \sum_{t \leq T} (y_t^* y_t^{*\top} - \mathbb{E}[y_t^* y_t^{*\top}]) \right) \Upsilon \left(T^{-1} \sum_{t \leq T} \mathbb{E}[y_t^* v_t] \right) \\ &\quad + O_p(T^{-1/2}). \end{aligned} \quad (\text{E.26})$$

The result then follows from (E.12), (E.13), (E.15), (E.18), and (E.26). *Q.E.D.*

Lemma E.4. *Under Assumptions 1, 2, 3 and E.1, we have*

$$\begin{aligned} T^{-1/2} \sum_{t \leq T} \tilde{A}^\top \hat{y}_t^* \varepsilon_t &= A^\top T^{-1/2} \sum_{t \leq T} (y_t^* \varepsilon_t - \mathbb{E}[y_t^* \varepsilon_t]) \\ &\quad + A^\top \left(T^{-1/2} \sum_{t \leq T} (y_t^* y_t^{*\top} - \mathbb{E}[y_t^* y_t^{*\top}]) \right) \Upsilon \left(T^{-1} \sum_{t \leq T} \mathbb{E}[y_t^* \varepsilon_t] \right) \\ &\quad + O_p(T^{-1/2}). \end{aligned}$$

PROOF OF LEMMA E.4. We begin with the decomposition

$$T^{-1/2} \sum_{t \leq T} \tilde{A}^\top \hat{y}_t^* \varepsilon_t = T^{-1/2} \sum_{t \leq T} A^\top y_t^* \varepsilon_t + (\tilde{A} - A)^\top T^{-1/2} \sum_{t \leq T} y_t^* \varepsilon_t + T^{-1/2} \sum_{t \leq T} \tilde{A}^\top (\hat{y}_t^* - y_t^*) \varepsilon_t. \quad (\text{E.27})$$

By arguments similar to those used to establish (C.75), we have

$$T^{-1/2} \sum_{t \leq T} A^\top \mathbb{E}[y_t^* \varepsilon_t] = \mathbf{0}_{n-\bar{r}}.$$

Therefore,

$$T^{-1/2} \sum_{t \leq T} A^\top y_t^* \varepsilon_t = A^\top T^{-1/2} \sum_{t \leq T} (y_t^* \varepsilon_t - \mathbb{E}[y_t^* \varepsilon_t]). \quad (\text{E.28})$$

We next study the last term on the right-hand side of (E.27). By definition,

$$T^{-1/2} \sum_{t \leq T} \tilde{A}^\top (\hat{y}_t^* - y_t^*) \varepsilon_t = -A^\top T^{-1/2} \sum_{t \leq T} x_t \varepsilon_t (\hat{\beta} - \beta) - (\tilde{A} - A)^\top T^{-1/2} \sum_{t \leq T} x_t \varepsilon_t (\hat{\beta} - \beta), \quad (\text{E.29})$$

where we use $\hat{y}_t^* - y_t^* = -x_t(\hat{\beta} - \beta)$. From Assumption E.1(iii), together with the condition $T^{-1} \sum_{t \leq T} \mathbb{E}[x_t \varepsilon_t] = 0$, it follows that

$$T^{-1/2} \sum_{t \leq T} x_t \varepsilon_t = O_p(1). \quad (\text{E.30})$$

This, combined with (E.5), (E.22), and (E.29), implies that

$$T^{-1/2} \sum_{t \leq T} \tilde{A}^\top (\hat{y}_t^* - y_t^*) \varepsilon_t = O_p(T^{-1/2}). \quad (\text{E.31})$$

Next, by arguments similar to those used in (E.24), we have

$$T^{-1} \sum_{t \leq T} \mathbb{E}[y_t^* \varepsilon_t] = O(1), \quad T^{-1} \sum_{t \leq T} (y_t^* \varepsilon_t - \mathbb{E}[y_t^* \varepsilon_t]) = O_p(1).$$

Combining this with (E.22) yields

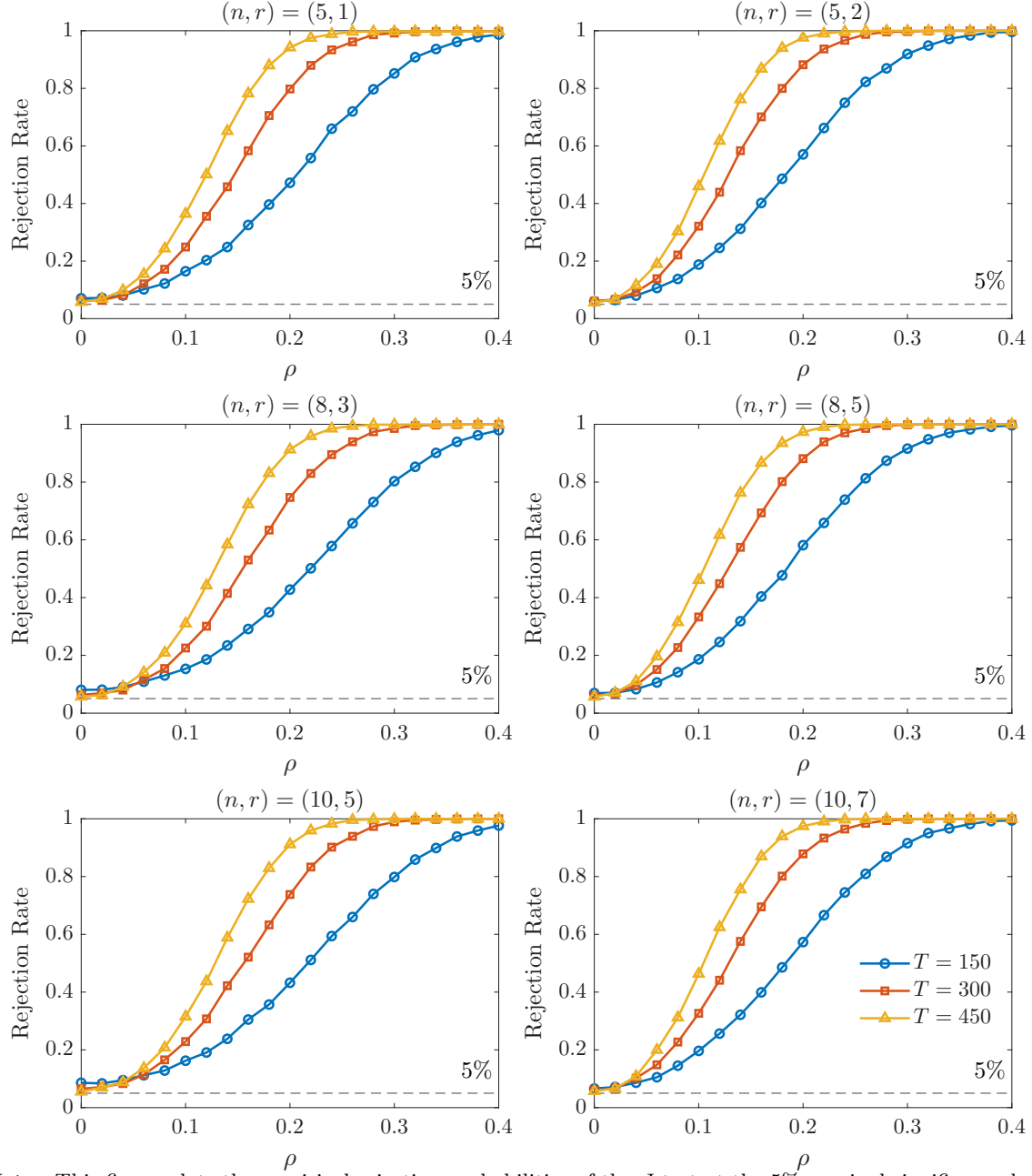
$$(\tilde{A} - A)^\top T^{-1/2} \sum_{t \leq T} y_t^* \varepsilon_t = A^\top \left(T^{-1/2} \sum_{t \leq T} (y_t^* y_t^{*\top} - \mathbb{E}[y_t^* y_t^{*\top}]) \right) \Upsilon \left(T^{-1} \sum_{t \leq T} \mathbb{E}[y_t^* \varepsilon_t] \right) + O_p(T^{-1/2}). \quad (\text{E.32})$$

Substituting (E.28), (E.31) and (E.32) into (E.27) completes the proof. *Q.E.D.*

F Additional Simulation Results

This appendix reports two additional sets of simulation results related to Section 5: the power of the feasible J -test in the extended design and the performance of the BIC criterion in estimating the number of latent factors.

Figure F.1: Empirical Power of the J -test: Extended Design ($d_x = 3$)



Notes. This figure plots the empirical rejection probabilities of the J -test at the 5% nominal significance level in the extended design with $d_x = 3$ exogenous regressors, based on 10,000 Monte Carlo replications. Each panel corresponds to one of the six configurations (n, r) listed in (74), and within each panel the three curves correspond to the sample sizes $T \in \{150, 300, 450\}$. The horizontal axis is ρ , while the vertical axis reports the rejection rate. The horizontal dashed line indicates the nominal 5% level. Power is reported for the feasible GIV estimator, which selects $\bar{r}$ using the BIC criterion in (47)–(48) and is implemented according to Algorithm 1 in Appendix A. In the extended design, the estimator additionally partials out $x_t\beta$ using OLS.

Table F.1: BIC Selection Diagnostics for $\hat{r}$.

(n, r)	T	Baseline ($d_x = 0$)			Extended ($d_x = 3$)		
		$\text{mean}(\hat{r})$	$\text{mode}(\hat{r})$	$\mathbb{P}(\hat{r} = \bar{r})$	$\text{mean}(\hat{r})$	$\text{mode}(\hat{r})$	$\mathbb{P}(\hat{r} = \bar{r})$
(5, 1)	150	2.008	2	0.994	2.008	2	0.994
	300	2.010	2	0.992	2.010	2	0.992
	450	2.009	2	0.993	2.010	2	0.991
(5, 2)	150	3.020	3	0.980	3.022	3	0.978
	300	3.021	3	0.979	3.021	3	0.979
	450	3.021	3	0.979	3.017	3	0.983
(8, 3)	150	4.000	4	1.000	4.000	4	1.000
	300	4.002	4	0.999	4.001	4	0.999
	450	4.002	4	0.998	4.002	4	0.998
(8, 5)	150	6.025	6	0.975	6.025	6	0.975
	300	6.022	6	0.978	6.022	6	0.978
	450	6.020	6	0.980	6.020	6	0.980
(10, 5)	150	6.001	6	0.999	6.001	6	1.000
	300	6.002	6	0.999	6.001	6	0.999
	450	6.003	6	0.998	6.002	6	0.998
(10, 7)	150	8.026	8	0.974	8.027	8	0.974
	300	8.024	8	0.976	8.022	8	0.978
	450	8.022	8	0.978	8.023	8	0.977

Notes. This table reports BIC selection diagnostics for the rank estimator $\hat{r}$, based on 10,000 Monte Carlo replications for each design configuration. Results are presented separately for the baseline design ($d_x = 0$) and the extended design with $d_x = 3$ exogenous regressors. The columns $\text{mean}(\hat{r})$, $\text{mode}(\hat{r})$, and $\mathbb{P}(\hat{r} = \bar{r})$ report the average value, modal value, and probability of correct selection of the BIC estimator $\hat{r}$ defined in (47)–(48), where $\bar{r} = r + 1$ denotes the true rank of $(\mathbf{1}_n, \lambda)$.

F.1 Power of the J -test in the Extended Design

In the extended design with $d_x = 3$ exogenous regressors, the feasible J -test exhibits behavior similar to that in the baseline design shown in Figure 1: the power increases monotonically with ρ , rises substantially as the sample size T increases, and approaches one by $\rho = 0.4$ across all configurations (see Figure F.1 for details). Partialling out the exogenous regressors x_t has only a minor effect on power. Relative to the baseline design, the power curves in the extended design lie only slightly below their baseline counterparts at intermediate values of ρ , reflecting the additional estimation error from estimating β .

F.2 BIC Rank Selection

The feasible GIV estimation and inference procedures determine the rank $\bar{r}$ using the BIC estimator proposed in (47)–(48). Table F.1 shows that the BIC criterion performs remarkably well in estimating the true rank across all configurations of (n, r) in both the baseline and extended designs. In particular, the modal value of $\hat{r}$ always coincides with the true rank $\bar{r} = r + 1$, and the average value of $\hat{r}$ is extremely close to $\bar{r}$ in every case.

The probability of correct selection is uniformly high across all configurations, typically exceeding 97% and often above 99%. For example, in the baseline design with $(n, r) = (5, 1)$, the correct selection probabilities are 0.994, 0.992, and 0.993 at $T = 150, 300$, and 450, respectively. Even in configurations with a larger number of latent factors, such as $(n, r) = (10, 7)$, the correct selection probabilities remain close to 98% in both designs. Moreover, the performance of the BIC criterion is nearly identical in the baseline and extended designs, indicating that the additional estimation of β in the extended model has little effect on rank determination.

Overall, these results suggest that the proposed BIC procedure provides a highly reliable estimator of the rank $\bar{r}$, supporting the practical implementation of the feasible GIV estimation and inference procedures.

G Additional Empirical Results

G.1 Variable Construction

We construct the sector-level holding growth $\Delta q_{i,t}$ and the aggregate price change Δp_t following Appendix C of Gabaix and Koijen (2023).

The Financial Accounts report the dollar value of each sector’s U.S. equity holdings, denoted by $w_{i,t}$. Changes in $w_{i,t}$ reflect both capital appreciation of existing holdings and net purchases or sales of equities. To isolate the latter component, let $q_{i,t}$ denote the number of shares held by sector i and let P_t denote the aggregate equity price. Since $w_{i,t} = q_{i,t}P_t$, a sector that does not trade between $t - 1$ and t would end the quarter holding $w_{i,t-1}R_t$ dollars of equity, where $R_t = P_t/P_{t-1}$ is the gross capital-appreciation return on the aggregate market. It follows that

$$\Delta q_{i,t} \equiv \frac{q_{i,t}}{q_{i,t-1}} - 1 = \frac{w_{i,t}/P_t}{w_{i,t-1}/P_{t-1}} - 1 = \frac{w_{i,t}}{w_{i,t-1}R_t} - 1,$$

which is the empirical measure used in (78). Define $\Delta f_{i,t} \equiv (w_{i,t} - w_{i,t-1}R_t)/w_{i,t-1}$. We can equivalently write $\Delta q_{i,t} = \Delta f_{i,t}/R_t$, which corresponds to the flow-based measure reported in Appendix C.1.2 of Gabaix and Koijen (2023). The gross capital-appreciation return R_t is constructed from the CRSP value-weighted index excluding dividends and compounded from monthly to quarterly frequency.

The aggregate price change Δp_t is measured by the quarterly simple return on the CRSP value-weighted index including dividends, compounded from monthly returns.

G.2 Sector Construction and Weights

We construct the sector classification from the corporate-equity holdings reported in Financial Accounts Table L.224, following the data construction in Appendix C of Gabaix and Koijen (2023). The resulting cross-section consists of twelve equity-holding sectors, listed in Table G.1. These sectors correspond closely to the investor categories used in Gabaix and Koijen (2023).

For each sector, we compute its time- t market share as

$$S_{i,t} \equiv \frac{w_{i,t-1}}{\sum_{j \leq 12} w_{j,t-1}}. \quad (\text{G.1})$$

Thus, $S_{i,t}$ measures sector i 's share of the total equity holdings of the twelve modeled sectors and satisfies $\sum_{i=1}^{12} S_{i,t} = 1$. Table G.1 reports the average market share, $\bar{S}_i = T^{-1} \sum_{t \leq T} S_{i,t}$, computed over the benchmark sample 1993Q1–2018Q4.

Table G.1: Institutional Investor Sectors Holding U.S. Equities

Rank	Sector	Avg. share (%)	Cum. share (%)	Cross-section
1	Households	41.63	41.63	6, 12
2	Mutual funds and ETFs	23.42	65.05	6, 12
3	Foreign sector	13.00	78.05	6, 12
4	Private pension funds	9.92	87.97	6, 12
5	State and local pension funds	7.54	95.51	6, 12
6	Life insurance companies	1.89	97.40	6, 12
7	Property and casualty insurers	1.17	98.57	12
8	Federal government retirement funds	0.44	99.01	12
9	State and local governments	0.43	99.44	12
10	Closed-end funds	0.38	99.82	12
11	Banks	0.12	99.94	12
12	Broker-dealers	0.06	100.00	12

Notes. The table lists the twelve institutional investor sectors holding U.S. equities in the Financial Accounts of the United States (Table L.224), ranked by average market share over 1993Q1–2018Q4. “Avg. share” denotes the sector’s average share of aggregate equity holdings (in percent), and the shares sum to 100. “Cum. share” denotes the cumulative share accounted for by the sector and all sectors ranked above it. “Cross-section” identifies the estimation samples that include the sector: $n = 6$ corresponds to the granular core formed by the six largest sectors, while $n = 12$ corresponds to the full twelve-sector panel.

We rank sectors according to their average market shares $\bar{S}_i$. The six largest sectors jointly account for 97.4% of total equity holdings and constitute the granular core ($n = 6$): households,

mutual funds and ETFs, the foreign sector, private pension funds, state and local pension funds, and life insurance companies. The remaining six sectors each account for less than 1.2% of total equity holdings and, together with the granular core, form the full twelve-sector panel ($n = 12$).

G.3 Implementation of the Factor-residual IV Estimator

This subsection describes our implementation of the FIV estimator of Gabaix and Koijen (2023). To facilitate comparison with their results, we follow the estimation procedure in Section 4.2 of their paper as closely as possible.

ALGORITHM 3 (FIV ESTIMATION)

Step 0. For each sector i , let $\hat{\sigma}_{q,i}^2$ denote the time-series sample variance of $\Delta q_{i,t}$. Construct the precision weights

$$E_i = \frac{1/\bar{\sigma}_{q,i}^2}{\sum_{j \leq n} 1/\bar{\sigma}_{q,j}^2}, \quad \bar{\sigma}_{q,i}^2 = \max\{\hat{\sigma}_{q,i}^2, \text{median}_{j \leq n} \hat{\sigma}_{q,j}^2\}.$$

Step 1. Estimate the two-way fixed-effects regression by weighted least squares:

$$(\hat{\alpha}, \hat{\gamma}, \hat{\beta}) = \arg \min_{\alpha_i, \gamma_t, \beta_i} \sum_{t \leq T} \sum_{i \leq n} E_i \left(\Delta q_{i,t} - \alpha_i - \gamma_t - \beta_i^\top \eta_{o,t} \right)^2,$$

where α_i and γ_t are sector and time fixed effects, respectively, and $\eta_{o,t}$ denotes the vector of observed factors with sector-specific loadings β_i . The residual is

$$\Delta \check{q}_{i,t} \equiv \Delta q_{i,t} - \hat{\alpha}_i - \hat{\gamma}_t - \hat{\beta}_i^\top \eta_{o,t}.$$

Step 2. Extract r latent factors, denoted by $\hat{\eta}_{l,t}$, as the principal components of the weighted residuals $E_i^{1/2} \Delta \check{q}_{i,t}$. Given the estimated factors $\hat{\eta}_{l,t}$, estimate the corresponding factor loadings $\hat{\lambda}_i$ and recover the idiosyncratic shock $\check{u}_{i,t}$ from the regression

$$\Delta \check{q}_{i,t} = \hat{\lambda}_i^\top \hat{\eta}_{l,t} + \check{u}_{i,t}.$$

Step 3. Construct the FIV as the size-weighted sum of the estimated idiosyncratic shocks

$$z_t = \sum_{i \leq n} S_{i,t} \check{u}_{i,t},$$

where $S_{i,t}$ is defined in (G.1), and estimate the market multiplier from the time-series regression

$$\Delta p_t = \alpha + \kappa z_t + \Gamma^\top \eta_t + e_t, \quad \eta_t = (\eta_{o,t}^\top, \hat{\eta}_{l,t}^\top)^\top \quad (\text{G.2})$$

where Δp_t denotes the aggregate market return and κ is the aggregate market multiplier.

Step 4. Recompute the precision weights using the time-series sample variance of the estimated idiosyncratic shocks $\check{u}_{i,t}$ and repeat Steps 1–3 once. Report the coefficient estimates and their

Newey–West standard errors from the time-series regression in (G.2), without adjusting for the estimation error in z_t and η_t .

In our empirical implementation of the FIV estimator, the observed factors $\eta_{o,t}$ consist of GDP growth together with a set of asset-pricing factors. Specifically, following Gabaix and Koijen (2023), we run monthly cross-sectional regressions of stock returns on log market capitalization, the log book-to-market ratio, and a momentum signal. Individual stock returns, prices, and shares outstanding are obtained from CRSP, while book equity used to construct the book-to-market ratio is obtained from Compustat, matching the data inputs used by Gabaix and Koijen (2023). GDP growth is computed from real gross domestic product (series GDPC1), produced by the U.S. Bureau of Economic Analysis and retrieved from the Federal Reserve Economic Data (FRED) database maintained by the Federal Reserve Bank of St. Louis. The resulting Fama–MacBeth slope estimates are aggregated to the quarterly frequency and used as the observed size, value, and momentum factors.

References

- BELLONI, A., V. CHERNOZHUKOV, D. CHETVERIKOV, AND K. KATO (2015): “Some New Asymptotic Theory for Least Squares Series: Pointwise and Uniform Results,” *Journal of Econometrics*, 186(2), 345 – 366, High Dimensional Problems in Econometrics.
- GABAIX, X., AND R. S. KOIJEN (2023): “In search of the origins of financial fluctuations: The inelastic markets hypothesis,” SSRN Working Paper No. 3686935.
- YU, Y., T. WANG, AND R. J. SAMWORTH (2015): “A useful variant of the Davis–Kahan theorem for statisticians,” *Biometrika*, 102(2), 315–323.