Supplemental Appendix for "Jackknife Estimation of a Cluster-Sample IV Regression Model with Many Weak Instruments" [1]

John C. Chao, Norman R. Swanson, and Tiemen Woutersen

First Draft: September 5, 2020 *This Version*: September 15, 2021

Abstract

This Supplemental Appendix is comprised of two sub-appendices. Appendix S1 provides proofs for Theorems 2 and 3 of the main paper. Appendix S2 states additional supporting lemmas used to prove the main theorems of the paper. Proofs for these additional lemmas are reported in a separate Online Appendix which can be viewed at the URL:
http://econweb.umd.edu/~chao/Research/research_files/
Additional_Online_Appendix_Jackknife_Estimation_Cluster_Sample_IV_Model.pdf

Appendix S1: Proof of Theorems 2 and 3

Proof of Theorem 2: Define $\mathcal{Y}_n = \Gamma' M^{(Z_1,Q)}\varepsilon/\sqrt{n} + D_\mu^{-1}\underline{U}'A\varepsilon$. Note that, by the result of Lemma S2-9 given in Appendix S2 below, we have that $D_\mu^{-1}\widehat{\Delta}(\delta_0) = \Gamma' M^{(Z_1,Q)}\varepsilon/\sqrt{n} + D_\mu^{-1}\underline{U}'A\varepsilon + o_p(1) = \mathcal{Y}_n + o_p(1)$.

We now establish the asymptotic normality of $\mathcal{Y}_n$, upon appropriate standardization, in the case where $K_{2,n}/\left(\mu_n^{\min}\right)^2 = O(1)$. To proceed, let $a \in \mathbb{R}^d$ such that $\|a\|_2 = 1$ and define $b_{1n} = \Sigma_n^{-1/2}a$. and $b_{2n} = \sqrt{K_{2,n}}D_\mu^{-1}\Sigma_n^{-1/2}a$. Now, let $\mathcal{L}_{(i,t),n} = b_{1n}'\Gamma' M^{(Z_1,Q)}e_{(i,t)}\varepsilon_{(i,t)}/\sqrt{n}$ and $\mathcal{N}_{(i,t),n} = K_{2,n}^{-1/2}\sum_{(j,s)=1}^{(i,t)-1} A_{(i,t),(j,s)}\left[\underline{u}_{2,(i,t),n}\varepsilon_{(j,s)} + \underline{u}_{2,(j,s),n}\varepsilon_{(i,t)}\right]$, where $\underline{u}_{2,(i,t),n} = b_{2n}'\underline{U}_{(i,t)}$, with $\underline{u}_{2,(j,s),n}$ similarly defined, and where $e_{(i,t)}$ denotes an $m_n \times 1$ elementary vector whose $(i,t)^{th}$ component is 1 and all other components are 0. In addition, write, as in the proof of part (d) of Lemma S2-3[2], $\Sigma_n = VC\left(\mathcal{Y}_n|\mathcal{F}_n^W\right) = \Sigma_{1,n} + \Sigma_{2,n}$, where $\Sigma_{1,n} = VC\left(\Gamma' M^{(Z_1,Q)}\varepsilon/\sqrt{n}|\mathcal{F}_n^W\right)$ and $\Sigma_{2,n} = VC\left(D_\mu^{-1}\underline{U}'A\varepsilon|\mathcal{F}_n^W\right)$ as previously defined. Using these notations, note that we can write

[1] Corresponding author: John C. Chao, Department of Economics, 7343 Preinkert Drive, University of Maryland, chao@econ.umd.edu. Norman R. Swanson, Department of Economics, 9500 Hamilton Street, Rutgers University, nswanson@econ.rutgers.edu. Tiemen Woutersen, Department of Economics, 1130 E Helen Street, University of Arizona, woutersen@arizona.edu. The authors owe special thanks to Jerry Hausman and Whitney Newey for many discussions on the topic of this paper over a number of years. In addition, thanks are owed to all of the particpants of the 2019 MIT Conference Honoring Whitney Newey for comments on and advice given about the topic of this paper. Finally, the authors wish to thank Miriam Arden for excellent research assistance. Chao thanks the University of Maryland for research support, and Woutersen's work was supported by an Eller College of Management Research Grant.

[2] A proof of Lemma S2-3 is given in section 1 of the Additional Online Appendix which can be viewed at the URL:
http://econweb.umd.edu/~chao/Research/research_files/
Additional_Online_Appendix_Jackknife_Estimation_Cluster_Sample_IV_Model.pdf

$a'\Sigma_n^{-1/2}\mathcal{Y}_n = \mathcal{L}_{(1,1),n} + \sum_{(i,t)=2}^{m_n} \left\{\mathcal{L}_{(i,t),n} + \mathcal{N}_{(i,t),n}\right\}$. Next, observe that

$$\begin{aligned} E\left[\mathcal{L}_{(1,1),n}^2|\mathcal{F}_n^W\right] &= E\left[\varepsilon_{(1,1)}^2|\mathcal{F}_n^W\right]\frac{\left[a'\Sigma_n^{-1/2}\Gamma' M^{(Z_1,Q)}e_{(1,1)}\right]^2}{n} \\ &\leq E\left[\varepsilon_{(1,1)}^2|\mathcal{F}_n^W\right]a'\Sigma_n^{-1}a\left(\frac{\left\|\Gamma' M^{(Z_1,Q)}e_{(1,1)}\right\|_2}{\sqrt{n}}\right)^2 \text{ (by CS inequality)} \\ &\leq \left(\max_{1\leq(i,t)\leq m_n} E\left[\varepsilon_{(i,t)}^2|\mathcal{F}_n^W\right]\right)a'\Sigma_n^{-1}a\left(\frac{\max_{1\leq(i,t)\leq m_n}\left\|\Gamma' M^{(Z_1,Q)}e_{(i,t)}\right\|_2}{\sqrt{n}}\right)^2 \\ &= o_p(1) \text{ (by Assumptions 2(i) and 7(iv) and part (d) of Lemma S2-3)} \end{aligned}$$

Moreover, under Assumptions 2 and 3(iii), there exists a positive constant C^* such that

$$\begin{aligned} E_{W_n}\left\{\left(E\left[\mathcal{L}_{(1,1),n}^2|\mathcal{F}_n^W\right]\right)^2\right\} &= \frac{E_{W_n}\left\{\left[a'\Sigma_n^{-1/2}\Gamma' M^{(Z_1,Q)}e_{(1,1)}\right]^4\left(E\left[\varepsilon_{(1,1)}^2|\mathcal{F}_n^W\right]\right)^2\right\}}{n^2} \\ &\leq \frac{C}{n^2}E\left(\left[a'\Sigma_n^{-1/2}\Gamma' M^{(Z_1,Q)}e_{(1,1)}\right]^4\right) \text{ (by Assumption 2(i))} \\ &\leq CE\left(\frac{a'\Sigma_n^{-1/2}\Gamma' M^{(Z_1,Q)}\Gamma\Sigma_n^{-1/2}a}{n}\right)^2 \text{ (by CS inequality)} \\ &\leq C\overline{C} = C^* < \infty \text{ (by Assumption 3(iii) and Lemma S2-3(d))} \end{aligned}$$

Since the upper bound above does not depend on n, we further deduce that $\sup_n E_{W_n}\left\{\left(E\left[\mathcal{L}_{(1,1),n}^2|\mathcal{F}_n^W\right]\right)^2\right\} < \infty$. It follows by the law of iterated expectations and by Theorem 25.12 of Billingsley (1995) that $E\left(\mathcal{L}_{(1,1),n}^2\right) = E_{W_n}\left(E\left[\mathcal{L}_{(1,1),n}^2|\mathcal{F}_n^W\right]\right) \to 0$. Application of Markov's inequality then allows us to deduce that $\mathcal{L}_{(1,1),n} = b_{1n}'\Gamma' M^{(Z_1,Q)}e_{(1,1)}\varepsilon_{(1,1)}/\sqrt{n}$ $= o_p(1)$, from which we obtain the representation $a'\Sigma_n^{-1/2}\mathcal{Y}_n = \mathcal{V}_n + o_p(1)$, where $\mathcal{V}_n = \sum_{(i,t)=2}^{m_n}\mathcal{V}_{(i,t),n}$ with $\mathcal{V}_{(i,t),n} = \mathcal{L}_{(i,t),n} + \mathcal{N}_{(i,t),n}$. Note we can also write $\mathcal{V}_n = \mathcal{L}_n + \mathcal{N}_n$, where $\mathcal{L}_n = \sum_{(i,t)=2}^{m_n}\mathcal{L}_{(i,t),n}$ and $\mathcal{N}_n = \sum_{(i,t)=2}^{m_n}\mathcal{N}_{(i,t),n}$.

Next, define the σ-fields $\mathcal{F}_{(i,t),n} = \sigma\left(\left\{\varepsilon_{(k,v)}, U_{(k,v)}\right\}_{(k,v)=1}^{(i,t)}, W_n\right)$ for $(i,t) = 1, 2, ..., m_n$, note that by construction $\mathcal{F}_{(i,t)-1,n} \subseteq \mathcal{F}_{(i,t),n}$ for $(i,t) = 2, ..., m_n$ and $\mathcal{V}_{(i,t),n}$ is $\mathcal{F}_{(i,t),n}$-measurable. Note also that, under Assumption 1, it is easily seen that $E\left[\mathcal{V}_{(i,t),n}|\mathcal{F}_{(i,t)-1,n}\right] = 0$. In addition, note that, by part (d) of Lemma S2-3 and Lemma S2-6, and Assumption 2(i);

$$\begin{aligned} E\left[\underline{u}_{2,(i,t),n}^2|\mathcal{F}_n^W\right] &\leq (b_{2n}'b_{2n})\max_{1\leq(i,t)\leq m_n} E\left[\left\|\underline{U}_{(i,t)}\right\|_2^2|\mathcal{F}_n^W\right] \\ &\leq \frac{K_{2,n}}{(\mu_n^{\min})^2}a'\Sigma_n^{-1}a\max_{1\leq(i,t)\leq m_n} E\left[\left\|\underline{U}_{(i,t)}\right\|_2^2|\mathcal{F}_n^W\right] = O_{a.s.}(1) \end{aligned} \tag{1}$$

since, for this theorem, we assume that $K_{2,n}/\left(\mu_n^{\min}\right)^2 = O(1)$. It follows then from straightforward calculations, from applying the triangle and CS inequalities, as well as from expression (1), part (d) of Lemma S2-1, part (d) of Lemma S2-3, and Assumptions 2(i) and 3(iii) that there exists a positive constant $\overline{C}$ such that

$$\begin{aligned}
& Var\left(\mathcal{V}_{(i,t),n}|\mathcal{F}_n^W\right) \\
= \; & E\left[\mathcal{L}_{(i,t),n}^2|\mathcal{F}_n^W\right] + E\left[\mathcal{N}_{(i,t),n}^2|\mathcal{F}_n^W\right] \\
\leq \; & \left(\max_{1\leq(i,t)\leq m_n} E\left[\varepsilon_{(i,t)}^2|\mathcal{F}_n^W\right]\right) a'\Sigma_n^{-1}a\lambda_{\max}\left(\frac{\Gamma'\Gamma}{n}\right) \\
& + \frac{4}{K_{2,n}}\left(\max_{1\leq(i,t)\leq m_n} E\left[\varepsilon_{(i,t)}^2|\mathcal{F}_n^W\right]\right)\left(\max_{1\leq(i,t)\leq m_n} E\left[\underline{u}_{2,(i,t),n}^2|\mathcal{F}_n^W\right]\right)\left(\max_{1\leq(i,t)\leq m_n}\sum_{(j,s)=1}^{m_n} A_{(i,t),(j,s)}^2\right) \\
= \; & O_{a.s.}(1) + O_{a.s.}\left(\frac{K_{2,n}}{\left(\mu_n^{\min}\right)^2 n}\right) = O_{a.s.}(1)
\end{aligned}$$

By the law of iterated expectations and Theorem 16.1 of Billingsley (1995), there exists a constant $\overline{C}$ such that $Var\left(\mathcal{V}_{(i,t),n}\right) = E\left(\mathcal{V}_{(i,t),n}^2\right) = E_{W_n}\left[E\left(\mathcal{V}_{(i,t),n}^2|\mathcal{F}_n^W\right)\right] \leq \overline{C} < \infty$ for all n sufficiently large. These results show that $\left\{\mathcal{V}_{(i,t),n}, \mathcal{F}_{(i,t),n}, 1 \leq (i,t) \leq m_n, n \geq 1\right\}$ forms a square-integrable martingale difference array.

To show the asymptotic normality of $\mathcal{V}_n$, we verify the conditions of the central limit theorem for martingale difference arrays given in Lemma S2-15. To proceed, first consider condition (22), which, as noted in the remark which follows Lemma S2-15, is a sufficient condition for condition (20) of Lemma S2-15. We shall verify (22) for the case where $\delta = 2$. Note first that, by applying Loève's c_r inequality, we get

$$\sum_{(i,t)=2}^{m_n} E\left[\mathcal{V}_{(i,t),n}^4\right] = \sum_{(i,t)=2}^{m_n} E\left[\left(\mathcal{L}_{(i,t),n} + \mathcal{N}_{(i,t),n}\right)^4\right] \leq 8\sum_{(i,t)=2}^{m_n} E\left[\mathcal{L}_{(i,t),n}^4\right] + 8\sum_{(i,t)=2}^{m_n} E\left[\mathcal{N}_{(i,t),n}^4\right]$$

Hence, to verify condition (22), it suffices to show that $\sum_{(i,t)=2}^{m_n} E\left[\mathcal{L}_{(i,t),n}^4\right] = o(1)$ and $\sum_{(i,t)=2}^{m_n} E\left[\mathcal{N}_{(i,t),n}^4\right] = o(1)$. To do this, we first focus on a conditional expectation analogue of $\sum_{(i,t)=2}^{m_n} E\left[\mathcal{L}_{(i,t),n}^4\right]$. Note that

$$\begin{aligned}
& \sum_{(i,t)=2}^{m_n} E\left[\mathcal{L}_{(i,t),n}^4|\mathcal{F}_n^W\right] \\
= \; & \frac{1}{n^2}\sum_{(i,t)=2}^{m_n}\left[a'\Sigma_n^{-1/2}\Gamma' M^{(Z_1,Q)} e_{(i,t)}\right]^4 E\left[\varepsilon_{(i,t)}^4|\mathcal{F}_n^W\right] \\
\leq \; & a'\Sigma_n^{-1}a\frac{1}{n}\sum_{(i,t)=2}^{m_n}\left[a'\Sigma_n^{-1/2}\Gamma' M^{(Z_1,Q)} e_{(i,t)}\right]^2\left(\frac{\left\|\Gamma' M^{(Z_1,Q)} e_{(i,t)}\right\|_2}{\sqrt{n}}\right)^2 E\left[\varepsilon_{(i,t)}^4|\mathcal{F}_n^W\right] \text{ (by CS inequality)}
\end{aligned}$$

$$
\begin{aligned}
&\leq a'\Sigma_n^{-1}a\left(\max_{1\leq(i,t)\leq m_n}E\left[\varepsilon_{(i,t)}^4|\mathcal{F}_n^W\right]\right)\left(\frac{\max_{1\leq(i,t)\leq m_n}\left\|\Gamma'M^{(Z_1,Q)}e_{(i,t)}\right\|_2}{\sqrt{n}}\right)^2 \\
&\quad\times\frac{1}{n}a'\Sigma_n^{-1/2}\Gamma'M^{(Z_1,Q)}\sum_{(i,t)=1}^{m_n}e_{(i,t)}e_{(i,t)}'M^{(Z_1,Q)}\Gamma\Sigma_n^{-1/2}a \\
&\leq a'\Sigma_n^{-1}a\left(\max_{1\leq(i,t)\leq m_n}E\left[\varepsilon_{(i,t)}^4|\mathcal{F}_n^W\right]\right)\left(\frac{\max_{1\leq(i,t)\leq m_n}\left\|\Gamma'M^{(Z_1,Q)}e_{(i,t)}\right\|_2}{\sqrt{n}}\right)^2 \\
&\quad\times\frac{a'\Sigma_n^{-1/2}\Gamma'M^{(Z_1,Q)}\Gamma\Sigma_n^{-1/2}a}{n} \\
&\leq \left(a'\Sigma_n^{-1}a\right)^2\left(\max_{1\leq(i,t)\leq m_n}E\left[\varepsilon_{(i,t)}^4|\mathcal{F}_n^W\right]\right)\left(\frac{\max_{1\leq(i,t)\leq m_n}\left\|\Gamma'M^{(Z_1,Q)}e_{(i,t)}\right\|_2}{\sqrt{n}}\right)^2\lambda_{\max}\left(\frac{\Gamma'\Gamma}{n}\right) \\
&\leq C\left(\frac{\max_{1\leq(i,t)\leq m_n}\left\|\Gamma'M^{(Z_1,Q)}e_{(i,t)}\right\|_2}{\sqrt{n}}\right)^2=o_p(1)
\end{aligned}
$$

where the last line above follows from Assumptions 2(i), 3(iii), and 7(iv) and by Lemma S2-3(d). Next, note that, under Assumptions 2 and 3(iii), there exists a positive constant C^* such that

$$
\begin{aligned}
&E_{W_n}\left(\sum_{(i,t)=2}^{m_n}E\left[\mathcal{L}_{(i,t),n}^4|\mathcal{F}_n^W\right]\right)^2 \\
&= \frac{1}{n^4}\sum_{(i,t)=2}^{m_n}\sum_{(j,s)=2}^{m_n}E\left(_{W_n}\left[a'\Sigma_n^{-1/2}\Gamma'M^{(Z_1,Q)}e_{(i,t)}\right]^4\left[a'\Sigma_n^{-1/2}\Gamma'M^{(Z_1,Q)}e_{(j,s)}\right]^4\right. \\
&\qquad\left.\times E\left[\varepsilon_{(i,t)}^4|\mathcal{F}_n^W\right]E\left[\varepsilon_{(j,s)}^4|\mathcal{F}_n^W\right]\right) \\
&\leq \frac{C}{n^4}\sum_{(i,t)=2}^{m_n}\sum_{(j,s)=2}^{m_n}E_{W_n}\left(\left[a'\Sigma_n^{-1/2}\Gamma'M^{(Z_1,Q)}e_{(i,t)}\right]^4\left[a'\Sigma_n^{-1/2}\Gamma'M^{(Z_1,Q)}e_{(j,s)}\right]^4\right) \\
&\leq \frac{C}{n^4}E_{W_n}\left\{a'\Sigma_n^{-1/2}\Gamma'M^{(Z_1,Q)}\sum_{(i,t)=1}^{m_n}e_{(i,t)}e_{(i,t)}'M^{(Z_1,Q)}\Gamma\Sigma_n^{-1/2}aa'\Sigma_n^{-1/2}\Gamma'M^{(Z_1,Q)}\right. \\
&\qquad\left.\times\sum_{(j,s)=1}^{m_n}e_{(j,s)}e_{(j,s)}'M^{(Z_1,Q)}\Gamma\Sigma_n^{-1/2}a\left(a'\Sigma_n^{-1/2}\Gamma'M^{(Z_1,Q)}\Gamma\Sigma_n^{-1/2}a\right)^2\right\} \\
&= CE_{W_n}\left(\frac{a'\Sigma_n^{-1/2}\Gamma'M^{(Z_1,Q)}\Gamma\Sigma_n^{-1/2}a}{n}\right)^4 \\
&\leq C\overline{C}=C^*<\infty\ \text{(by Assumption 3(iii) and Lemma S2-3(d))}
\end{aligned}
$$

where the second inequality above follows from applying the CS inequality. Since the upper bound above does not depend on n, we further deduce that

$\sup_n E_{W_n}\left(\sum_{(i,t)=2}^{m_n} E\left[\mathcal{L}_{(i,t),n}^4|\mathcal{F}_n^W\right]\right)^2 < \infty$. It follows by the law of iterated expectations and by Theorem 25.12 of Billingsley (1995) that $\sum_{(i,t)=2}^{m_n} E\left[\mathcal{L}_{(i,t),n}^4\right] = \sum_{(i,t)=2}^{m_n} E_{W_n}\left(E\left[\mathcal{L}_{(i,t),n}^4|\mathcal{F}_n^W\right]\right) \to 0$.

Turning our attention to the bilinear term, note that by Loève's c_r inequality we have $\sum_{(i,t)=2}^{m_n} E\left[\mathcal{N}_{(i,t),n}^4|\mathcal{F}_n^W\right] \leq \mathcal{R}_1 + \mathcal{R}_2$, where $\mathcal{R}_1 = \sum_{(i,t)=2}^{m_n} \left(8/K_{2,n}^2\right) E\left[\left(\sum_{(j,s)=1}^{(i,t)-1} A_{(i,t),(j,s)}\underline{u}_{2,(i,t)}\varepsilon_{(j,s)}\right)^4 |\mathcal{F}_n^W\right]$ and $\mathcal{R}_2 = \sum_{(i,t)=2}^{m_n} \left(8/K_{2,n}^2\right) E\left[\left(\sum_{(j,s)=1}^{(i,t)-1} A_{(i,t),(j,s)}\underline{u}_{2,(j,s)}\varepsilon_{(i,t)}\right)^4 |\mathcal{F}_n^W\right]$. Focusing first on the term $\mathcal{R}_1$, note that, by straightforward calculations as well as by making use of Assumptions 2(i) and 5(ii), parts (b) and (c) of Lemma S2-1, part (d) of Lemma S2-3, and Lemma S2-6; we deduce that, there exists a positive constant $\overline{C}$ such that

$$\begin{aligned}
\frac{\left(\mu_n^{\min}\right)^4 n}{K_{2,n}^2}\mathcal{R}_1 &\leq 24n\left(a'\Sigma_n^{-1}a\right)^2\left(\max_{1\leq(i,t)\leq m_n} E\left[\left\|\underline{U}_{(i,t)}\right\|_2^4|\mathcal{F}_n^W\right]\right)\left(\max_{1\leq(i,t)\leq m_n} E\left[\varepsilon_{(i,t)}^4|\mathcal{F}_n^W\right]\right) \\
&\quad\times\left[\frac{1}{K_{2,n}^2}\sum_{\substack{(i,t),(j,s)=1\\(i,t)\neq(j,s)}}^{m_n} A_{(i,t),(j,s)}^4 + \frac{1}{K_{2,n}^2}\sum_{(i,t)=1}^{m_n}\sum_{\substack{(j,s),(k,v)=1\\(j,s)\neq(i,t),(k,v)\neq(i,t)}}^{m_n} A_{(i,t),(j,s)}^2 A_{(i,t),(k,v)}^2\right] \\
&\leq \overline{C}n\left[\frac{1}{K_{2,n}^2}\sum_{\substack{(i,t),(j,s)=1\\(i,t)\neq(j,s)}}^{m_n} A_{(i,t),(j,s)}^4 + \frac{1}{K_{2,n}^2}\sum_{(i,t)=1}^{m_n}\sum_{\substack{(j,s),(k,v)=1\\(j,s)\neq(i,t),(k,v)\neq(i,t)}}^{m_n} A_{(i,t),(j,s)}^2 A_{(i,t),(k,v)}^2\right] \\
&= O_{a.s.}\left(\frac{K_{2,n}}{n}\right) + O_{a.s.}(1) = O_{a.s.}(1).
\end{aligned}$$

Applying the law of iterated expectations and Theorem 16.1 of Billingsley (1995), we then have

$$\begin{aligned}
&\frac{\left(\mu_n^{\min}\right)^4 n}{K_{2,n}^2}E_{W_n}(\mathcal{R}_1) \\
&\leq \overline{C}nE_{W_n}\left[\frac{1}{K_{2,n}^2}\sum_{\substack{(i,t),(j,s)=1\\(i,t)\neq(j,s)}}^{m_n} A_{(i,t),(j,s)}^4 + \frac{1}{K_{2,n}^2}\sum_{(i,t)=1}^{m_n}\sum_{\substack{(j,s),(k,v)=1\\(j,s)\neq(i,t),(k,v)\neq(i,t)}}^{m_n} A_{(i,t),(j,s)}^2 A_{(i,t),(k,v)}^2\right] \\
&= O(1)
\end{aligned}$$

from which we further deduce that

$$E_{W_n}(\mathcal{R}_1) = \sum_{(i,t)=2}^{m_n}\frac{8}{K_{2,n}^2}E\left[\left(\sum_{(j,s)=1}^{(i,t)-1} A_{(i,t),(j,s)}\underline{u}_{2,(i,t)}\varepsilon_{(j,s)}\right)^4\right] = O\left(\frac{K_{2,n}^2}{\left(\mu_n^{\min}\right)^4 n}\right) = o(1)$$

In a similar way, we can also show that
$E_{W_n}(\mathcal{R}_2) = \left(8/K_{2,n}^2\right) E\left[\sum_{(i,t)=2}^{m_n}\left(\sum_{(j,s)=1}^{(i,t)-1} A_{(j,s),(i,t)}\underline{u}_{2,(j,s)}\varepsilon_{(i,t)}\right)^4\right] = o(1)$. It follows that $\sum_{(i,t)=2}^{m_n} E\left[\mathcal{N}_{(i,t),n}^4\right] \le E_{W_n}(\mathcal{R}_1) + E_{W_n}(\mathcal{R}_2) = o(1)$. This verifies condition (22).

Next, we verify condition (21) of Lemma S2-15. To proceed, first let $s_W^2 = Var\left[\mathcal{V}_n|\mathcal{F}_n^W\right] = Var\left(\sum_{(i,t)=2}^{m_n}\mathcal{V}_{(i,t),n}|\mathcal{F}_n^W\right)$, and note that

$$s_W^2 = Var\left(\frac{b_{1n}'\Gamma' M^{(Z_1,Q)}\varepsilon}{\sqrt{n}} + \frac{b_{2n}'\underline{U}'A\varepsilon}{\sqrt{K_{2,n}}}|\mathcal{F}_n^W\right) + o_p(1) = a'\Sigma_n^{-1/2}\Sigma_n\Sigma_n^{-1/2}a + o_p(1) = 1 + o_p(1) \quad (2)$$

On the other hand, by straightforward calculation, we can write

$$\begin{aligned} s_W^2 &= \frac{1}{n}\sum_{(i,t)=2}^{m_n}\left[b_{1n}'\Gamma' M^{(Z_1,Q)}e_{(i,t)}\right]^2 E\left[\varepsilon_{(i,t)}^2|\mathcal{F}_n^W\right] \\ &+\frac{1}{K_{2,n}}\sum_{(i,t)=2}^{m_n}\sum_{(j,s)=1}^{(i,t)-1} A_{(i,t),(j,s)}^2\left\{E\left[\underline{u}_{2,(i,t)}^2|\mathcal{F}_n^W\right]E\left[\varepsilon_{(j,s)}^2|\mathcal{F}_n^W\right] + E\left[\underline{u}_{2,(j,s)}^2|\mathcal{F}_n^W\right]E\left[\varepsilon_{(i,t)}^2|\mathcal{F}_n^W\right]\right\} \\ &+\frac{2}{K_{2,n}}\sum_{(i,t)=2}^{m_n}\sum_{(j,s)=1}^{(i,t)-1} A_{(i,t),(j,s)}^2 E\left[\underline{u}_{2,(i,t)}\varepsilon_{(i,t)}|\mathcal{F}_n^W\right]E\left[\underline{u}_{2,(j,s)}\varepsilon_{(j,s)}|\mathcal{F}_n^W\right] \qquad (3) \end{aligned}$$

Making use of expression (3), we obtain, after some further calculations,

$$\begin{aligned} &\sum_{(i,t)=2}^{m_n} E\left[\mathcal{V}_{(i,t),n}^2|\mathcal{F}_{(i,t)-1,n}\right] - s_{W_n}^2 \\ =\ & \frac{2}{\sqrt{n}}\sum_{(i,t)=2}^{m_n}\sum_{(j,s)=1}^{(i,t)-1}\left[b_{1n}'\Gamma' M^{(Z_1,Q)}e_{(i,t)}\right]\frac{A_{(i,t),(j,s)}}{\sqrt{K_{2,n}}}\left\{\varepsilon_{(j,s)}E\left[\varepsilon_{(i,t)}\underline{u}_{2,(i,t)}|\mathcal{F}_n^W\right] + \underline{u}_{2,(j,s)}E\left[\varepsilon_{(i,t)}^2|\mathcal{F}_n^W\right]\right\} \\ &+\sum_{(i,t)=2}^{m_n}\sum_{(j,s)=1}^{(i,t)-1}\frac{A_{(i,t),(j,s)}^2}{K_{2,n}}\left(\varepsilon_{(j,s)}^2 - E\left[\varepsilon_{(j,s)}^2|\mathcal{F}_n^W\right]\right)E\left[\underline{u}_{2,(i,t)}^2|\mathcal{F}_n^W\right] \\ &+\sum_{(i,t)=2}^{m_n}\sum_{(j,s)=1}^{(i,t)-1}\frac{A_{(i,t),(j,s)}^2}{K_{2,n}}\left(\underline{u}_{2,(j,s)}^2 - E\left[\underline{u}_{2,(j,s)}^2|\mathcal{F}_n^W\right]\right)E\left[\varepsilon_{(i,t)}^2|\mathcal{F}_n^W\right] \\ &+2\sum_{(i,t)=2}^{m_n}\sum_{(j,s)=1}^{(i,t)-1}\frac{A_{(i,t),(j,s)}^2}{K_{2,n}}\left(\varepsilon_{(j,s)}\underline{u}_{2,(j,s)} - E\left[\underline{u}_{2,(j,s)}\varepsilon_{(j,s)}|\mathcal{F}_n^W\right]\right)E\left[\underline{u}_{2,(i,t)}\varepsilon_{(i,t)}|\mathcal{F}_n^W\right] \\ &+2\sum_{(i,t)=3}^{m_n}\sum_{(j,s)=2}^{(i,t)-1}\sum_{(k,\upsilon)=1}^{(j,s)-1}\frac{A_{(i,t),(j,s)}A_{(i,t),(k,\upsilon)}}{K_{2,n}}E\left[\underline{u}_{2,(i,t)}\varepsilon_{(i,t)}|\mathcal{F}_n^W\right]\left\{\underline{u}_{2,(j,s)}\varepsilon_{(k,\upsilon)} + \varepsilon_{(j,s)}\underline{u}_{2,(k,\upsilon)}\right\} \end{aligned}$$

$$+2\sum_{(i,t)=3}^{m_n}\sum_{(j,s)=2}^{(i,t)-1}\sum_{(k,v)=1}^{(j,s)-1}\frac{A_{(i,t),(j,s)}A_{(i,t),(k,v)}}{K_{2,n}}\varepsilon_{(j,s)}\varepsilon_{(k,v)}E\left[\underline{u}_{2,(i,t)}^2|\mathcal{F}_n^W\right]$$
$$+2\sum_{(i,t)=3}^{m_n}\sum_{(j,s)=2}^{(i,t)-1}\sum_{(k,v)=1}^{(j,s)-1}\frac{A_{(i,t),(j,s)}A_{(i,t),(k,v)}}{K_{2,n}}\underline{u}_{2,(j,s)}\underline{u}_{2,(k,v)}E\left[\varepsilon_{(i,t)}^2|\mathcal{F}_n^W\right]$$
$$= \mathcal{T}_1+\mathcal{T}_2+\mathcal{T}_3+\mathcal{T}_4+\mathcal{T}_5+\mathcal{T}_6+\mathcal{T}_7,\ \ (say)$$

Note first that, by applying parts (a)-(c) of Lemma S2-14, we have $\mathcal{T}_1 \xrightarrow{p} 0$, $\mathcal{T}_2 \xrightarrow{p} 0$, and $\mathcal{T}_3 \xrightarrow{p} 0$. Consider next the term

$$\mathcal{T}_4 = 2\sum_{(i,t)=2}^{m_n}\sum_{(j,s)=1}^{(i,t)-1}\frac{A_{(i,t),(j,s)}^2}{K_{2,n}}\left(\underline{u}_{2,(j,s)}\varepsilon_{(j,s)} - E\left[\underline{u}_{2,(j,s)}\varepsilon_{(j,s)}|\mathcal{F}_n^W\right]\right)E\left[\underline{u}_{2,(i,t)}\varepsilon_{(i,t)}|\mathcal{F}_n^W\right].$$

In this case, we apply part (a) of Lemma S2-8 with $u_{(j,s)} = \underline{u}_{2,(j,s)}$, $\overline{\psi}_{(j,s)} = E\left[\underline{u}_{2,(j,s)}\varepsilon_{(j,s)}|\mathcal{F}_n^W\right]$, and $\phi_{(i,t)} = E\left[\underline{u}_{2,(i,t)}\varepsilon_{(i,t)}|\mathcal{F}_n^W\right]$. Note that, in this case, $\left\{\left(\underline{u}_{2,(i,t)},\varepsilon_{(i,t)}\right)\right\}_{(i,t)=1}^{m_n}$ is independent conditional on $\mathcal{F}_n^W$, and $\sup_{1\leq(i,t)\leq m_n}E\left[\varepsilon_{(i,t)}^4|\mathcal{F}_n^W\right] \leq C$ *a.s.* by Assumptions 1 and 2(i), respectively. Moreover, note that Assumption 2, part (d) of Lemma S2-3, Lemma S2-6, and the fact that $K_{2,n}/\left(\mu_n^{\min}\right)^2 = O(1)$ in this case together imply that there exists a constant $C \geq 1$ such that $E\left[\underline{u}_{2,(i,t)}^4|\mathcal{F}_n^W\right] \leq \left[K_{2,n}^2/\left(\mu_n^{\min}\right)^4\right]E\left[\left\|\underline{U}_{(i,t)}\right\|_2^4|\mathcal{F}_n^W\right]\left(a'\Sigma_n^{-1}a\right)^2 \leq C < \infty$ *a.s.* for all $(i,t)\in\{1,2,...,m_n\}$ and for all n sufficiently large, so that $\max_{1\leq(i,t)\leq m_n}E\left[\underline{u}_{2,(i,t)}^4|\mathcal{F}_n^W\right] \leq C$ *a.s.n.* Finally, using the upper bound derived in expression (27) in the proof of part (a) of Lemma S2-14[3], we obtain $\max_{1\leq(i,t)\leq m_n}\left|\phi_{(i,t)}\right| \leq \max_{1\leq(i,t)\leq m_n}E\left[\left|\underline{u}_{2,(i,t)}\varepsilon_{(i,t)}\right||\mathcal{F}_n^W\right] \leq C$ *a.s.n.* and $\max_{1\leq(j,s)\leq m_n}\left|\overline{\psi}_{(j,s)}\right| \leq \max_{1\leq(i,t)\leq m_n}E\left[\left|\underline{u}_{2,(j,s)}\varepsilon_{(j,s)}\right||\mathcal{F}_n^W\right] \leq C$ *a.s.n.* It follows by part (a) of Lemma S2-8 that $\mathcal{T}_4 \xrightarrow{p} 0$.

Now, consider $\mathcal{T}_5$. Here, we apply part (b) of Lemma S2-8 with $u_{(j,s)} = \underline{u}_{2,(j,s)}$ and $\phi_{(i,t)} = E\left[\underline{u}_{2,(i,t)}\varepsilon_{(i,t)}|\mathcal{F}_n^W\right]$. Note again that $\left\{\left(\underline{u}_{2,(i,t)},\varepsilon_{(i,t)}\right)\right\}_{(i,t)=1}^{m_n}$ is independent conditional on $\mathcal{F}_n^W$, and $\max_{1\leq(i,t)\leq m_n}E\left[\varepsilon_{(i,t)}^4|\mathcal{F}_n^W\right] \leq C$ *a.s.* by Assumptions 1 and 2(i), respectively. Moreover, previously, we have shown that $E\left[\underline{u}_{2,(i,t)}^4|\mathcal{F}_n^W\right] \leq C$ *a.s.n.* and $\max_{1\leq(i,t)\leq m_n}\left|\phi_{(i,t)}\right| \leq C$ *a.s.n.* Hence, applying part (b) of Lemma S2-8, we deduce that $\mathcal{T}_5 \xrightarrow{p} 0$.

Turning our attention to $\mathcal{T}_6$, we note that, for this term, we can apply part (c) of Lemma S2-8 with $\phi_{(i,t)} = E\left[\underline{u}_{2,(i,t)}^2|\mathcal{F}_n^W\right]$. From (1), there exists a positive constant C such that $E\left[\underline{u}_{2,(i,t)}^2|\mathcal{F}_n^W\right] \leq C < \infty$ *a.s.* for all $(i,t)\in\{1,2,...,m_n\}$ and for all n sufficiently large, so that $\max_{1\leq(i,t)\leq m_n}\left|\phi_{(i,t)}\right| = \max_{1\leq(i,t)\leq m_n}E\left[\underline{u}_{2,(i,t)}^2|\mathcal{F}_n^W\right] \leq C$ *a.s.n.* Hence, applying part (c) of Lemma

[3] A proof of Lemma S2-14 is given in section 1 of the Additional Online Appendix which can be viewed at the URL: http://econweb.umd.edu/~chao/Research/research_files/Additional_Online_Appendix_Jackknife_Estimation_Cluster_Sample_IV_Model.pdf

S2-8, we obtain $\mathcal{T}_6 \xrightarrow{p} 0$.

Finally, consider $\mathcal{T}_7$. In this case, we apply part (d) of Lemma S2-8 with $u_{(j,s)} = \underline{u}_{2,(j,s)}$, $u_{(k,v)} = \underline{u}_{2,(k,v)}$, and $\phi_{(i,t)} = E\left[\varepsilon_{(i,t)}^2 | \mathcal{F}_n^W\right]$. Using a conditional version of Liapounov's inequality and Assumption 2(i), we obtain $E\left[\varepsilon_{(i,t)}^2 | \mathcal{F}_n^W\right] \leq \left(E\left[\varepsilon_{(i,t)}^4 | \mathcal{F}_n^W\right]\right)^{1/2} \leq C < \infty$ *a.s.* for all $(i,t) \in \{1, 2, ..., m_n\}$ and for all n, so that $\max_{1 \leq (i,t) \leq m_n} \left|\phi_{(i,t)}\right|$
$= \max_{1 \leq (i,t) \leq m_n} E\left[\varepsilon_{(i,t)}^2 | \mathcal{F}_n^W\right] \leq C$ *a.s.* Moreover, as noted previously, Assumption 2, part (d) of Lemma S2-3, Lemma S2-6, and the fact that $K_{2,n} / \left(\mu_n^{\min}\right)^2 = O(1)$ together imply that
$\max_{1 \leq (i,t) \leq m_n} E\left[\underline{u}_{2,(i,t)}^4 | \mathcal{F}_n^W\right] \leq C$ *a.s.n.* It follows by applying part (d) of Lemma S2-8 that $\mathcal{T}_7 \xrightarrow{p} 0$.

The above argument shows that $\sum_{(i,t)=2}^{m_n} E\left[\mathcal{V}_{(i,t),n}^2 | \mathcal{F}_{(i,t)-1,n}\right] - s_{W_n}^2 = \sum_{k=1}^{7} \mathcal{T}_k = o_p(1)$. On the other hand, expression (2) above implies that $s_{W_n}^2 - 1 = o_p(1)$. Putting these two results together, we obtain $\sum_{(i,t)=2}^{m_n} E\left[\mathcal{V}_{(i,t),n}^2 | \mathcal{F}_{(i,t)-1,n}\right] - 1 = o_p(1)$, which establishes condition (21) of Lemma S2-15.

It now follows from Lemma S2-15 that
$\mathcal{V}_n = \sum_{(i,t)=2}^{m_n} \left\{ b_{1n}' \Gamma' M^Q e_{(i,t)} \varepsilon_{(i,t)} / \sqrt{n} + \sum_{(j,s)=1}^{(i,t)-1} A_{(i,t),(j,s)} \left[\underline{u}_{2,(i,t)} \varepsilon_{(j,s)} + \underline{u}_{2,(j,s)} \varepsilon_{(i,t)}\right] \right\} \xrightarrow{d} N(0,1)$.
Since, previously, we have shown that $a' \Sigma_n^{-1/2} \mathcal{Y}_n = \mathcal{V}_n + o_p(1)$, this further implies that $a' \Sigma_n^{-1/2} \mathcal{Y}_n \xrightarrow{d} N(0,1)$. Given that this result holds for all $a \in \mathbb{R}^d$ such that $\|a\|_2 = 1$, we can then apply the Cramér-Wold device to obtain

$$\Sigma_n^{-1/2} \mathcal{Y}_n = \Sigma_n^{-1/2} \left(\frac{\Gamma' M^{(Z_1,Q)} \varepsilon}{\sqrt{n}} + D_\mu^{-1} \underline{U}' A \varepsilon \right) \xrightarrow{d} N(0, I_d) \tag{4}$$

Next, let $H_n = \Gamma' M^{(Z_1,Q)} \Gamma / n$, $\Lambda_{I,n} = H_n^{-1} \Sigma_n H_n^{-1}$, and $\mathcal{Y}_n = \Gamma' M^{(Z_1,Q)} \varepsilon / \sqrt{n} + D_\mu^{-1} \underline{U}' A \varepsilon$, as given above. Consider first $\widehat{\delta}_{L,n}$. Theorem 1 has already shown that $\widehat{\delta}_{L,n} \xrightarrow{p} \delta_0$. To show asymptotic normality of $\widehat{\delta}_L$, note first that, by Lemma S2-11, $\widehat{\delta}_{L,n}$ satisfies the set of (normalized) first-order conditions $\widehat{\Delta}\left(\widehat{\delta}_{L,n}\right) = 0$, where
$\widehat{\Delta}(\delta) = -\left[(y - X\delta)' M^{(Z_1,Q)} (y - X\delta) / 2\right] \left[\partial \widehat{Q}_{FELIM}(\delta) / \partial \delta\right]$ with
$\widehat{\ell}(\delta) = \left[(y - X\delta)' A (y - X\delta)\right] / \left[(y - X\delta)' M^{(Z_1,Q)} (y - X\delta)\right]$. Applying the mean-value theorem to each component of $\widehat{\Delta}(\delta)$ and expanding it around the point $\delta = \delta_0$, we obtain $0 = \widehat{\Delta}\left(\widehat{\delta}_{L,n}\right) = \widehat{\Delta}(\delta_0) + \left(\partial \widehat{\Delta}\left(\overline{\delta}_n\right) / \partial \delta'\right) \left(\widehat{\delta}_{L,n} - \delta_0\right)$, with $\overline{\delta}_n$ lying on the line segment between $\widehat{\delta}_{L,n}$ and δ_0. Multiplying both sides of this equation by D_μ^{-1}, we further obtain

$$0 = D_\mu^{-1} \widehat{\Delta}(\delta_0) + D_\mu^{-1} \frac{\partial \widehat{\Delta}\left(\overline{\delta}_n\right)}{\partial \delta'} \left(\widehat{\delta}_{L,n} - \delta_0\right) = D_\mu^{-1} \widehat{\Delta}(\delta_0) + D_\mu^{-1} \frac{\partial \widehat{\Delta}\left(\overline{\delta}_n\right)}{\partial \delta'} D_\mu^{-1} D_\mu \left(\widehat{\delta}_{L,n} - \delta_0\right) \tag{5}$$

From the result of Lemma S2-10, we have $-D_\mu^{-1} \left(\partial \widehat{\Delta}\left(\overline{\delta}_n\right) / \partial \delta'\right) D_\mu^{-1} = H_n + o_p(1)$, where $H_n = \Gamma' M^{(Z_1,Q)} \Gamma / n$ is a positive definite matrix *a.s.n.* by Assumption 3(iii), which, in turn, implies that $D_\mu^{-1}\left(\partial\widehat{\Delta}\left(\overline{\delta}_n\right)/\partial\delta'\right)D_\mu^{-1}$ is nonsingular and, thus, invertible w.p.a.1. It follows that, for all n sufficiently large, we can solve for $D_\mu\left(\widehat{\delta}_{L,n}-\delta_0\right)$ in (5) above to get

$$\begin{aligned} D_\mu\left(\widehat{\delta}_{L,n}-\delta_0\right) &= -\left[D_\mu^{-1}\left(\frac{\partial\widehat{\Delta}\left(\overline{\delta}_n\right)}{\partial\delta'}\right)D_\mu^{-1}\right]^{-1}D_\mu^{-1}\widehat{\Delta}\left(\delta_0\right) \\ &= H_n^{-1}\left(\frac{\Gamma' M^{(Z_1,Q)}\varepsilon}{\sqrt{n}}+D_\mu^{-1}\underline{U}'A\varepsilon\right)\left[1+o_p\left(1\right)\right], \end{aligned} \tag{6}$$

where the last equality follows by applying Lemma S2-9. By part (d) of Lemma S2-3, Σ_n is positive definite *a.s.n.*, so that Σ_n^{-1} is well-defined for all n sufficiently large, and both $\Sigma_n^{1/2}$ and $\Sigma_n^{-1/2}$ can be taken to be symmetric matrices. Since H_n is also symmetric, it further follows that $\Lambda_{I,n}=H_n^{-1}\Sigma_n H_n^{-1}$ is symmetric and positive definite *a.s.n.*, and both $\Lambda_{I,n}^{-1}=\left(H_n^{-1}\Sigma_n H_n^{-1}\right)^{-1}$ and $\Lambda_{I,n}^{-1/2}=\left(H_n^{-1}\Sigma_n H_n^{-1}\right)^{-1/2}$ are well-defined for all n sufficiently large. Multiplying both sides of the equation above by $\Lambda_{I,n}^{-1/2}$, we then get $\Lambda_{I,n}^{-1/2}D_\mu\left(\widehat{\delta}_{L,n}-\delta_0\right)=\left(H_n^{-1}\Sigma_n H_n^{-1}\right)^{-1/2}H_n^{-1}\mathcal{Y}_n\left[1+o_p\left(1\right)\right]$, where $\mathcal{Y}_n=\Gamma' M^{(Z_1,Q)}\varepsilon/\sqrt{n}+D_\mu^{-1}\underline{U}'A\varepsilon$. Let $R_{W,n}=\left(H_n^{-1}\Sigma_n H_n^{-1}\right)^{-1/2}H_n^{-1}\Sigma_n^{1/2}$, and note that $R_{W,n}R_{W,n}'=I_d$ for all n sufficiently large. It, thus, follows from the result given in (4) above and the continuous mapping theorem that $\Lambda_{I,n}^{-1/2}D_\mu\left(\widehat{\delta}_{L,n}-\delta_0\right)\xrightarrow{d}N\left(0,I_d\right)$, as $n\to\infty$, as required.

Turning our attention now to $\widehat{\delta}_{F,n}$, note that we can write this estimator, appropriately standardized, as

$$\begin{aligned} & D_\mu\left(\widehat{\delta}_{F,n}-\delta_0\right) \\ = & \left(D_\mu^{-1}X'\left[A-\widehat{\ell}_{F,n}M^{(Z_1,Q)}\right]XD_\mu^{-1}\right)^{-1}D_\mu^{-1}X'\left[A-\widehat{\ell}_{F,n}M^{(Z_1,Q)}\right]\left(y-X\delta_0\right) \end{aligned} \tag{7}$$

so that, multiplying by $\Lambda_{I,n}^{-1/2}=\left(H_n^{-1}\Sigma_n H_n^{-1}\right)^{-1/2}$ and applying Lemmas S2-12 and S2-13, we obtain $\Lambda_{I,n}^{-1/2}D_\mu\left(\widehat{\delta}_{F,n}-\delta_0\right)=\left(H_n^{-1}\Sigma_n H_n^{-1}\right)^{-1/2}H_n^{-1}\mathcal{Y}_n\left[1+o_p\left(1\right)\right]$. It follows from the result given in (4) above and the continuous mapping theorem that $\Lambda_{I,n}^{-1/2}D_\mu\left(\widehat{\delta}_{F,n}-\delta_0\right)\xrightarrow{d}N\left(0,I_d\right)$, as $n\to\infty$, as required. □

Proof of Theorem 3: To proceed, note that, in this case, $\left(\mu_n^{\min}\right)/\sqrt{K_{2,n}}=o\left(1\right)$ but $\sqrt{K_{2,n}}/\left(\mu_n^{\min}\right)^2\to 0$, so that, by the result given in Lemma S2-9, we have

$$\frac{\mu_n^{\min}}{\sqrt{K_{2,n}}}D_\mu^{-1}\widehat{\Delta}\left(\delta_0\right)=\frac{\mu_n^{\min}}{\sqrt{K_{2,n}}}D_\mu^{-1}\underline{U}'A\varepsilon+o_p\left(1\right) \tag{8}$$

where $\underline{U}=U-\varepsilon\rho'$. Again, let $H_n=\Gamma' M^{(Z_1,Q)}\Gamma/n$, and $\Sigma_{2,n}=VC\left(D_\mu^{-1}\underline{U}'A\varepsilon|\mathcal{F}_n^W\right)$ $=D_\mu^{-1}VC\left(\underline{U}'A\varepsilon|\mathcal{F}_n^W\right)D_\mu^{-1}$. Now, by assumption, $\widetilde{L}_n$ can be any sequence of bounded $(l\times d)$ non-random matrices such that $\lambda_{\min}\left(\left(\mu_n^{\min}\right)^2\widetilde{L}_nH_n^{-1}\Sigma_{2,n}H_n^{-1}\widetilde{L}_n'/K_{2,n}\right)\geq\underline{C}$ *a.s.n.* for some constant $\underline{C}>0$. It follows that $\left(\mu_n^{\min}\right)^2\widetilde{L}_nH_n^{-1}\Sigma_{2,n}H_n^{-1}\widetilde{L}_n'/K_{2,n}$ is positive definite *a.s.n.*, so that, with

probability one, $\left(\left(\mu_n^{\min}\right)^2 \widetilde{L}_n H_n^{-1} \Sigma_{2,n} H_n^{-1} \widetilde{L}_n' / K_{2,n}\right)^{-1/2}$ is well-defined for all n sufficiently large. Hence, we can let $\widetilde{\mathcal{N}}_n = \left(\left(\mu_n^{\min}\right)^2 \widetilde{L}_n H_n^{-1} \Sigma_{2,n} H_n^{-1} \widetilde{L}_n' / K_{2,n}\right)^{-1/2} \widetilde{L}_n H_n^{-1} \left(\mu_n^{\min}/\sqrt{K_{2,n}}\right) D_\mu^{-1} \underline{U}' A \varepsilon$ and construct the linear combination $\mathcal{J}_n = a'\widetilde{\mathcal{N}}_n$ for any $a \in \mathbb{R}^d$ such that $\|a\|_2 = 1$. Next, define $\underline{u}_{(i,t),n} = a' \left(\left(\mu_n^{\min}\right)^2 \widetilde{L}_n H_n^{-1} \Sigma_{2,n} H_n^{-1} \widetilde{L}_n' / K_{2,n}\right)^{-1/2} \widetilde{L}_n H_n^{-1} D_\mu^{-1} \underline{U}_{(i,t)}$, with $\underline{u}_{(j,s),n}$ similarly defined, and we can write $\mathcal{J}_n = \left(\mu_n^{\min}/\sqrt{K_{2,n}}\right) \sum_{(i,t)=2}^{m_n} \sum_{(j,s)=1}^{(i,t)-1} A_{(i,t),(j,s)} \left[\underline{u}_{(i,t),n} \varepsilon_{(j,s)} + \underline{u}_{(j,s),n} \varepsilon_{(i,t)}\right] = \sum_{(i,t)=2}^{m_n} \mathcal{J}_{(i,t),n}$, where $\mathcal{J}_{(i,t),n} = \left(\mu_n^{\min}/\sqrt{K_{2,n}}\right) \sum_{(j,s)=1}^{(i,t)-1} A_{(i,t),(j,s)} \left[\underline{u}_{(i,t),n} \varepsilon_{(j,s)} + \underline{u}_{(j,s),n} \varepsilon_{(i,t)}\right]$. Again, define the σ-fields $\mathcal{F}_{(i,t),n} = \sigma\left(\left\{\varepsilon_{(k,v)}, U_{(k,v)}\right\}_{(k,v)=1}^{(i,t)}, W_n\right)$ for $(i,t) = 1, 2, ..., m_n$, noting that by construction $\mathcal{F}_{(i,t)-1,n} \subseteq \mathcal{F}_{(i,t),n}$ for $(i,t) = 2, ..., m_n$ and $\mathcal{J}_{(i,t),n}$ is $\mathcal{F}_{(i,t),n}-$measurable. In addition, note that, using Assumption 1, it is easily seen that $E\left[\underline{u}_{(i,t),n} | \mathcal{F}_{(i,t)-1,n}\right] = 0$ and $E\left[\varepsilon_{(i,t)} | \mathcal{F}_{(i,t)-1,n}\right] = 0$, from which it follows that $E\left[\mathcal{J}_{(i,t),n} | \mathcal{F}_{(i,t)-1,n}\right] =$
$\left(\mu_n^{\min}/\sqrt{K_{2,n}}\right) \sum_{(j,s)=1}^{(i,t)-1} A_{(i,t),(j,s)} \left\{\varepsilon_{(j,s)} E\left[\underline{u}_{(i,t),n} | \mathcal{F}_{(i,t)-1,n}\right] + \underline{u}_{(j,s),n} E\left[\varepsilon_{(i,t)} | \mathcal{F}_{(i,t)-1,n}\right]\right\} = 0$. Moreover, applying the CS inequality and making use of the fact that

$$E\left[\underline{u}_{(i,t),n}^2 | \mathcal{F}_n^W\right] \leq \frac{\max_{1 \leq (i,t) \leq m_n} E\left[\left\|\underline{U}_{(i,t)}\right\|_2^2 | \mathcal{F}_n^W\right] \left\|\widetilde{L}_n\right\|_F^2}{\lambda_{\min}\left(\left(\mu_n^{\min}\right)^2 \widetilde{L}_n H_n^{-1} \Sigma_{2,n} H_n^{-1} \widetilde{L}_n' / K_{2,n}\right) \left[\lambda_{\min}\left(H_n\right)\right]^2} \left(\frac{1}{\mu_n^{\min}}\right)^2 = O_{a.s.}\left(\frac{1}{\left(\mu_n^{\min}\right)^2}\right) \tag{9}$$

and that $E\left[\varepsilon_{(i,t)}^2 | \mathcal{F}_n^W\right] \leq \overline{C}$ *a.s.* by Assumption 2(i), we see that

$$\begin{aligned}
& Var\left(\mathcal{J}_{(i,t),n} | \mathcal{F}_n^W\right) \\
\leq \;& \frac{\left(\mu_n^{\min}\right)^2}{K_{2,n}} \sum_{(j,s)=1}^{(i,t)-1} A_{(i,t),(j,s)}^2 \left(E\left[\underline{u}_{(i,t),n}^2 | \mathcal{F}_n^W\right] E\left[\varepsilon_{(j,s)}^2 | \mathcal{F}_n^W\right] + E\left[\varepsilon_{(i,t)}^2 | \mathcal{F}_n^W\right] E\left[\underline{u}_{(j,s),n}^2 | \mathcal{F}_n^W\right]\right. \\
& \left. +2\sqrt{E\left[\underline{u}_{(i,t),n}^2 | \mathcal{F}_n^W\right] E\left[\varepsilon_{(i,t)}^2 | \mathcal{F}_n^W\right]} \sqrt{E\left[\underline{u}_{(j,s),n}^2 | \mathcal{F}_n^W\right] E\left[\varepsilon_{(j,s)}^2 | \mathcal{F}_n^W\right]}\right) \\
\leq \;& \frac{4\overline{C}^2}{\left(\mu_n^{\min}\right)^2} \frac{\left(\mu_n^{\min}\right)^2}{K_{2,n}} \sum_{(j,s)=1}^{(i,t)-1} A_{(i,t),(j,s)}^2 = \frac{4\overline{C}^2}{K_{2,n}} \sum_{(j,s)=1}^{(i,t)-1} A_{(i,t),(j,s)}^2 \; a.s.n.
\end{aligned} \tag{10}$$

Hence, applying the law of iterated expectations, part (d) of Lemma S2-1, and Theorem 16.1 of Billingsley (1995), we further deduce that $Var\left(\mathcal{J}_{(i,t),n}\right) = E_W\left[E\left(\mathcal{J}_{(i,t),n}^2 | \mathcal{F}_n^W\right)\right]$
$\leq \left(4\overline{C}^2/K_{2,n}\right) \sum_{(j,s)=1}^{(i,t)-1} E_W\left[A_{(i,t),(j,s)}^2\right] \leq C$ for some positive constant C for all n sufficiently large. These results show that $\left\{\mathcal{J}_{(i,t),n}, \mathcal{F}_{(i,t),n}, 1 \leq (i,t) \leq m_n, n \geq 1\right\}$ forms a square-integrable martingale difference array.

Next, we verify condition (22) of the central limit theorem for martingale difference arrays given in Lemma S2-15 below. By Loève's c_r inequality we have

$$\begin{aligned}
&\sum_{(i,t)=2}^{m_n} E\left[\left(\frac{\mu_n^{\min}}{\sqrt{K_{2,n}}}\sum_{(j,s)=1}^{(i,t)-1} A_{(i,t),(j,s)}\left[\underline{u}_{(i,t),n}\varepsilon_{(j,s)}+\underline{u}_{(j,s),n}\varepsilon_{(i,t)}\right]\right)^4 |\mathcal{F}_n^W\right] \\
\leq\ & 8\sum_{(i,t)=2}^{m_n}\frac{\left(\mu_n^{\min}\right)^4}{K_{2,n}^2}E\left[\left(\sum_{(j,s)=1}^{(i,t)-1} A_{(i,t),(j,s)}\underline{u}_{(i,t),n}\varepsilon_{(j,s)}\right)^4 |\mathcal{F}_n^W\right] \\
&+8\sum_{(i,t)=2}^{m_n}\frac{\left(\mu_n^{\min}\right)^4}{K_{2,n}^2}E\left[\left(\sum_{(j,s)=1}^{(i,t)-1} A_{(i,t),(j,s)}\underline{u}_{(j,s),n}\varepsilon_{(i,t)}\right)^4 |\mathcal{F}_n^W\right] \\
=\ & \mathcal{E}_1+\mathcal{E}_2,\ \ (say)\,.
\end{aligned}\tag{11}$$

Focusing first on $\mathcal{E}_1$, it is easy to see that there exists some positive constant C such that

$$\begin{aligned}
&\mathcal{E}_1 \\
=\ & \frac{8\left(\mu_n^{\min}\right)^4}{K_{2,n}^2}E\left[\sum_{(i,t)=2}^{m_n}\left(\sum_{(j,s)=1}^{(i,t)-1} A_{(i,t),(j,s)}\underline{u}_{(i,t),n}\varepsilon_{(j,s)}\right)^4 |\mathcal{F}_n^W\right] \\
\leq\ & \frac{8\left(\mu_n^{\min}\right)^4}{K_{2,n}^2}\sum_{(i,t)=1}^{m_n}\sum_{\substack{(j,s)=1\\(j,s)\neq(i,t)}}^{m_n} A_{(i,t),(j,s)}^4 E\left[\underline{u}_{(i,t),n}^4|\mathcal{F}_n^W\right]E\left[\varepsilon_{(j,s)}^4|\mathcal{F}_n^W\right] \\
&+\frac{24\left(\mu_n^{\min}\right)^4}{K_{2,n}^2}\sum_{(i,t)=1}^{m_n}\sum_{\substack{(j,s),(k,v)=1\\(j,s)\neq(i,t),(k,v)\neq(i,t)\\(j,s)\neq(k,v)}}^{m_n} A_{(i,t),(j,s)}^2A_{(i,t),(k,v)}^2E\left[\underline{u}_{(i,t),n}^4|\mathcal{F}_n^W\right]E\left[\varepsilon_{(j,s)}^2|\mathcal{F}_n^W\right]E\left[\varepsilon_{(k,v)}^2|\mathcal{F}_n^W\right] \\
\leq\ & \frac{C}{K_{2,n}}\left[\frac{1}{K_{2,n}}\sum_{\substack{(i,t),(j,s)=1\\(j,s)\neq(i,t)}}^{m_n} A_{(i,t),(j,s)}^4+\frac{1}{K_{2,n}}\sum_{(i,t)=1}^{m_n}\sum_{\substack{(j,s),(k,v)=1\\(j,s)\neq(i,t),(k,v)\neq(i,t)}}^{m_n} A_{(i,t),(j,s)}^2A_{(i,t),(k,v)}^2\right]
\end{aligned}$$

where the second inequality above follows from Assumption 2(i) and from an upper bound on the conditional fourth moment of
$\underline{u}_{(i,t),n}=a'\left(\left(\mu_n^{\min}\right)^2\widetilde{L}_nH_n^{-1}\Sigma_{2,n}H_n^{-1}\widetilde{L}_n'/K_{2,n}\right)^{-1/2}\widetilde{L}_nH_n^{-1}D_\mu^{-1}\underline{U}_{(i,t)}$ given by

$$\begin{aligned}
E\left[\underline{u}_{(i,t),n}^4|\mathcal{F}_n^W\right] \ \leq\ & \frac{1}{\left(\mu_n^{\min}\right)^4}\left(\max_{1\leq(i,t)\leq m_n}E\left[\left\|\underline{U}_{(i,t)}\right\|_2^4|\mathcal{F}_n^W\right]\right)\frac{1}{\left[\lambda_{\min}\left(H_n\right)\right]^4} \\
&\times\left\|\widetilde{L}_n\right\|_F^4\left(\frac{1}{\lambda_{\min}\left(\left(\mu_n^{\min}\right)^2\widetilde{L}_nH_n^{-1}\Sigma_{2,n}H_n^{-1}\widetilde{L}_n'/K_{2,n}\right)}\right)^2 \\
\leq\ & \frac{C^*}{\left(\mu_n^{\min}\right)^4}\ \ a.s.n.,\text{ for some constant }C^*>0.
\end{aligned}\tag{12}$$

Note also that, in deriving the upper bound given in (12), we have applied Assumption 3(iii), Lemma S2-6, the boundedness of $\left\|\widetilde{L}_n\right\|_F^2$, and the assumption that $\lambda_{\min}\left(\left(\mu_n^{\min}\right)^2 \widetilde{L}_n H_n^{-1}\Sigma_{2,n}H_n^{-1}\widetilde{L}_n'/K_{2,n}\right) \geq \underline{C} > 0 \;\; a.s.n.$ Moreover, by parts (b) and (c) of Lemma S2-1, we have that $K_{2,n}^{-1}\sum_{(i,t),(j,s)=1,(i,t)\neq(j,s)}^{m_n} A_{(i,t),(j,s)}^4 = O_{a.s.}\left(K_{2,n}^2/n^2\right)$ and $K_{2,n}^{-1}\sum_{(i,t)=1}^{m_n}\sum_{(j,s),(k,v)=1,(j,s)\neq(i,t),(k,v)\neq(i,t)}^{m_n} A_{(i,t),(j,s)}^2 A_{(i,t),(k,v)}^2 = O_{a.s.}\left(K_{2,n}/n\right)$. from which it follows that $n\mathcal{E}_1 = O_{a.s.}(1)$ in light of Assumption 5(ii). Hence, by applying the law of iterated expectations and Theorem 16.1 of Billingsley (1995), we obtain

$$\begin{aligned}
& nE_{W_n}\left[\mathcal{E}_1\right] \\
= \; & \frac{8n\left(\mu_n^{\min}\right)^4}{K_{2,n}^2} E_{W_n}\left\{E\left[\sum_{(i,t)=2}^{m_n}\left(\sum_{(j,s)=1}^{(i,t)-1} A_{(i,t),(j,s)}\underline{u}_{(i,t),n}\varepsilon_{(j,s)}\right)^4 |\mathcal{F}_n^W\right]\right\} \\
\leq \; & \frac{Cn}{K_{2,n}}\left\{E_{W_n}\left[\frac{1}{K_{2,n}}\sum_{\substack{(i,t),(j,s)=1\\(i,t)\neq(j,s)}}^{m_n} A_{(i,t),(j,s)}^4 + \frac{1}{K_{2,n}}\sum_{(i,t)=1}^{m_n}\sum_{\substack{(j,s),(k,v)=1\\(j,s)\neq(i,t),(k,v)\neq(i,t)}}^{m_n} A_{(i,t),(j,s)}^2 A_{(i,t),(k,v)}^2\right]\right\} \\
& \text{(for all } n \text{ sufficiently large)} \\
= \; & O(1),
\end{aligned}$$

which shows that $E_{W_n}\left[\mathcal{E}_1\right] = O(1/n) = o(1)$. In a similar way, we can also show that $E_{W_n}\left[\mathcal{E}_2\right] = 8\left[\left(\mu_n^{\min}\right)^4/K_{2,n}^2\right] E\left[\sum_{(i,t)=2}^{m_n}\left(\sum_{(j,s)=1}^{(i,t)-1} A_{(j,s),(i,t)}\underline{u}_{(i,t),n}\varepsilon_{(i,t)}\right)^4\right] = o(1)$. Condition (22) of Lemma S2-15 then follows from these calculations since

$$\sum_{(i,t)=2}^{m_n} E\left[\left(\frac{\mu_n^{\min}}{\sqrt{K_{2,n}}}\sum_{(j,s)=1}^{(i,t)-1} A_{(i,t),(j,s)}\left[\underline{u}_{(i,t),n}\varepsilon_{(j,s)} + \underline{u}_{(j,s),n}\varepsilon_{(i,t)}\right]\right)^4\right] \leq E_{W_n}\left[\mathcal{E}_1\right] + E_{W_n}\left[\mathcal{E}_2\right] = o(1)$$

Next, we verify condition (21) of Lemma S2-15. Note first that, by construction, $Var\left(\mathcal{J}_n|\mathcal{F}_n^W\right)$ $= a'\left(\widetilde{L}_n\Lambda_{II,n}\widetilde{L}_n'\right)^{-1/2}\widetilde{L}_n\Lambda_{II,n}\widetilde{L}_n'\left(\widetilde{L}_n\Lambda_{II,n}\widetilde{L}_n'\right)^{-1/2} a = 1$, with $\Lambda_{II,n} = \left[\left(\mu_n^{\min}\right)^2/K_{2,n}\right] H_n^{-1}\Sigma_{2,n}H_n^{-1}$. This, in turn, implies that $Var\left(\mathcal{J}_n\right) = E_{W_n}\left[E\left(\mathcal{J}_n^2|\mathcal{F}_n^W\right)\right] = E_{W_n}\left[Var\left(\mathcal{J}_n|\mathcal{F}_n^W\right)\right] = 1$. On the other hand, by direct calculation, we obtain

$$\begin{aligned}
1 \; = \; & Var\left(\mathcal{J}_n|\mathcal{F}_n^W\right) \\
= \; & \frac{\left(\mu_n^{\min}\right)^2}{K_{2,n}}\sum_{(i,t)=2}^{m_n}\sum_{(j,s)=1}^{(i,t)-1} A_{(i,t),(j,s)}^2 E\left[\varepsilon_{(j,s)}^2|\mathcal{F}_n^W\right] E\left[\underline{u}_{(i,t),n}^2|\mathcal{F}_n^W\right] \\
& + \frac{\left(\mu_n^{\min}\right)^2}{K_{2,n}}\sum_{(i,t)=2}^{m_n}\sum_{(j,s)=1}^{(i,t)-1} A_{(i,t),(j,s)}^2 E\left[\underline{u}_{(j,s),n}^2|\mathcal{F}_n^W\right] E\left[\varepsilon_{(i,t)}^2|\mathcal{F}_n^W\right]
\end{aligned}$$

$$+2\frac{\left(\mu_n^{\min}\right)^2}{K_{2,n}}\sum_{(i,t)=2}^{m_n}\sum_{(j,s)=1}^{(i,t)-1}A^2_{(i,t),(j,s)}E\left[\underline{u}_{(j,s),n}\varepsilon_{(j,s)}|\mathcal{F}_n^W\right]E\left[\underline{u}_{(i,t),n}\varepsilon_{(i,t)}|\mathcal{F}_n^W\right] \quad (13)$$

Making use of expression (13), we obtain, after some further calculations,

$$\begin{aligned}
&\sum_{(i,t)=2}^{m_n}E\left[\mathcal{J}^2_{(i,t),n}|\mathcal{F}_{(i,t)-1,n}\right]-1\\
=\;&\frac{\left(\mu_n^{\min}\right)^2}{K_{2,n}}\sum_{(i,t)=2}^{m_n}\sum_{(j,s)=1}^{(i,t)-1}A^2_{(i,t),(j,s)}\left(\varepsilon^2_{(j,s)}-E\left[\varepsilon^2_{(j,s)}|\mathcal{F}_n^W\right]\right)E\left[\underline{u}^2_{(i,t),n}|\mathcal{F}_n^W\right]\\
&+\frac{\left(\mu_n^{\min}\right)^2}{K_{2,n}}\sum_{(i,t)=2}^{m_n}\sum_{(j,s)=1}^{(i,t)-1}A^2_{(i,t),(j,s)}\left(\underline{u}^2_{(j,s),n}-E\left[\underline{u}^2_{(j,s),n}|\mathcal{F}_n^W\right]\right)E\left[\varepsilon^2_{(i,t)}|\mathcal{F}_n^W\right]\\
&+\frac{2\left(\mu_n^{\min}\right)^2}{K_{2,n}}\sum_{(i,t)=2}^{m_n}\sum_{(j,s)=1}^{(i,t)-1}A^2_{(i,t),(j,s)}\left(\underline{u}_{(j,s),n}\varepsilon_{(j,s)}-E\left[\underline{u}_{(j,s),n}\varepsilon_{(j,s)}|\mathcal{F}_n^W\right]\right)E\left[\underline{u}_{(i,t),n}\varepsilon_{(i,t)}|\mathcal{F}_n^W\right]\\
&+\frac{2\left(\mu_n^{\min}\right)^2}{K_{2,n}}\sum_{(i,t)=3}^{m_n}\sum_{(j,s)=2}^{(i,t)-1}\sum_{(k,v)=1}^{(j,s)-1}A_{(i,t),(j,s)}A_{(i,t),(k,v)}E\left[\underline{u}_{(i,t),n}\varepsilon_{(i,t)}|\mathcal{F}_n^W\right]\left\{\underline{u}_{(j,s),n}\varepsilon_{(k,v)}+\varepsilon_{(j,s)}\underline{u}_{(k,v),n}\right\}\\
&+\frac{2\left(\mu_n^{\min}\right)^2}{K_{2,n}}\sum_{(i,t)=3}^{m_n}\sum_{(j,s)=2}^{(i,t)-1}\sum_{(k,v)=1}^{(j,s)-1}A_{(i,t),(j,s)}A_{(i,t),(k,v)}\varepsilon_{(j,s)}\varepsilon_{(k,v)}E\left[\underline{u}^2_{(i,t),n}|\mathcal{F}_n^W\right]\\
&+\frac{2\left(\mu_n^{\min}\right)^2}{K_{2,n}}\sum_{(i,t)=3}^{m_n}\sum_{(j,s)=2}^{(i,t)-1}\sum_{(k,v)=1}^{(j,s)-1}A_{(i,t),(j,s)}A_{(i,t),(k,v)}\underline{u}_{(j,s),n}\underline{u}_{(k,v),n}E\left[\varepsilon^2_{(i,t)}|\mathcal{F}_n^W\right]\\
=\;&\mathcal{TT}_1+\mathcal{TT}_2+\mathcal{TT}_3+\mathcal{TT}_4+\mathcal{TT}_5+\mathcal{TT}_6 \qquad (14)
\end{aligned}$$

To analyze the terms $\mathcal{TT}_k$ $(k=1,..,6)$, note first that, by applying parts (b) and (a) of Lemma S2-16, we obtain $\mathcal{TT}_1\xrightarrow{p}0$ and $\mathcal{TT}_2\xrightarrow{p}0$, respectively. Consider now the term

$$\mathcal{TT}_3=\frac{2\left(\mu_n^{\min}\right)^2}{K_{2,n}}\sum_{(i,t)=2}^{m_n}\sum_{(j,s)=1}^{(i,t)-1}A^2_{(i,t),(j,s)}\left(\underline{u}_{(j,s),n}\varepsilon_{(j,s)}-E\left[\underline{u}_{(j,s),n}\varepsilon_{(j,s)}|\mathcal{F}_n^W\right]\right)E\left[\underline{u}_{(i,t),n}\varepsilon_{(i,t)}|\mathcal{F}_n^W\right]$$

In this case, we apply part (a) of Lemma S2-8 with $u_{(j,s),n}=\left(\mu_n^{\min}\right)\underline{u}_{(j,s),n}$, $\overline{\psi}_{(j,s)}=E\left[\left(\mu_n^{\min}\right)\underline{u}_{(j,s),n}\varepsilon_{(j,s)}|\mathcal{F}_n^W\right]$, and $\phi_{(i,t)}=E\left[\left(\mu_n^{\min}\right)\underline{u}_{(i,t),n}\varepsilon_{(i,t)}|\mathcal{F}_n^W\right]$. Note that, in this case, $\left\{\left(u_{(i,t),n},\varepsilon_{(i,t)}\right)\right\}_{(i,t)=1}^{m_n}$ is independent conditional on $\mathcal{F}_n^W=\sigma\left(W_n\right)$, and $\max_{1\leq(i,t)\leq m_n}E\left[\varepsilon^4_{(i,t)}|\mathcal{F}_n^W\right]\leq C\ a.s.$ by Assumptions 1(i) and 2(i), respectively. Moreover, the upper bound given by (12) implies that there exists a constant $C^*>0$ such that $\max_{1\leq(i,t)\leq m_n}E\left[u^4_{(i,t),n}|\mathcal{F}_n^W\right]=\max_{1\leq(i,t)\leq m_n}\left(\mu_n^{\min}\right)^4E\left[\underline{u}^4_{(i,t),n}|\mathcal{F}_n^W\right]\leq\left(\mu_n^{\min}\right)^4C^*/\left(\mu_n^{\min}\right)^4=C^*\ \ a.s.n.$ Finally, note that, by using the fact that

$\underline{u}_{(i,t),n} = a'\left(\left(\mu_n^{\min}\right)^2 \widetilde{L}_n H_n^{-1}\Sigma_{2,n}H_n^{-1}\widetilde{L}_n'/K_{2,n}\right)^{-1/2}\widetilde{L}_n H_n^{-1}D_\mu^{-1}\underline{U}_{(i,t)}$ and by applying Assumption 2(i), Lemma S2-6, and the assumption that $\lambda_{\min}\left(\left(\mu_n^{\min}\right)^2 \widetilde{L}_n H_n^{-1}\Sigma_{2,n}H_n^{-1}\widetilde{L}_n'/K_{2,n}\right) \geq \underline{C} > 0$ *a.s.n.*; we can show that there exists a constant $C > 0$ such that

$$
\begin{aligned}
& E\left[\left|\left(\mu_n^{\min}\right)\underline{u}_{(i,t),n}\varepsilon_{(i,t)}\right| |\mathcal{F}_n^W\right] \\
= & \left(\mu_n^{\min}\right) E\left[\left|\varepsilon_{(i,t)}\underline{U}_{(i,t)}' D_\mu^{-1}H_n^{-1}\widetilde{L}_n'\left(\frac{\left(\mu_n^{\min}\right)^2 \widetilde{L}_n H_n^{-1}\Sigma_{2,n}H_n^{-1}\widetilde{L}_n'}{K_{2,n}}\right)^{-1/2} a\right| |\mathcal{F}_n^W\right] \\
\leq & \left(\mu_n^{\min}\right)\sqrt{E\left[\varepsilon_{(i,t)}^2|\mathcal{F}_n^W\right]}\left[a'\left(\frac{\left(\mu_n^{\min}\right)^2 \widetilde{L}_n H_n^{-1}\Sigma_{2,n}H_n^{-1}\widetilde{L}_n'}{K_{2,n}}\right)^{-1/2}\widetilde{L}_n H_n^{-1}D_\mu^{-1}\right. \\
& \left.\times E\left[\underline{U}_{(i,t)}\underline{U}_{(i,t)}'|\mathcal{F}_n^W\right] D_\mu^{-1}H_n^{-1}\widetilde{L}_n'\left(\frac{\left(\mu_n^{\min}\right)^2 \widetilde{L}_n H_n^{-1}\Sigma_{2,n}H_n^{-1}\widetilde{L}_n'}{K_{2,n}}\right)^{-1/2} a\right]^{1/2} \quad \text{(by CS inequality)} \\
\leq & \left(\mu_n^{\min}\right)\sqrt{E\left[\varepsilon_{(i,t)}^2|\mathcal{F}_n^W\right]}\frac{1}{\left(\mu_n^{\min}\right)}\left(\sqrt{\max_{1\leq(i,t)\leq m_n} E\left[\left\|\underline{U}_{(i,t)}\right\|_2^2 |\mathcal{F}_n^W\right]}\right) \\
& \times\frac{1}{\lambda_{\min}\left(\Gamma' M^{(Z_1,Q)}\Gamma/n\right)}\left\|\widetilde{L}_n\right\|_F\left(\frac{1}{\sqrt{\lambda_{\min}\left(\left(\mu_n^{\min}\right)^2 \widetilde{L}_n H_n^{-1}\Sigma_{2,n}H_n^{-1}\widetilde{L}_n'/K_{2,n}\right)}}\right) \\
\leq & C < \infty \;\; a.s. \text{ for all } (i,t)\in\{1,2,...,m_n\} \text{ and for all } n \text{ sufficiently large} \qquad (15)
\end{aligned}
$$

from which we further deduce that $\max_{(i,t)}\left|\phi_{(i,t)}\right| \leq \max_{(i,t)} E\left[\left|\left(\mu_n^{\min}\right)\underline{u}_{(i,t),n}\varepsilon_{(i,t)}\right| |\mathcal{F}_n^W\right] \leq C$ *a.s.n.*and also that $\max_{(j,s)}\left|\overline{\psi}_{(j,s)}\right| \leq \max_{(j,s)} E\left[\left|\left(\mu_n^{\min}\right)\underline{u}_{(j,s),n}\varepsilon_{(j,s)}\right| |\mathcal{F}_n^W\right] \leq C$ *a.s.n.* Hence, applying part (a) of Lemma S2-8, we have $\mathcal{TT}_3 \xrightarrow{p} 0$.

Next, consider the term

$$
\mathcal{TT}_4 = \frac{2\left(\mu_n^{\min}\right)^2}{K_{2,n}}\sum_{(i,t)=3}^{m_n}\sum_{(j,s)=2}^{(i,t)-1}\sum_{(k,v)=1}^{(j,s)-1} A_{(i,t),(j,s)}A_{(i,t),(k,v)}E\left[\underline{u}_{(i,t),n}\varepsilon_{(i,t)}|\mathcal{F}_n^W\right]\left\{\underline{u}_{(j,s),n}\varepsilon_{(k,v)} + \varepsilon_{(j,s)}\underline{u}_{(k,v),n}\right\}
$$

Here, we apply part (b) of Lemma S2-8 with $u_{(j,s),n} = \left(\mu_n^{\min}\right)\underline{u}_{(j,s),n}$ and $\phi_{(i,t)} = E\left[\left(\mu_n^{\min}\right)\underline{u}_{(i,t),n}\varepsilon_{(i,t)}|\mathcal{F}_n^W\right]$. Note that $\left\{\left(u_{(i,t),n},\varepsilon_{(i,t)}\right)\right\}_{(i,t)=1}^{m_n}$ is independent conditional on $\mathcal{F}_n^W = \sigma\left(W_n\right)$, and $\max_{1\leq(i,t)\leq m_n} E\left[\varepsilon_{(i,t)}^4|\mathcal{F}_n^W\right] \leq C$ *a.s.* by Assumptions 1 and 2(i), respectively. Moreover, from calculations given previously, we have $\max_{1\leq(i,t)\leq m_n}\left(\mu_n^{\min}\right)^4 E\left[\underline{u}_{(i,t),n}^4|\mathcal{F}_n^W\right] \leq C$ *a.s.n.* and $\max_{(i,t)}\left|\phi_{(i,t)}\right| \leq C$ *a.s.n.* Hence, by applying part (b) of Lemma S2-8, we deduce that $\mathcal{TT}_4 \xrightarrow{p} 0$.

Turning our attention to the term

$$\mathcal{TT}_5 = \frac{2\left(\mu_n^{\min}\right)^2}{K_{2,n}} \sum_{(i,t)=3}^{m_n} \sum_{(j,s)=2}^{(i,t)-1} \sum_{(k,\upsilon)=1}^{(j,s)-1} A_{(i,t),(j,s)} A_{(i,t),(k,\upsilon)} \varepsilon_{(j,s)} \varepsilon_{(k,\upsilon)} E\left[\underline{u}_{(i,t),n}^2 | \mathcal{F}_n^W\right]$$

For this term, we apply part (c) of Lemma S2-8 with $\phi_{(i,t)} = E\left[u_{(i,t),n}^2 | \mathcal{F}_n^W\right]$ with $u_{(i,t),n} = \left(\mu_n^{\min}\right) \underline{u}_{(i,t),n}$. From (9), there exists a positive constant C such that $E\left[u_{(i,t),n}^2 | \mathcal{F}_n^W\right]$ $= \left(\mu_n^{\min}\right)^2 E\left[\underline{u}_{(i,t),n}^2 | \mathcal{F}_n^W\right] \leq C < \infty$ $a.s.$ for all $(i,t) \in \{1, 2, ..., m_n\}$ and for all n sufficiently large, so that $\max_{(i,t)} \left|\phi_{(i,t)}\right| = \max_{1 \leq (i,t) \leq m_n} E\left[u_{(i,t),n}^2 | \mathcal{F}_n^W\right] \leq C$ $a.s.n.$ Hence, applying part (c) of Lemma S2-8, we obtain $\mathcal{TT}_5 \xrightarrow{p} 0$.

Finally, consider the term

$$\mathcal{TT}_6 = \frac{2\left(\mu_n^{\min}\right)^2}{K_{2,n}} \sum_{(i,t)=3}^{m_n} \sum_{(j,s)=2}^{(i,t)-1} \sum_{(k,\upsilon)=1}^{(j,s)-1} A_{(i,t),(j,s)} A_{(i,t),(k,\upsilon)} \underline{u}_{(j,s),n} \underline{u}_{(k,\upsilon),n} E\left[\varepsilon_{(i,t)}^2 | \mathcal{F}_n^W\right]$$

In this case, we apply part (d) of Lemma S2-8 with $u_{(j,s)} = \left(\mu_n^{\min}\right) \underline{u}_{(j,s),n}$, $u_{(k,\upsilon)} = \left(\mu_n^{\min}\right) \underline{u}_{(k,\upsilon),n}$, and $\phi_{(i,t)} = E\left[\varepsilon_{(i,t)}^2 | \mathcal{F}_n^W\right]$. Using a conditional version of Liapounov's inequality and Assumption 2(i), we obtain $E\left[\varepsilon_{(i,t)}^2 | \mathcal{F}_n^W\right] \leq \left(E\left[\varepsilon_{(i,t)}^4 | \mathcal{F}_n^W\right]\right)^{1/2} \leq C < \infty$ $a.s.$ for all $(i,t) \in \{1, 2, ..., m_n\}$ and for all n sufficiently large, so that $\max_{(i,t)} \left|\phi_{(i,t)}\right| = \max_{(i,t)} E\left[\varepsilon_{(i,t)}^2 | \mathcal{F}_n^W\right] \leq C$ $a.s.n.$ Moreover, the upper bound in (12) implies that $\max_{1 \leq (i,t) \leq m_n} E\left[u_{(i,t),n}^4 | \mathcal{F}_n^W\right] = \max_{1 \leq (i,t) \leq m_n} \left(\mu_n^{\min}\right)^4 E\left[\underline{u}_{(i,t),n}^4 | \mathcal{F}_n^W\right] \leq C$ $a.s.n.$ It follows by applying part (d) of Lemma S2-8 that $\mathcal{TT}_6 \xrightarrow{p} 0$.

It follows from the above calculations that the terms $\mathcal{TT}_k \xrightarrow{p} 0$ for each $k \in \{1, ..., 6\}$, which in light of equation (14) implies that $\sum_{(i,t)=2}^{m_n} E\left[\mathcal{J}_{(i,t),n}^2 | \mathcal{F}_{(i,t)-1,n}\right] - 1 = o_p(1)$. This establishes condition (21) of Lemma S2-15. It now follows from Lemma S2-15 that $\mathcal{J}_n$ $= \left(\mu_n^{\min}/\sqrt{K_{2,n}}\right) a' \left(\left(\mu_n^{\min}\right)^2 \widetilde{L}_n H_n^{-1} \Sigma_{2,n} H_n^{-1} \widetilde{L}_n' / K_{2,n}\right)^{-1/2} \widetilde{L}_n H_n^{-1} D_\mu^{-1} \underline{U}' A \varepsilon \xrightarrow{d} N(0,1)$. Since this result holds for all $a \in \mathbb{R}^d$ such that $\|a\|_2 = 1$, applying the Cramér-Wold device, we further deduce that

$$\left(\mu_n^{\min}/\sqrt{K_{2,n}}\right) \left(\widetilde{L}_n \Lambda_{II,n} \widetilde{L}_n'\right)^{-1/2} \widetilde{L}_n H_n^{-1} D_\mu^{-1} \underline{U}' A \varepsilon \xrightarrow{d} N(0, I_d), \tag{16}$$

where $\Lambda_{II,n} = \left(\mu_n^{\min}\right)^2 H_n^{-1} \Sigma_{2,n} H_n^{-1} / K_{2,n}$ with $H_n = \Gamma' M^{(Z_1,Q)} \Gamma / n$. Next, recall that $\widehat{\Delta}(\delta) = -\left[(y - X\delta)' M^{(Z_1,Q)} (y - X\delta)/2\right] \left[\partial \widehat{Q}_{FELIM}(\delta)/\partial \delta\right]$; and note that, by Lemma S2-10, we have $-D_\mu^{-1} \left(\partial \widehat{\Delta}\left(\overline{\delta}_n\right)/\partial \delta'\right) D_\mu^{-1} = H_n + o_p(1)$, with H_n being positive definite given Assumption 3(iii), so that upon inverting the expansion given in expression (5) above and multiplying by $\left(\mu_n^{\min}\right)/\sqrt{K_{2,n}}$, we obtain

$$\begin{aligned}\left(\mu_n^{\min}/\sqrt{K_{2,n}}\right) D_\mu \left(\widehat{\delta}_{L,n} - \delta_0\right) &= \left(\mu_n^{\min}/\sqrt{K_{2,n}}\right) H_n^{-1} D_\mu^{-1} \widehat{\Delta}\left(\delta_0\right)\left[1 + o_p\left(1\right)\right] \\ &= \left(\mu_n^{\min}/\sqrt{K_{2,n}}\right) H_n^{-1} D_\mu^{-1} \underline{U}' A \varepsilon \left[1 + o_p\left(1\right)\right],\end{aligned}$$

where the last equality comes from applying expression (8). It follows by multiplying both sides of the equation above by $\left(\widetilde{L}_n \Lambda_{II,n} \widetilde{L}_n\right)^{-1/2} \widetilde{L}_n$ and applying the result given in expression (16) that $\left(\mu_n^{\min}/\sqrt{K_{2,n}}\right)\left(\widetilde{L}_n \Lambda_{II,n} \widetilde{L}_n\right)^{-1/2} \widetilde{L}_n D_\mu \left(\widehat{\delta}_{L,n} - \delta_0\right) \xrightarrow{d} N\left(0, I_d\right)$.

Turning our attention now to $\widehat{\delta}_{F,n}$, note that, using expression (7) above, we can write

$$\begin{aligned}&\frac{\left(\mu_n^{\min}\right) D_\mu \left(\widehat{\delta}_{F,n} - \delta_0\right)}{\sqrt{K_{2,n}}} \\ =\; &\frac{\left(\mu_n^{\min}\right)\left(D_\mu^{-1} X'\left[A - \widehat{\ell}_{F,n} M^{(Z_1,Q)}\right] X D_\mu^{-1}\right)^{-1} D_\mu^{-1} X'\left[A - \widehat{\ell}_{F,n} M^{(Z_1,Q)}\right]\left(y - X\delta_0\right)}{\sqrt{K_{2,n}}}\end{aligned}$$

It follows by applying Lemmas S2-12 and S2-13 that

$$\frac{\mu_n^{\min}}{\sqrt{K_{2,n}}} D_\mu \left(\widehat{\delta}_{F,n} - \delta_0\right) = \frac{\mu_n^{\min}}{\sqrt{K_{2,n}}} H_n^{-1} D_\mu^{-1} \underline{U}' A \varepsilon + o_p\left(1\right), \tag{17}$$

noting that, in this case, $\left(\mu_n^{\min}\right)/\sqrt{K_{2,n}} = o\left(1\right)$ but $\sqrt{K_{2,n}}/\left(\mu_n^{\min}\right)^2 \to 0$. Again, let $\Lambda_{II,n} = \left(\mu_n^{\min}\right)^2 H_n^{-1} \Sigma_{2,n} H_n^{-1}/K_{2,n}$ and let $\widetilde{L}_n$ be any sequence of bounded $(l \times d)$ non-random matrices such that $\lambda_{\min}\left(\widetilde{L}_n \Lambda_{II,n} \widetilde{L}_n'\right) \geq \underline{C}$ *a.s.n.* It follows by multiplying both sides of equation (17) above by $\left(\widetilde{L}_n \Lambda_{II,n} \widetilde{L}_n\right)^{-1/2} \widetilde{L}_n$ and applying the result given in expression (16) that $\left(\mu_n^{\min}/\sqrt{K_{2,n}}\right)\left(\widetilde{L}_n \Lambda_{II,n} \widetilde{L}_n\right)^{-1/2} \widetilde{L}_n D_\mu \left(\widehat{\delta}_{F,n} - \delta_0\right) \xrightarrow{d} N\left(0, I_d\right)$. □

Appendix S2: Key Lemmas Used in Proving the Main Theorems

In this appendix, we state a number of lemmas that are used in the proofs of the main theorems of the paper. Proofs for these lemmas are available in a separate online appendix which can be viewed at the URL: http://econweb.umd.edu/~chao/Research/research_files/Additional_Online_Appendix_Jackknife_Estimation_Cluster_Sample_IV_Model.pdf

Lemma S2-1: Let $A = P^{\perp} - M^{(Z,Q)} D_{\widehat{\vartheta}} M^{(Z,Q)}$. Then, under Assumptions 2-7, the following statements hold as $K_{2,n}$, $n \to \infty$.

(a) $\sum_{(i,t),(j,s)=1,(i,t)\neq(j,s)}^{m_n} A_{(i,t),(j,s)}^2 = O_{a.s.}\left(K_{2,n}\right)$.

(b) $\sum_{(i,t),(j,s)=1,(i,t)\neq(j,s)}^{m_n} A_{(i,t),(j,s)}^4 = O_{a.s.}\left(K_{2,n}^3/n^2\right)$.

(c) $\sum_{(j,s)=1}^{m_n} \sum_{(i,t),(k,v)=1,(i,t)\neq(j,s),(k,v)\neq(j,s)}^{m_n} A_{(i,t),(j,s)}^2 A_{(j,s),(k,v)}^2 = O_{a.s.}\left(K_{2,n}^2/n\right)$.

(d) $\max_{1\leq(i,t)\leq m_n}\left(\sum_{(j,s)=1}^{m_n} A^2_{(i,t),(j,s)}\right) = O_{a.s.}\left(K_{2,n}/n\right)$.

Lemma S2-2: Suppose that Assumptions 1-7 are satisfied. Then, the following statements are true: (a) $D_\mu^{-1}X'M^{(Z_1,Q)}XD_\mu^{-1} = O_p\left(n\left(\mu_n^{\min}\right)^{-2}\right)$; (b) $D_\mu^{-1}X'AXD_\mu^{-1} = H_n + o_p(1)$, where $H_n = \Gamma' M^{(Z_1,Q)}\Gamma/n = O_p(1)$.

Lemma S2-3: Let $\underline{U} = U - \varepsilon\rho'$ and $\underline{U}_{(i,t)} = U_{(i,t)} - \rho\varepsilon_{(i,t)}$ and let $VC\left(X|\mathcal{F}_n^W\right)$ denote the conditional covariance matrix of the random vector X given $\mathcal{F}_n^W$. Under Assumptions 1-2, 5-6, and 8; there exists positive constants $0 < \underline{C} \leq \overline{C} < \infty$ such that the following statements are true.
(a) $\lambda_{\max}\left[VC\left(\Gamma'M^{(Z_1,Q)}\varepsilon/\sqrt{n}|\mathcal{F}_n^W\right)\right] \leq \overline{C}$ $a.s.$ and $\lambda_{\min}\left[VC\left(\Gamma'M^{(Z_1,Q)}\varepsilon/\sqrt{n}|\mathcal{F}_n^W\right)\right] \geq \underline{C}$ $a.s.$ for all n sufficiently large.
(b) $VC\left(\underline{U}'A\varepsilon/\sqrt{K_{2,n}}|\mathcal{F}_n^W\right) \geq \underline{C}I_d > \underset{d\times d}{0}$ $a.s.$, for all n sufficiently large.
(c) $\lambda_{\max}\left(VC\left[\underline{U}'A\varepsilon/\sqrt{K_{2,n}}|\mathcal{F}_n^W\right]\right) \leq \overline{C}$ $a.s.$, $\lambda_{\max}\left(VC\left[\underline{U}'A\varepsilon/\sqrt{K_{2,n}}\right]\right) \leq \overline{C}$, $\lambda_{\max}\left(VC\left[U'A\varepsilon/\sqrt{K_{2,n}}|\mathcal{F}_n^W\right]\right) \leq \overline{C}$ $a.s.$, and $\lambda_{\max}\left(VC\left[U'A\varepsilon/\sqrt{K_{2,n}}\right]\right) \leq \overline{C}$, for all n sufficiently large.
(d) For any $a \in \mathbb{R}^d$ with $\|a\|_2 = 1$ and for all n sufficiently large, $\lambda_{\min}(\Sigma_n) \geq \underline{C} > 0$ $a.s.$ and $a'\Sigma_n^{-1}a \leq \overline{C} < \infty$ $a.s.$, where $\Sigma_n = VC\left(\mathcal{Y}_n|\mathcal{F}_n^W\right) = \Sigma_{1,n} + \Sigma_{2,n}$, as defined in section 3 of the main paper, and where $\mathcal{Y}_n = \Gamma'M^{(Z_1,Q)}\varepsilon/\sqrt{n} + D_\mu^{-1}\underline{U}'A\varepsilon$.

Lemma S2-4: Under Assumptions 1-7, the following results hold: (a) $D_\mu^{-1}X'A\varphi_n = O_p\left(\tau_n/K_{1,n}^{\varrho_g}\right)$;
(b) $D_\mu^{-1}X'A\varepsilon = \frac{\Gamma'M^{(Z_1,Q)}\varepsilon}{\sqrt{n}} + D_\mu^{-1}U'A\varepsilon + O_p\left(K_{2,n}^{-\varrho_\gamma}\right) + O_p\left(K_{2,n}^{-(\varrho_\gamma-1)}n^{-1}\right) + O_p\left(\kappa_n^{\max}/\left(\mu_n^{\min}K_{1,n}^{\varrho_f}\right)\right) = O_p\left(\max\left\{1, \sqrt{K_{2,n}}/\left(\mu_n^{\min}\right)\right\}\right)$

Lemma S2-5: Under Assumptions 1-7, the following results hold: (a) $D_\mu^{-1}X'M^{(Z_1,Q)}\varphi_n = O_p\left(\tau_n/K_{1,n}^{\varrho_g}\right)$; (b) $D_\mu^{-1}X'M^{(Z_1,Q)}\varepsilon = O_p\left(n/\mu_n^{\min}\right)$.

Lemma S2-6: Suppose that Assumptions 2 and 8 hold. For $1 \leq p \leq 8$ and for all n, there exists a positive constant C such that $\max_{1\leq(i,t)\leq m_n} E\left[\left\|\underline{U}_{(i,t)}\right\|_2^p|\mathcal{F}_n^W\right] \leq C < \infty$ $a.s.$, where $\underline{U}_{(i,t)} = U_{(i,t)} - \rho\varepsilon_{(i,t)}$.

Lemma S2-7: Under Assumptions 1-7, the following results hold: (a) $\widehat{\ell}_{L,n} = o_p\left(\left[\mu_n^{\min}\right]^2/n\right)$; (b) $\widehat{\ell}_{F,n} = o_p\left(\left[\mu_n^{\min}\right]^2/n\right)$.

Lemma S2-8: Let A be as defined above. Suppose that i) $\left(u_{(1,1),n}, \varepsilon_{(1,1)}\right), ..., \left(u_{(1,T_1),n}, \varepsilon_{(1,T_1)}\right)$, $\left(u_{(2,1),n}, \varepsilon_{(2,1),n}\right), ..., \left(u_{(2,T_2),n}, \varepsilon_{(2,T_2),n}\right), ..., \left(u_{(n,1),n}, \varepsilon_{(n,1),n}\right) ...., \left(u_{(n,T_n),n}, \varepsilon_{(n,T_n),n}\right)$ are independent conditional on $\mathcal{F}_n^W = \sigma(W_n)$; ii) there exists a constant C such that, almost surely for all n sufficiently large, $\max_{1\leq(i,t)\leq m_n} E\left(u^4_{(i,t),n}|\mathcal{F}_n^W\right) \leq C$, $\max_{1\leq(i,t)\leq m_n} E\left(\varepsilon^4_{(i,t),n}|\mathcal{F}_n^W\right) \leq C$, and $\max_{1\leq(i,t)\leq m_n}\left|\phi_{(i,t),n}\right| \leq C$. In addition, define $\overline{\psi}_{(j,s),n} = E\left[u_{(j,s),n}\varepsilon_{(j,s),n}|\mathcal{F}_n^W\right]$ for $(j,s) = 1, ..., m_n$. Then, under Assumptions 5 and 6, the following statements are true:
(a) $K_{2,n}^{-1}\sum_{1\leq(j,s)<(i,t)\leq m_n} A^2_{(i,t),(j,s)}\phi_{(i,t),n}\left\{u_{(j,s),n}\varepsilon_{(j,s),n} - \overline{\psi}_{(j,s),n}\right\} \xrightarrow{p} 0$;
(b) $K_{2,n}^{-1}\sum_{1\leq (k,v)<(j,s)<(i,t)\leq m_n} A_{(i,t),(j,s)}A_{(i,t),(k,v)}\phi_{(i,t),n}\left\{u_{(j,s),n}\varepsilon_{(k,v),n} + \varepsilon_{(j,s),n}u_{(k,v),n}\right\} \xrightarrow{p} 0$;
(c) $K_{2,n}^{-1}\sum_{1\leq (k,v)<(j,s)<(i,t)\leq m_n} A_{(i,t),(j,s)}A_{(i,t),(k,v)}\phi_{(i,t),n}\varepsilon_{(j,s),n}\varepsilon_{(k,v),n} \xrightarrow{p} 0$;
(d) $K_{2,n}^{-1}\sum_{1\leq (k,v)<(j,s)<(i,t)\leq m_n} A_{(i,t),(j,s)}A_{(i,t),(k,v)}\phi_{(i,t),n}u_{(j,s),n}u_{(k,v),n} \xrightarrow{p} 0$.

Lemma S2-9: Let

$$\widehat{\Delta}(\delta_0) = -\frac{(y - X\delta_0)' M^{(Z_1,Q)} (y - X\delta_0)}{2} \frac{\partial}{\partial \delta} \left\{ \frac{(y - X\delta)' A (y - X\delta)}{(y - X\delta)' M^{(Z_1,Q)} (y - X\delta)} \right\} \Bigg|_{\delta=\delta_0}.$$

Suppose that Assumptions 1-8 hold; then, $D_\mu^{-1}\widehat{\Delta}(\delta_0) = \Gamma' M^{(Z_1,Q)}\varepsilon/\sqrt{n} + D_\mu^{-1}\underline{U}' A\varepsilon + o_p(1)$, where $\underline{U} = U - \varepsilon\rho'$ and where $\rho = \lim_{n\to\infty} E\left[U' M^Q \varepsilon\right] / E\left[\varepsilon' M^Q \varepsilon\right]$.

Lemma S2-10: Suppose that Assumptions 1-7 are satisfied. Let $\overline{\delta}_n$ be any estimator such that, as $n \to \infty$, $D_\mu\left(\overline{\delta}_n - \delta_0\right)/\mu_n^{\min} = o_p(1)$. Then, $-D_\mu^{-1}\left(\partial\widehat{\Delta}\left(\overline{\delta}_n\right)/\partial\delta'\right) D_\mu^{-1} = H_n + o_p(1)$, where $H_n = \Gamma' M^{(Z_1,Q)}\Gamma/n$ and where
$\widehat{\Delta}(\delta) = -\left[(y - X\delta)' M^{(Z_1,Q)} (y - X\delta)/2\right]\left[\partial\widehat{Q}_{FELIM}(\delta)/\partial\delta\right]$
$= X' A (y - X\delta) - \widehat{\ell}(\delta) X' M^{(Z_1,Q)} (y - X\delta)$, with
$\widehat{\ell}(\delta) = (y - X\delta)' A (y - X\delta) / \left[(y - X\delta)' M^{(Z_1,Q)} (y - X\delta)\right]$. In addition, we also have

$$D_\mu^{-1} X'\left[A - \widehat{\ell}\left(\overline{\delta}_n\right) M^{(Z_1,Q)}\right] X D_\mu^{-1} = H_n + o_p(1). \tag{18}$$

Lemma S2-11: Let $\widehat{\ell}_L = Q\left(\widetilde{\beta}\right) = \min_{\beta\in\overline{B}} Q(\beta)$, where $Q(\beta)$ is as defined in Assumption 9. Then, $\widehat{\ell}_L$ is also the smallest root of the determinantal equation $\det\left[\overline{X}' A\overline{X} - \ell\overline{X}' M^{(Z_1,Q)}\overline{X}\right] = 0$, where $\overline{X} = [y, X]$. Suppose in addition that condition (11) in Assumption 9 is satisfied; then, $\widehat{\ell}_L$ has the representation

$$\widehat{\ell}_L = \frac{\left(y - X\widehat{\delta}_L\right)' A \left(y - X\widehat{\delta}_L\right)}{\left(y - X\widehat{\delta}_L\right)' M^{(Z_1,Q)} \left(y - X\widehat{\delta}_L\right)}, \tag{19}$$

where $\widehat{\delta}_L$ denotes the FELIM estimator. Moreover, $\overline{X}' A\left(y - X\widehat{\delta}_L\right) - \widehat{\ell}_L\overline{X}' M^{(Z_1,Q)}\left(y - X\widehat{\delta}_L\right)$ $= 0$. In particular, this implies that $\widehat{\Delta}\left(\widehat{\delta}_L\right) = 0$, where
$\widehat{\Delta}(\delta) = -\left[(y - X\delta)' M^{(Z_1,Q)} (y - X\delta)/2\right]\left(\partial\widehat{Q}_{FELIM}(\delta)/\partial\delta\right)$, so that $\widehat{\delta}_L$ satisfies the set of (normalized) first-order conditions for minimizing the variance ratio objective function $\widehat{Q}_{FELIM}(\delta) = (y - X\delta)' A (y - X\delta) / \left[(y - X\delta)' M^{(Z_1,Q)} (y - X\delta)\right]$.

Lemma S2-12: Suppose that Assumptions 1-7 are satisfied. Then,
$D_\mu^{-1} X'\left[A - \widehat{\ell}_{F,n} M^{(Z_1,Q)}\right] X D_\mu^{-1} = H_n + o_p(1)$, where $H_n = \Gamma' M^{(Z_1,Q)}\Gamma/n$,
$\widehat{\ell}_{F,n} = \left[\widehat{\ell}_{L,n} - \left(1 - \widehat{\ell}_{L,n}\right)(C/m_n)\right] / \left[1 - \left(1 - \widehat{\ell}_{L,n}\right)(C/m_n)\right]$, and $\widehat{\ell}_{L,n}$ is smallest root of the determinantal equation $\det\left\{\overline{X}' A\overline{X} - \ell\overline{X}' M^{(Z_1,Q)}\overline{X}\right\} = 0$, with $\overline{X} = \left[\begin{array}{cc} y & X \end{array}\right]$.

Lemma S2-13: Suppose that Assumptions 1-8 hold. Then, $D_\mu^{-1} X'\left[A - \widehat{\ell}_{F,n} M^{(Z_1,Q)}\right](y - X\delta_0) = \mathcal{Y}_n\left[1 + o_p(1)\right]$, where $\mathcal{Y}_n = \Gamma' M^{(Z_1,Q)}\varepsilon/\sqrt{n} + D_\mu^{-1}\underline{U}' A\varepsilon$ with $\underline{U} = U - \varepsilon\rho'$ and
$\rho = \lim_{n\to\infty} E\left[U' M^Q\varepsilon\right] / E\left[\varepsilon' M^Q\varepsilon\right]$.

Lemma S2-14: For any $a \in \mathbb{R}^d$ such that $\|a\| = 1$, define $b_{1n} = \Sigma_n^{-1/2} a$, $\underline{u}_{2,(i,t),n} = b_{2n}'\underline{U}_{(i,t)}$ $= \sqrt{K_{2,n}} a' \Sigma_n^{-1/2} D_\mu^{-1}\underline{U}_{(i,t)}$, $\sigma_{(i,t),n}^2 = E\left[\varepsilon_{(i,t)}^2 | \mathcal{F}_n^W\right]$, $\widetilde{\psi}_{(i,t),n} = E\left[\underline{u}_{2,(i,t),n}\varepsilon_{(i,t)} | \mathcal{F}_n^W\right]$, and $\widetilde{\omega}_{(i,t)}^2 =$

$E\left[\underline{u}^2_{2,(i,t),n}|\mathcal{F}^W_n\right]$. Suppose that Assumptions 1-2 and 5-6 are satisfied. Then, the following statements are true.

(a) $\sum_{(i,t)=2}^{m_n}\sum_{(j,s)=1}^{(i,t)-1}\left[b'_{1n}\Gamma' M^{(Z_1,Q)}e_{(i,t)}/\sqrt{n}\right]\left(A_{(i,t),(j,s)}/\sqrt{K_{2,n}}\right)\left\{\varepsilon_{(j,s)}\widetilde{\psi}_{(i,t),n}+\underline{u}_{2,(j,s)}\sigma^2_{(i,t),n}\right\} = O_p\left(K^{1/4}_{2,n}/\mu^{\min}_n\right) = o_p(1)$.

(b) $\sum_{(i,t)=2}^{m_n}\sum_{(j,s)=1}^{(i,t)-1}\left(A^2_{(i,t),(j,s)}/K_{2,n}\right)\left(\varepsilon^2_{(j,s)}-\sigma^2_{(j,s),n}\right)\widetilde{\omega}^2_{(i,t),n} = O_p\left(K_{2,n}\left(\mu^{\min}_n\right)^{-2}n^{-1/2}\right) = o_p(1)$.

(c) $\sum_{(i,t)=2}^{m_n}\sum_{(j,s)=1}^{(i,t)-1}\left(A^2_{(i,t),(j,s)}/K_{2,n}\right)\left(\underline{u}^2_{2,(j,s),n}-\widetilde{\omega}^2_{(j,s),n}\right)\sigma^2_{(i,t),n} = O_p\left(K_{2,n}\left(\mu^{\min}_n\right)^{-2}n^{-1/2}\right) = o_p(1)$.

Lemma S2-15 (Gänsler and Stute, 1977): Let $\{X_{i,n},\mathcal{F}_{i,n},1\le i\le k_n,n\ge 1\}$ be a square integrable martingale difference array. Suppose that for all $\epsilon>0$

$$\sum_{i=1}^{k_n}E\left[X^2_{i,n}\mathbb{I}\{|X_{i,n}|>\epsilon\}|\mathcal{F}_{i-1,n}\right]\xrightarrow{p}0 \tag{20}$$

and

$$\sum_{i=1}^{k_n}E\left[X^2_{i,n}|\mathcal{F}_{i-1,n}\right]\xrightarrow{p}1. \tag{21}$$

Then, $\sum_{i=1}^{k_n}X_{i,n}\xrightarrow{d}N(0,1)$.

Remark: Note that a sufficient condition for condition (20), which we will verify in lieu of (20) in the proof of Theorems 2 and 3 in Appendix S1, is the following

$$\sum_{i=1}^{k_n}E\left[|X_{i,n}|^{2+\delta}\right]\xrightarrow{p}0,\text{ for some }\delta>0. \tag{22}$$

Lemma S2-16: Let $\widetilde{L}_n$ be a sequence of $l\times d$, nonrandom matrices (with $l\le d$) such that $\left\|\widetilde{L}_n\right\|^2_F\le\overline{C}<\infty$ for some constant $\overline{C}$, and let $\Sigma_{2,n}=VC\left(D^{-1}_\mu\underline{U}'A\varepsilon|\mathcal{F}^W_n\right)$
$=D^{-1}_\mu VC\left(\underline{U}'A\varepsilon|\mathcal{F}^W_n\right)D^{-1}_\mu$. Suppose that there exists a positive constant $\underline{C}$ such that
$\lambda_{\min}\left(\left(\mu^{\min}_n\right)^2\widetilde{L}_nH^{-1}_n\Sigma_{2,n}H^{-1}_n\widetilde{L}'_n/K_{2,n}\right)\ge\underline{C}>0\ a.s.n.$ Furthermore, let $a\in\mathbb{R}^d$ such that $\|a\|_2=1$ and let $\underline{u}_{a,(i,t),n}=a'\left(\left(\mu^{\min}_n\right)^2\widetilde{L}_nH^{-1}_n\Sigma_{2,n}H^{-1}_n\widetilde{L}'_n/K_{2,n}\right)^{-1/2}\widetilde{L}_nD^{-1}_\mu\underline{U}_{(i,t)}$. Suppose that Assumptions 1-2 and 5-6 are satisfied and that $\left(\mu^{\min}_n\right)^2/K_{2,n}=o(1)$ but
$\sqrt{K_{2,n}}/\left(\mu^{\min}_n\right)^2\to 0$. Under these conditions, the following statements are true:

(a) $\left[\left(\mu^{\min}_n\right)^2/K_{2,n}\right]\sum_{(i,t)=2}^{m_n}\sum_{(j,s)=1}^{(i,t)-1}A^2_{(i,t),(j,s)}\left(\underline{u}^2_{a,(j,s),n}-E\left[\underline{u}^2_{a,(j,s),n}|\mathcal{F}^W_n\right]\right)E\left[\varepsilon^2_{(i,t)}|\mathcal{F}^W_n\right]$
$= O_p\left(n^{-1/2}\right)=o_p(1)$;

(b) $\left[\left(\mu^{\min}_n\right)^2/K_{2,n}\right]\sum_{(i,t)=2}^{m_n}\sum_{(j,s)=1}^{(i,t)-1}A^2_{(i,t),(j,s)}\left(\varepsilon^2_{(j,s)}-E\left[\varepsilon^2_{(j,s)}|\mathcal{F}^W_n\right]\right)E\left[\underline{u}^2_{a,(i,t),n}|\mathcal{F}^W_n\right]$
$= O_p\left(n^{-1/2}\right)=o_p(1)$.

Lemma S2-17 Under Assumptions 1-7, $D_{\mu}^{-1}X'AD\left(\varepsilon\circ\varepsilon\right)AXD_{\mu}^{-1}=\Sigma_{1,n}$ $+\sum_{(i,t),(j,s)=1,(i,t)\neq(j,s)}^{m_n}A_{(i,t),(j,s)}^{2}\sigma_{(i,t)}^{2}D_{\mu}^{-1}\Psi_{(j,s)}D_{\mu}^{-1}+o_p\left(1\right)$, where $\Sigma_{1,n}=\Gamma'M^{(Z_1,Q)}D_{\sigma^2}M^{(Z_1,Q)}\Gamma/n$, $\sigma_{(i,t)}^{2}=E\left[\varepsilon_{(i,t)}^{2}|\mathcal{F}_n^W\right]$, $D_{\sigma^2}=diag\left(\sigma_{(1,1)}^{2},...,\sigma_{(n,T_n)}^{2}\right)$, and $\Psi_{(j,s)}=E\left[U_{(j,s)}U_{(j,s)}'|\mathcal{F}_n^W\right]$.

Lemma S2-18 Suppose that Assumptions 1-8 are satisfied, and let $\left\{\widehat{\delta}_n\right\}$ be any sequence of estimators such that $\left\|\widehat{\delta}_n-\delta_0\right\|_2\xrightarrow{p}0$ as $n\to\infty$, as long as $\sqrt{K_{2,n}}/\left(\mu_n^{\min}\right)^2\to0$. Also, define the following notations: let $\widehat{\varepsilon}=M^{(Z,Q)}\left(y-X\widehat{\delta}_n\right)$, $J=\left[M^Q\circ M^Q\right]^{-1}$, $S_1=X'AD\left(J\left[\widehat{\varepsilon}\circ\widehat{\varepsilon}\right]\right)AX$, $S_2=\left(\widehat{\varepsilon}\circ\widehat{\varepsilon}\right)'J\left(A\circ A\right)J\left(\widehat{\varepsilon}\iota_d'\circ M^{(Z,Q)}X\right)$, $\underline{S}_2=\left(\widehat{\varepsilon}\circ\widehat{\varepsilon}\right)'J\left(A\circ A\right)J\left(\widehat{\varepsilon}\iota_d'\circ\underline{\widehat{U}}\right)$ with $\underline{\widehat{U}}=M^{(Z,Q)}X-\widehat{\varepsilon}\widehat{\rho}_n'$, $S_3=\left(\widehat{\varepsilon}\circ\widehat{\varepsilon}\right)'J\left(A\circ A\right)J\left(\widehat{\varepsilon}\circ\widehat{\varepsilon}\right)$, $S_4=\left(\widehat{\varepsilon}\iota_d'\circ M^{(Z,Q)}X\right)'J\left(A\circ A\right)J\left(\widehat{\varepsilon}\iota_d'\circ M^{(Z,Q)}X\right)$, $\underline{S}_4=\left(\widehat{\varepsilon}\iota_d'\circ\underline{\widehat{U}}\right)'J\left(A\circ A\right)J\left(\widehat{\varepsilon}\iota_d'\circ\underline{\widehat{U}}\right)$, and $\Sigma_{1,n}=\Gamma'M^{(Z_1,Q)}D_{\sigma^2}M^{(Z_1,Q)}\Gamma/n$. In addition, define $\sigma_{(i,t)}^{2}=E\left[\varepsilon_{(i,t)}^{2}|\mathcal{F}_n^W\right]$, $D_{\sigma^2}=diag\left(\sigma_{(1,1)}^{2},...,\sigma_{(n,T_n)}^{2}\right)$, $\phi_{(i,t)}=E\left[U_{(i,t)}\varepsilon_{(i,t)}|\mathcal{F}_n^W\right]$, $\Psi_{(i,t)}=E\left[U_{(i,t)}U_{(i,t)}'|\mathcal{F}_n^W\right]$, $\underline{\phi}_{(i,t)}=E\left[\underline{U}_{(i,t)}\varepsilon_{(i,t)}|\mathcal{F}_n^W\right]$, and $\underline{\Psi}_{(i,t)}=E\left[\underline{U}_{(i,t)}\underline{U}_{(i,t)}'|\mathcal{F}_n^W\right]$ where $\underline{U}_{(i,t)}=U_{(i,t)}-\rho\varepsilon_{(i,t)}$ and where for notational convenience we suppress the dependence of $\sigma_{(i,t)}^{2}$, $\phi_{(i,t)}$, $\Psi_{(i,t)}$, $\underline{\phi}_{(i,t)}$, and $\underline{\Psi}_{(i,t)}$ on $\mathcal{F}_n^W=\sigma\left(W_n\right)$. Then, under the above conditions, the following statements are true.

(a) $D_{\mu}^{-1}S_1D_{\mu}^{-1}=\Sigma_{1,n}+\sum_{(i,t),(j,s)=1,(i,t)\neq(j,s)}^{m_n}A_{(i,t),(j,s)}^{2}\sigma_{(i,t)}^{2}D_{\mu}^{-1}\Psi_{(j,s)}D_{\mu}^{-1}$ $+o_p\left(\max\left\{1,K_{2,n}\left(\mu_n^{\min}\right)^{-2}\right\}\right)$.

(b) $S_3/K_{2,n}-K_{2,n}^{-1}\sum_{(i,t),(j,s)=1,(i,t)\neq(j,s)}^{m_n}A_{(i,t),(j,s)}^{2}\sigma_{(i,t)}^{2}\sigma_{(j,s)}^{2}=o_p\left(1\right)$.

(c) $D_{\mu}^{-1}S_4D_{\mu}^{-1}-\sum_{(i,t),(j,s)=1,(i,t)\neq(j,s)}^{m_n}A_{(i,t),(j,s)}^{2}D_{\mu}^{-1}\phi_{(i,t)}\phi_{(j,s)}'D_{\mu}^{-1}=o_p\left(K_{2,n}\left(\mu_n^{\min}\right)^{-2}\right)$.

(d) $\left(\mu_n^{\min}/K_{2,n}\right)S_2D_{\mu}^{-1}-\left(\mu_n^{\min}/K_{2,n}\right)\sum_{(i,t),(j,s)=1,(i,t)\neq(j,s)}^{m_n}A_{(i,t),(j,s)}^{2}\sigma_{(i,t)}^{2}\phi_{(j,s)}'D_{\mu}^{-1}=o_p\left(1\right)$.

(e) $D_{\mu}^{-1}\widehat{\rho}_n=O_p\left(\left(\mu_n^{\min}\right)^{-1}\right)$ and $D_{\mu}^{-1}\left(\widehat{\rho}_n-\rho\right)=o_p\left(\left(\mu_n^{\min}\right)^{-1}\right)$, where $\rho=\lim_{n\to\infty}\rho_n=\lim_{n\to\infty}\left(E\left[U'M^Q\varepsilon\right]/n\right)/\left(E\left[\varepsilon'M^Q\varepsilon\right]/n\right)$.

(f) $D_{\mu}^{-1}\underline{S}_4D_{\mu}^{-1}-\sum_{(i,t),(j,s)=1,(i,t)\neq(j,s)}^{m_n}A_{(i,t),(j,s)}^{2}D_{\mu}^{-1}\underline{\phi}_{(i,t)}\underline{\phi}_{(j,s)}'D_{\mu}^{-1}=o_p\left(K_{2,n}\left(\mu_n^{\min}\right)^{-2}\right)$.

(g) $\left(\mu_n^{\min}/K_{2,n}\right)-\left(\mu_n^{\min}/K_{2,n}\right)\sum_{(i,t),(j,s)=1,(i,t)\neq(j,s)}^{m_n}A_{(i,t),(j,s)}^{2}\sigma_{(i,t)}^{2}\underline{\phi}_{(j,s)}'D_{\mu}^{-1}=o_p(1)$.

References

[1] Billingsley, P. (1995). *Probability and Measure.* New York: John Wiley & Sons.

[2] Gänsler, P. and W. Stute (1977). *Wahrscheinlichkeitstheorie*. New York: Springer-Verlag