Supplemental Appendix to Jackknife Estimation of a Cluster-Sample IV Regression Model with Many Weak Instruments (not intended for publication)*

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

April 8, 2021

Abstract

This Supplemental Appendix is comprised of two sub-appendices. Appendix S1 gives the proofs for Theorems 2 and 3 stated in section 3 of the main paper. Appendix S2, on the other hand, states and proves additional supporting lemmas whose results are used to prove the main theorems of the paper.

Appendix S1: Proofs 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 $\mathcal{L}_{(i,t),n} = b_{1n}'\Gamma' M^{(Z_1,Q)} e_{(i,t)}\varepsilon_{(i,t)}/\sqrt{n}$

*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.

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, $\Sigma_n = VC\left(\mathcal{Y}_n | \mathcal{F}_n^W\right) = \Sigma_{1,n} + \Sigma_{2,n}$, 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 $a'\Sigma_n^{-1/2} \mathcal{Y}_n = \mathcal{L}_{(1,1),n} + \sum_{(i,t)=2}^{m_n} \left\{ \mathcal{L}_{(i,t),n} + \mathcal{N}_{(i,t),n} \right\}$. 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);

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

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 (37), which, as noted in the remark which follows Lemma S2-15, is a sufficient condition for condition (35) of Lemma S2-15. We shall verify (37) 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 (37), 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

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

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

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

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

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

In a similar way, we can also show that $E_{W_n}(\mathcal{R}_2)=\left(8/K_{2,n}^2\right)E\left[\sum\nolimits_{(i,t)=2}^{m_n}\left(\sum\nolimits_{(j,s)=1}^{(i,t)-1} A_{(j,s),(i,t)}\underline{u}_{2,(j,s)}\varepsilon_{(i,t)}\right)^4\right]=o(1)$. It follows that $\sum\nolimits_{(i,t)=2}^{m_n} E\left[\mathcal{N}_{(i,t),n}^4\right]\leq E_{W_n}(\mathcal{R}_1)+E_{W_n}(\mathcal{R}_2)=o(1)$. This verifies condition (37).

Next, we verify condition (36) 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

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

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

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

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

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

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

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

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

In this case, we apply part (a) of Lemma S2-8 with $u_{(j,s)} = \underline{u}_{2,(j,s)}$, $\overline{\psi}_{(j,s)} = E\left[\underline{u}_{2,(j,s)}\varepsilon_{(j,s)}|\mathcal{F}^W_n\right]$, and $\phi_{(i,t)} = E\left[\underline{u}_{2,(i,t)}\varepsilon_{(i,t)}|\mathcal{F}^W_n\right]$. 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 $\mathcal{F}^W_n$, and $\sup_{1\leq(i,t)\leq m_n}E\left[\varepsilon^4_{(i,t)}|\mathcal{F}^W_n\right] \leq C$ $a.s.$ 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 $K_{2,n}/\left(\mu^{\min}_n\right)^2 = O(1)$ in this case together imply that there exists a constant $C \geq 1$ such that $E\left[\underline{u}^4_{2,(i,t)}|\mathcal{F}^W_n\right] \leq \left[K^2_{2,n}/\left(\mu^{\min}_n\right)^4\right]E\left[\left\|\underline{U}_{(i,t)}\right\|^4_2|\mathcal{F}^W_n\right]\left(a'\Sigma^{-1}_n a\right)^2 \leq C < \infty$ $a.s.$ for all $(i,t) \in \{1,2,...,m_n\}$ and for all n sufficiently large, so that $\max_{1\leq(i,t)\leq m_n}E\left[\underline{u}^4_{2,(i,t)}|\mathcal{F}^W_n\right] \leq C$ $a.s.n.$ Finally, using the upper bound derived in expression (30) in the proof of part (a) of Lemma S2-14, we obtain $\max_{1\leq(i,t)\leq m_n}\left|\phi_{(i,t)}\right| \leq \max_{1\leq(i,t)\leq m_n}E\left[\left|\underline{u}_{2,(i,t)}\varepsilon_{(i,t)}\right||\mathcal{F}^W_n\right] \leq C$ $a.s.n.$ and $\max_{1\leq(j,s)\leq m_n}\left|\overline{\psi}_{(j,s)}\right| \leq \max_{1\leq(i,t)\leq m_n}E\left[\left|\underline{u}_{2,(j,s)}\varepsilon_{(j,s)}\right||\mathcal{F}^W_n\right] \leq C$ $a.s.n.$ 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 $\phi_{(i,t)} = E\left[\underline{u}_{2,(i,t)}\varepsilon_{(i,t)}|\mathcal{F}^W_n\right]$. Note again that $\left\{\left(\underline{u}_{2,(i,t)},\varepsilon_{(i,t)}\right)\right\}_{(i,t)=1}^{m_n}$ is independent conditional on $\mathcal{F}^W_n$, and $\max_{1\leq(i,t)\leq m_n}E\left[\varepsilon^4_{(i,t)}|\mathcal{F}^W_n\right] \leq C$ $a.s.$ by Assumptions 1 and 2(i), respectively. Moreover, previously, we have shown that $E\left[\underline{u}^4_{2,(i,t)}|\mathcal{F}^W_n\right] \leq C$ $a.s.n.$ 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 $\phi_{(i,t)} = E\left[\underline{u}^2_{2,(i,t)}|\mathcal{F}^W_n\right]$. From (1), there exists a positive constant C such that $E\left[\underline{u}^2_{2,(i,t)}|\mathcal{F}^W_n\right] \leq C < \infty$ $a.s.$ for all $(i,t) \in \{1,2,...,m_n\}$ and for all n sufficiently large, so that $\max_{1\leq(i,t)\leq m_n}\left|\phi_{(i,t)}\right| = \max_{1\leq(i,t)\leq m_n}E\left[\underline{u}^2_{2,(i,t)}|\mathcal{F}^W_n\right] \leq C$ $a.s.n.$ Hence, applying part (c) of Lemma 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 $\phi_{(i,t)} = E\left[\varepsilon^2_{(i,t)}|\mathcal{F}^W_n\right]$. 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 (36) 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

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

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

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

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. $\square$

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

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

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, $\left(\left(\mu_n^{\min}\right)^2\widetilde{L}_nH_n^{-1}\Sigma_{2,n}H_n^{-1}\widetilde{L}_n'/K_{2,n}\right)^{-1/2}$ is well-defined for all n sufficiently large. Hence, we can let $\widetilde{\mathcal{N}}_n=\left(\left(\mu_n^{\min}\right)^2\widetilde{L}_nH_n^{-1}\Sigma_{2,n}H_n^{-1}\widetilde{L}_n'/K_{2,n}\right)^{-1/2}\widetilde{L}_nH_n^{-1}\left(\mu_n^{\min}/\sqrt{K_{2,n}}\right)D_\mu^{-1}\underline{U}'A\varepsilon$ 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 $\underline{u}_{(i,t),n}=a'\left(\left(\mu_n^{\min}\right)^2\widetilde{L}_nH_n^{-1}\Sigma_{2,n}H_n^{-1}\widetilde{L}_n'/K_{2,n}\right)^{-1/2}\widetilde{L}_nH_n^{-1}D_\mu^{-1}\underline{U}_{(i,t)}$, 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

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

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 $Var\left(\mathcal{J}_{(i,t),n}\right)=E_W\left[E\left(\mathcal{J}_{(i,t),n}^2|\mathcal{F}_n^W\right)\right]\leq$

$\left(4\overline{C}^2/K_{2,n}\right)\sum_{(j,s)=1}^{(i,t)-1} E_W\left[A^2_{(i,t),(j,s)}\right] \leq C$ 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 (37) 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

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

where the second inequality above follows from Assumption 2(i) and from an upper bound on the conditional fourth moment of

$\underline{u}_{(i,t),n} = a' \left(\left(\mu_n^{\min} \right)^2 \widetilde{L}_n H_n^{-1} \Sigma_{2,n} H_n^{-1} \widetilde{L}_n' / K_{2,n} \right)^{-1/2} \widetilde{L}_n H_n^{-1} D_\mu^{-1} \underline{U}_{(i,t)}$ given by

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

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 $\lambda_{\min}\left(\left(\mu_n^{\min} \right)^2 \widetilde{L}_n H_n^{-1} \Sigma_{2,n} H_n^{-1} \widetilde{L}_n' / K_{2,n} \right) \geq \underline{C} > 0 \quad a.s.n.$ 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 (37) of Lemma S2-15 then follows from these calculations since

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

Next, we verify condition (36) 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

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

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

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

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

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

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

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 $\max_{1 \leq (i,t) \leq m_n} E\left[\varepsilon_{(i,t)}^4 | \mathcal{F}_n^W\right] \leq C$ $a.s.$ 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 $\max_{1 \leq (i,t) \leq m_n} E\left[u_{(i,t),n}^4 | \mathcal{F}_n^W\right]=\max_{1 \leq (i,t) \leq m_n}\left(\mu_n^{\min}\right)^4 E\left[\underline{u}_{(i,t),n}^4 | \mathcal{F}_n^W\right] \leq\left(\mu_n^{\min}\right)^4 C^* /\left(\mu_n^{\min}\right)^4=C^*$ $a.s.n.$ Finally, note that, by using the fact that $\underline{u}_{(i,t),n}=a^{\prime}\left(\left(\mu_n^{\min}\right)^2 \widetilde{L}_n H_n^{-1} \Sigma_{2,n} H_n^{-1} \widetilde{L}_n^{\prime} / K_{2,n}\right)^{-1/2} \widetilde{L}_n H_n^{-1} D_\mu^{-1} \underline{U}_{(i,t)}$ and by applying Assumption 2(i), Lemma S2-6, and the assumption that $\lambda_{\min}\left(\left(\mu_n^{\min}\right)^2 \widetilde{L}_n H_n^{-1} \Sigma_{2,n} H_n^{-1} \widetilde{L}_n^{\prime} / K_{2,n}\right) \geq \underline{C}>0$ $a.s.n.$; we can show that there exists a constant $C>0$ such that

$$
\begin{aligned}
& E\left[\left|\left(\mu_n^{\min}\right) \underline{u}_{(i,t),n} \varepsilon_{(i,t)}\right| | \mathcal{F}_n^W\right] \\
=\ & \left(\mu_n^{\min}\right) E\left[\left|\varepsilon_{(i,t)} \underline{U}_{(i,t)}^{\prime} D_\mu^{-1} H_n^{-1} \widetilde{L}_n^{\prime}\left(\frac{\left(\mu_n^{\min}\right)^2 \widetilde{L}_n H_n^{-1} \Sigma_{2,n} H_n^{-1} \widetilde{L}_n^{\prime}}{K_{2,n}}\right)^{-1/2} a\right| | \mathcal{F}_n^W\right] \\
\leq\ & \left(\mu_n^{\min}\right) \sqrt{E\left[\varepsilon_{(i,t)}^2 | \mathcal{F}_n^W\right]}\left[a^{\prime}\left(\frac{\left(\mu_n^{\min}\right)^2 \widetilde{L}_n H_n^{-1} \Sigma_{2,n} H_n^{-1} \widetilde{L}_n^{\prime}}{K_{2,n}}\right)^{-1/2} \widetilde{L}_n H_n^{-1} D_\mu^{-1}\right. \\
& \left.\times E\left[\underline{U}_{(i,t)} \underline{U}_{(i,t)}^{\prime} | \mathcal{F}_n^W\right] D_\mu^{-1} H_n^{-1} \widetilde{L}_n^{\prime}\left(\frac{\left(\mu_n^{\min}\right)^2 \widetilde{L}_n H_n^{-1} \Sigma_{2,n} H_n^{-1} \widetilde{L}_n^{\prime}}{K_{2,n}}\right)^{-1/2} a\right]^{1/2} \\
& \text{(by CS inequality)}
\end{aligned}
$$

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

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

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

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

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

In this case, we apply part (d) of Lemma S2-8 with $u_{(j,s)} = \left(\mu_n^{\min}\right) \underline{u}_{(j,s),n}$, $u_{(k,v)} = \left(\mu_n^{\min}\right) \underline{u}_{(k,v),n}$, 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 (36) 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

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

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 $\left(\mu_n^{\min}/\sqrt{K_{2,n}}\right) \left(\widetilde{L}_n \Lambda_{II,n} \widetilde{L}_n\right)^{-1/2} \widetilde{L}_n D_\mu \left(\widehat{\delta}_{L,n} - \delta_0\right) \xrightarrow{d} N(0, I_d)$.

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

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

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

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

noting that, in this case, $\left(\mu_n^{\min}\right) / \sqrt{K_{2,n}} = o(1)$ 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 $\left(\mu_n^{\min} / \sqrt{K_{2,n}}\right) \left(\widetilde{L}_n \Lambda_{II,n} \widetilde{L}_n\right)^{-1/2} \widetilde{L}_n D_\mu \left(\widehat{\delta}_{F,n} - \delta_0\right) \xrightarrow{d} N(0, I_d)$. $\square$

Appendix S2: Key Lemmas Used in Proving the Main Theorems

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) $\sum_{(i,t),(j,s)=1,(i,t)\neq(j,s)}^{m_n} A_{(i,t),(j,s)}^2 = O_{a.s.}(K_{2,n})$.
(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) $\max_{1\leq(i,t)\leq m_n} \left(\sum_{(j,s)=1}^{m_n} A_{(i,t),(j,s)}^2\right) = O_{a.s.}(K_{2,n}/n)$.
Proof of Lemma S2-1:

To show part (a), note first that, given Lemma 1, there exists a constant $C > 0$ such that

$$\begin{aligned}
tr\left\{D_{\widehat{\vartheta}}^2\right\} &= d_{P^\perp}' \left(M^{(Z,Q)} \circ M^{(Z,Q)}\right)^{-2} d_{P^\perp} \leq \frac{d_{P^\perp}' d_{P^\perp}}{\left[\lambda_{\min}\left(M^{(Z,Q)} \circ M^{(Z,Q)}\right)\right]^2} \\
&\leq \left(\frac{1}{C}\right)^2 d_{P^\perp}' d_{P^\perp} \quad a.s. \text{ (by Lemma 1)}
\end{aligned}$$

$$
\begin{aligned}
&\leq \left(\frac{1}{C}\right)^2 \left(\max_{1\leq(i,t)\leq m_n} P^{\perp}_{(i,t),(i,t)}\right) \sum_{(i,t)=1}^{m_n} P^{\perp}_{(i,t),(i,t)} \\
&= O_{a.s.}\left(\frac{K_{2,n}^2}{n}\right) \quad \text{(by Assumption 5(iv))}.
\end{aligned}
$$

Now, by straightforward calculations and by making use of the inequality $\left|\sum_{i=1}^{G} a_i\right|^r \leq G^{r-1}\sum_{i=1}^{G} |a_i|^r$, we get

$$
\begin{aligned}
\sum_{\substack{(i,t),(j,s)=1 \\ (i,t)\neq(j,s)}}^{m_n} A^2_{(i,t),(j,s)} &= \sum_{\substack{(i,t),(j,s)=1 \\ (i,t)\neq(j,s)}}^{m_n} \left(P^{\perp}_{(i,t),(j,s)} - e'_{(i,t)} M^{(Z,Q)} D_{\widehat{\vartheta}} M^{(Z,Q)} e_{(j,s)}\right)^2 \\
&\leq 2 \sum_{\substack{(i,t),(j,s)=1 \\ (i,t)\neq(j,s)}}^{m_n} \left(P^{\perp}_{(i,t),(j,s)}\right)^2 + 2 \sum_{\substack{(i,t),(j,s)=1 \\ (i,t)\neq(j,s)}}^{m_n} \left(e'_{(i,t)} M^{(Z,Q)} D_{\widehat{\vartheta}} M^{(Z,Q)} e_{(j,s)}\right)^2 \\
&\leq 2\left[K_{2,n} + \sum_{(i,t)=1}^{m_n} e'_{(i,t)} M^{(Z,Q)} D_{\widehat{\vartheta}} M^{(Z,Q)} D_{\widehat{\vartheta}} M^{(Z,Q)} e_{(i,t)}\right] \\
&\leq 2\left[K_{2,n} + tr\left\{M^{(Z,Q)} D^2_{\widehat{\vartheta}} M^{(Z,Q)}\right\}\right] = 2\left[K_{2,n} + tr\left\{D_{\widehat{\vartheta}} M^{(Z,Q)} D_{\widehat{\vartheta}}\right\}\right] \\
&\leq 2\left[K_{2,n} + tr\left\{D^2_{\widehat{\vartheta}}\right\}\right] = O_{a.s.}\left(K_{2,n}\right) + O_{a.s.}\left(\frac{K_{2,n}^2}{n}\right) = O_{a.s.}\left(K_{2,n}\right).
\end{aligned}
$$

Parts (b)-(d) can be shown by using arguments that are analogous to that given to show part (a) above. Hence, for the sake of brevity, we will not provide an explicit proof for these parts here, but a proof can be obtained from the authors upon request. □

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

Proof of Lemma S2-2:

To show part (a), note first that $D_\mu^{-1} X' M^{(Z_1,Q)} X D_\mu^{-1}$ $\leq 3\left[\Gamma' M^{(Z_1,Q)} \Gamma/n + D_\mu^{-1} D_\kappa F' M^{(Z_1,Q)} F D_\kappa D_\mu^{-1}/n + D_\mu^{-1} U' M^{(Z_1,Q)} U D_\mu^{-1}\right]$, where $F = \left(f\left(W_{1,(1,1)}\right), .., f\left(W_{1,(1,T_1)}\right), ..., f\left(W_{1,(n,1)}\right), .., f\left(W_{1,(n,T_n)}\right)\right)'$ and where we take $A \leq B$ for two square matrices A and B to mean that $A - B$ is negative semi-definite, or, alternatively, $B - A$ is positive semidefinite. Now, for $a, b \in \mathbb{R}^d$ such that $\|a\|_2 = \|b\|_2 = 1$, we can apply the CS inequality and Assumption 3(iii) to obtain $\left|a'\Gamma' M^{(Z_1,Q)} \Gamma b/n\right| \leq \sqrt{a'\Gamma' M^{(Z_1,Q)} \Gamma a/n}\sqrt{b'\Gamma' M^{(Z_1,Q)} \Gamma b/n} \leq \sqrt{a'\Gamma'\Gamma a/n}\sqrt{b'\Gamma'\Gamma b/n} = O_{a.s.}(1)$. Since the above argument holds for all a, $b \in \mathbb{R}^d$ such that $\|a\|_2 = \|b\|_2 = 1$, we further deduce that

$\Gamma' M^{(Z_1,Q)} \Gamma / n = O_p(1)$. Next, note that, for $a, b \in \mathbb{R}^d$ such that $\|a\|_2 = \|b\|_2 = 1$, we can apply the CS inequality along with Assumptions 3(ii), 4(i), 5(i), and 7(ii) to obtain

$$\begin{aligned}
\left| \frac{a' D_\mu^{-1} D_\kappa F' M^{(Z_1,Q)} F D_\kappa D_\mu^{-1} b}{n} \right| &\leq \sqrt{\frac{a' D_\mu^{-1} D_\kappa \left(F - Z_1 \Theta^{K_{1,n}}\right)' M^{(Z_1,Q)} \left(F - Z_1 \Theta^{K_{1,n}}\right) D_\kappa D_\mu^{-1} a}{n}} \\
&\quad \times \sqrt{\frac{b' D_\mu^{-1} D_\kappa \left(F - Z_1 \Theta^{K_{1,n}}\right)' M^{(Z_1,Q)} \left(F - Z_1 \Theta^{K_{1,n}}\right) D_\kappa D_\mu^{-1} b}{n}} \\
&\leq \frac{m_n d}{n} \left\| f(\cdot) - \Theta^{K_{1,n}\prime} Z_1(\cdot) \right\|_{\infty,d}^2 \frac{\left(\kappa_n^{\max}\right)^2}{\left(\mu_n^{\min}\right)^2} = O_{a.s.}\left(\frac{\left(\kappa_n^{\max}\right)^2}{\left(\mu_n^{\min}\right)^2 K_{1,n}^{2\varrho_f}} \right)
\end{aligned}$$

Since the above argument holds for all a, $b \in \mathbb{R}^d$ such that $\|a\|_2 = \|b\|_2 = 1$, we further deduce that $D_\mu^{-1} D_\kappa F' M^{(Z_1,Q)} F D_\kappa D_\mu^{-1} / n = O_{a.s.}\left(\left(\kappa_n^{\max}\right)^2 \left(\mu_n^{\min}\right)^{-2} K_{1,n}^{-2\varrho_f} \right)$. Finally, note that, by the CS inequality, $\left| a' D_\mu^{-1} U' M^{(Z_1,Q)} U D_\mu^{-1} b \right| \leq \sqrt{a' D_\mu^{-1} U' M^{(Z_1,Q)} U D_\mu^{-1} a} \sqrt{b' D_\mu^{-1} U' M^{(Z_1,Q)} U D_\mu^{-1} b}$. Now, by Assumptions 2(i), 3(ii), and 6(ii),

$$\begin{aligned}
E\left[a' D_\mu^{-1} U' M^{(Z_1,Q)} U D_\mu^{-1} a | \mathcal{F}_n^W \right] &\leq a' D_\mu^{-1} E\left[U'U | \mathcal{F}_n^W \right] D_\mu^{-1} a \\
&\leq \frac{\overline{T} \left(\max_{1 \leq (i,t) \leq m_n} E\left[\left\| U_{(i,t)} \right\|_2^2 | \mathcal{F}_n^W \right] \right) n}{\left(\mu_n^{\min}\right)^2} = O_{a.s.}\left(\frac{n}{\left(\mu_n^{\min}\right)^2} \right),
\end{aligned}$$

where $\overline{T} = \max_{1 \leq (i,t) \leq m_n} T_i$. Hence, by Theorem 16.1 of Billingsley (1995), there exists a constant $\overline{C} < \infty$ such that for all n sufficiently large $E\left[\left(\left(\mu_n^{\min}\right)^2 / n \right) a' D_\mu^{-1} U' M^{(Z_1,Q)} U D_\mu^{-1} a \right] = E_{W_n}\left(\left(\left(\mu_n^{\min}\right)^2 / n \right) E\left[\frac{a' U' M^{(Z_1,Q)} U a}{n} | \mathcal{F}_n^W \right] \right) \leq \overline{C}$. It follows from the Markov's inequality that $a' D_\mu^{-1} U' M^{(Z_1,Q)} U D_\mu^{-1} a = O_p\left(n \left(\mu_n^{\min}\right)^{-2} \right)$. In the same way, we also have $b' D_\mu^{-1} U' M^{(Z_1,Q)} U D_\mu^{-1} b = O_p\left(n / \left(\mu_n^{\min}\right)^2 \right)$, so that $\left| a' D_\mu^{-1} U' M^{(Z_1,Q)} U D_\mu^{-1} b \right| \leq \sqrt{a' D_\mu^{-1} U' M^{(Z_1,Q)} U D_\mu^{-1} a} \sqrt{b' D_\mu^{-1} U' M^{(Z_1,Q)} U D_\mu^{-1} b} = O_p\left(n \left(\mu_n^{\min}\right)^{-2} \right)$. Since this result holds for all a, $b \in \mathbb{R}^d$ such that $\|a\|_2 = \|b\|_2 = 1$, we further deduce that $D_\mu^{-1} U' M^{(Z_1,Q)} U D_\mu^{-1} = O_p\left(n \left(\mu_n^{\min}\right)^{-2} \right)$. Putting these results together, it follows that $D_\mu^{-1} X' M^{(Z_1,Q)} X D_\mu^{-1} = O_p(1) + O_p\left(\left(\kappa_n^{\max}\right)^2 \left(\mu_n^{\min}\right)^{-2} K_{1,n}^{-2\varrho_f} \right) + O_p\left(n \left(\mu_n^{\min}\right)^{-2} \right) = O_p\left(n \left(\mu_n^{\min}\right)^{-2} \right)$, as required to show part (a).

Part (b) of this lemma can be proved by generalizing the argument for Lemma A2 of Hausman et al (2012) to our setting here with cluster sampling, fixed effects, and possibly many covariates. For the sake of brevity, we will not include an explicit proof here, but the details of a proof can be obtained from the authors upon request. □

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 $\mathcal{Y}_n = \Gamma' M^{(Z_1,Q)}\varepsilon/\sqrt{n} + D_\mu^{-1}\underline{U}' A\varepsilon$.

Proof of Lemma S2-3:

For part (a), note that, by Assumptions 1, 2, and 3(iii); there exists a pair of constants $0 < \underline{C} \leq \overline{C} < \infty$ such that, for any $b \in \mathbb{R}^d$ such that $\|b\| = 1$ and for all n sufficiently large, $b'VC\left(\Gamma' M^{(Z_1,Q)}\varepsilon/\sqrt{n}|\mathcal{F}_n^W\right)b = b'\Gamma' M^{(Z_1,Q)}E\left[\varepsilon\varepsilon'|\mathcal{F}_n^W\right]M^{(Z_1,Q)}\Gamma b/n \leq \left(\max_{1\leq(i,t)\leq m_n} E\left[\varepsilon_{(i,t)}^2|\mathcal{F}_n^W\right]\right)\lambda_{\max}\left(\Gamma'\Gamma/n\right) \leq \overline{C}$ $a.s.$ and $b'VC\left(\Gamma' M^{(Z_1,Q)}\varepsilon/\sqrt{n}|\mathcal{F}_n^W\right)b \geq \left(\min_{1\leq(i,t)\leq m_n} E\left[\varepsilon_{(i,t)}^2|\mathcal{F}_n^W\right]\right)b'\Gamma' M^{(Z_1,Q)}\Gamma b/n \geq \underline{C}$ $a.s.$ Since the above bounds hold for any $b \in \mathbb{R}^d$ such that $\|b\|_2 = 1$, it follows that, almost surely, $\lambda_{\max}\left[VC\left(\Gamma' M^{(Z_1,Q)}\varepsilon/\sqrt{n}|\mathcal{F}_n^W\right)\right] = \max_{\|b\|=1} b'VC\left(\Gamma' M^{(Z_1,Q)}\varepsilon/\sqrt{n}|\mathcal{F}_n^W\right)b \leq \overline{C} < \infty$ and $\lambda_{\min}\left[VC\left(\Gamma' M^{(Z_1,Q)}\varepsilon/\sqrt{n}|\mathcal{F}_n^W\right)\right] = \min_{\|b\|=1} b'VC\left(\Gamma' M^{(Z_1,Q)}\varepsilon/\sqrt{n}|\mathcal{F}_n^W\right)b \geq \underline{C} > 0$ for all n sufficiently large, which establishes the required result.

To show part (b), let $a \in \mathbb{R}^d$ such that $\|a\|_2 = 1$, and we define $\underline{U}_{(i,t)} = U_{(i,t)} - \rho\varepsilon_{(i,t)}$, $\underline{u}_{a,(i,t)} = a'\underline{U}_{(i,t)}$, $\sigma_{(i,t)}^2 = E\left[\varepsilon_{(i,t)}^2|\mathcal{F}_n^W\right]$, $\widetilde{\omega}_{a,(i,t)}^2 = E\left[\underline{u}_{a,(i,t)}^2|\mathcal{F}_n^W\right]$, $\widetilde{\psi}_{a,(i,t)} = E\left[\varepsilon_{(i,t)}\underline{u}_{a,(i,t)}|\mathcal{F}_n^W\right]$, and $\varrho_{a,(i,t)} = \widetilde{\psi}_{a,(i,t)}/\left(\sigma_{(i,t)}\widetilde{\omega}_{a,(i,t)}\right)$; for $(i,t) = 1,..,m_n$, where we have suppressed the dependence of $\sigma_{(i,t)}^2$, $\widetilde{\omega}_{a,(i,t)}^2$, $\widetilde{\psi}_{a,(i,t)}$, and $\varrho_{a,(i,t)}$ on $\mathcal{F}_n^W = \sigma(W_n)$ for notational convenience. Note also that we can write $a'VC\left(\underline{U}' A\varepsilon/\sqrt{K_{2,n}}|\mathcal{F}_n^W\right)a =$
$K_{2,n}^{-1}\sum_{(i,t)=2}^{m_n}\sum_{(j,s)=1}^{(i,t)-1} A_{(i,t),(j,s)}^2 E\left[\left(\varepsilon_{(j,s)}\underline{u}_{a,(i,t)} + \varepsilon_{(i,t)}\underline{u}_{a,(j,s)}\right)^2|\mathcal{F}_n^W\right]$, since, by construction, $A_{(i,t),(i,t)} = 0$ for $(i,t) = 1,...,m_n$. Moreover, define $\delta_{a,(i,t),(j,s)} = \begin{pmatrix} \sigma_{(j,s)}\widetilde{\omega}_{a,(i,t)} & \sigma_{(i,t)}\widetilde{\omega}_{a,(j,s)} \end{pmatrix}'$ and

$$\Delta_{(i,t),(j,s)}^a = \begin{pmatrix} 1 & \varrho_{a,(i,t)}\varrho_{a,(j,s)} \\ \varrho_{a,(i,t)}\varrho_{a,(j,s)} & 1 \end{pmatrix}$$

and note that, given that $(j,s) < (i,t)$, we have $E\left[\left(\varepsilon_{(j,s)}\underline{u}_{a,(i,t)} + \varepsilon_{(i,t)}\underline{u}_{a,(j,s)}\right)^2|\mathcal{F}_n^W\right] = \delta_{a,(i,t),(j,s)}'\Delta_{(i,t),(j,s)}^a\delta_{a,(i,t),(j,s)}$. Now, by the quadratic formula, the smallest eigenvalue of

$\Delta^a_{(i,t),(j,s)}$ is given by $\lambda_{\min}\left(\Delta^a_{(i,t),(j,s)}\right) = 1 - \left|\varrho_{a,(i,t)}\right| \left|\varrho_{a,(j,s)}\right|$. In addition, write

$$\begin{aligned}
\widetilde{\Omega}_{(i,t)} &= \begin{pmatrix} E\left[\varepsilon^2_{(i,t)}|\mathcal{F}^W_n\right] & E\left[\varepsilon_{(i,t)}\underline{U}'_{(i,t)}|\mathcal{F}^W_n\right] \\ E\left[\underline{U}_{(i,t)}\varepsilon_{(i,t)}|\mathcal{F}^W_n\right] & E\left[\underline{U}_{(i,t)}\underline{U}'_{(i,t)}|\mathcal{F}^W_n\right] \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ -\rho & I_d \end{pmatrix} \begin{pmatrix} E\left[\varepsilon^2_{(i,t)}|\mathcal{F}^W_n\right] & E\left[\varepsilon_{(i,t)}U'_{(i,t)}|\mathcal{F}^W_n\right] \\ E\left[U_{(i,t)}\varepsilon_{(i,t)}|\mathcal{F}^W_n\right] & E\left[U_{(i,t)}U'_{(i,t)}|\mathcal{F}^W_n\right] \end{pmatrix} \begin{pmatrix} 1 & -\rho' \\ 0 & I_d \end{pmatrix} \\
&= L_\rho \Omega_{(i,t)} L'_\rho, \text{ where } L_\rho = \begin{pmatrix} 1 & 0 \\ -\rho & I_d \end{pmatrix}.
\end{aligned}$$

Note that L_ρ is nonsingular, so that $L_\rho L'_\rho$ is positive definite. Hence, by Assumption 2 part (ii) and by the fact that $L_\rho L'_\rho$ is a fixed, finite-dimensional positive definite matrix, there exists some constant $C_1 > 1$ such that

$$\min_{1\leq(i,t)\leq m_n} \lambda_{\min}\left(\widetilde{\Omega}_{(i,t)}\right) \geq \min_{1\leq(i,t)\leq m_n} \lambda_{\min}\left(\Omega_{(i,t)}\right) \lambda_{\min}\left(L_\rho L'_\rho\right) \geq 1/C_1 > 0 \ a.s.n. \tag{18}$$

Next, let

$$\underset{(d+1)\times 2}{D_a} = \begin{pmatrix} 1 & 0 \\ 0 & a \end{pmatrix}, \ D_{SD,(i,t)} = \begin{pmatrix} \sigma_{(i,t)} & 0 \\ 0 & \widetilde{\omega}_{a,(i,t)} \end{pmatrix}, \text{ and } D_{\varrho,(i,t)} = \begin{pmatrix} 1 & \varrho_{a,(i,t)} \\ \varrho_{a,(i,t)} & 1 \end{pmatrix}$$

and note that

$$\begin{aligned}
D'_a \widetilde{\Omega}_{(i,t)} D_a &= \begin{pmatrix} 1 & 0 \\ 0 & a' \end{pmatrix} \begin{pmatrix} E\left[\varepsilon^2_{(i,t)}|\mathcal{F}^W_n\right] & E\left[\varepsilon_{(i,t)}\underline{U}'_{(i,t)}|\mathcal{F}^W_n\right] \\ E\left[\underline{U}_{(i,t)}\varepsilon_{(i,t)}|\mathcal{F}^W_n\right] & E\left[\underline{U}_{(i,t)}\underline{U}'_{(i,t)}|\mathcal{F}^W_n\right] \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & a \end{pmatrix} \\
&= \begin{pmatrix} \sigma_{(i,t)} & 0 \\ 0 & \widetilde{\omega}_{a,(i,t)} \end{pmatrix} \begin{pmatrix} 1 & \varrho_{a,(i,t)} \\ \varrho_{a,(i,t)} & 1 \end{pmatrix} \begin{pmatrix} \sigma_{(i,t)} & 0 \\ 0 & \widetilde{\omega}_{a,(i,t)} \end{pmatrix} = D_{SD,(i,t)} D_{\varrho,(i,t)} D_{SD,(i,t)},
\end{aligned}$$

Now, as can be seen from expression (18) above, an implication of Assumption 2(ii) is that

$$\min_{1\leq(i,t)\leq m_n} \sigma^2_{(i,t)} = e'_{1,d+1}\widetilde{\Omega}_{(i,t)}e_{1,d+1} \geq 1/C_1 > 0 \ a.s.n. \tag{19}$$

$$\min_{1\leq(i,t)\leq m_n} \widetilde{\omega}^2_{a,(i,t)} = \underline{a}'\widetilde{\Omega}_{(i,t)}\underline{a} \geq 1/C_1 > 0 \ a.s.n. \tag{20}$$

where $\underset{(d+1)\times 1}{e_{1,d+1}} = \begin{pmatrix} 1 & 0' \end{pmatrix}'$ and $\underset{(d+1)\times 1}{\underline{a}} = \begin{pmatrix} 0 & a' \end{pmatrix}'$, from which we deduce that $D_{SD,(i,t)}$ is invertible almost surely for each $(i,t) \in \{1,...,m_n\}$ and for all n sufficiently large. The invertibility of $D_{SD,(i,t)}$ then allows us to write $D_{\varrho,(i,t)} = D^{-1}_{SD,(i,t)} D'_a \widetilde{\Omega}_{(i,t)} D_a D^{-1}_{SD,(i,t)}$. On the other hand, Assumption 2(i) implies that there exists some constant $C_2 > 1$ such that

$$\min_{1\leq(i,t)\leq m_n} \lambda_{\min}\left(D_{SD,(i,t)}^{-1}\right) = \frac{1}{\max_{1\leq(i,t)\leq m_n}\lambda_{\max}\left(D_{SD,(i,t)}\right)} \geq \frac{1}{C_2} > 0 \quad a.s. \tag{21}$$

It follows from the fact that $\lambda_{\min}\left(D_a' D_a\right) = \lambda_{\min}\left(I_2\right) = 1$ and from making use of Assumptions 2(i) and (ii) and the lower bounds given in (18) and (21) that

$$\begin{aligned}
\min_{1\leq(i,t)\leq m_n} \lambda_{\min}\left(D_{\varrho,(i,t)}\right) &\geq \min_{1\leq(i,t)\leq m_n} \lambda_{\min}\left(\widetilde{\Omega}_{(i,t)}\right)\lambda_{\min}\left(D_a' D_a\right)\lambda_{\min}\left(D_{SD,(i,t)}^{-2}\right) \\
&\geq \frac{\min_{1\leq(i,t)\leq m_n}\lambda_{\min}\left(\widetilde{\Omega}_{(i,t)}\right)}{\max_{1\leq(i,t)\leq m_n}\left(\lambda_{\max}\left(D_{SD,(i,t)}\right)\right)^2} \geq \frac{1}{C^3} > 0 \quad a.s.n.,
\end{aligned}$$

where $C = \max\{C_1, C_2\}$. Moreover, by solving the characteristic equation of $D_{\varrho,(i,t)}$, we see that the smallest eigenvalue of $D_{\varrho,(i,t)}$ is given by $\lambda_{\min}\left(D_{\varrho,(i,t)}\right) = 1 - \left|\varrho_{a,(i,t)}\right|$, so that $\min_{1\leq(i,t)\leq m_n}\lambda_{\min}\left(D_{\varrho,(i,t)}\right) = 1 - \max_{1\leq(i,t)\leq m_n}\left|\rho_{a,(i,t)}\right| \geq 1/C^3 > 0$ $a.s.n.$, from which we further deduce that $\max_{1\leq(i,t)\leq m_n}\left|\varrho_{a,(i,t)}\right| \leq 1-(1/C^3) < 1$ $a.s.n.$ Applying this upper bound along with the lower bounds given by (19) and (20) as well as the fact that $\lambda_{\min}\left(\Delta^a_{(i,t),(j,s)}\right) = 1 - \left|\varrho_{a,(i,t)}\right|\left|\varrho_{a,(j,s)}\right|$, as derived earlier, we have

$$\begin{aligned}
E\left[\left(\varepsilon_{(j,s)}\underline{u}_{a,(i,t)} + \varepsilon_{(i,t)}\underline{u}_{a,(j,s)}\right)^2 |\mathcal{F}_n^W\right] &= \delta'_{a,(i,t),(j,s)}\Delta^a_{(i,t),(j,s)}\delta_{a,(i,t),(j,s)} \\
&\geq \left[1 - \left|\varrho_{a,(i,t)}\right|\left|\varrho_{a,(j,s)}\right|\right]\left[\sigma^2_{(j,s)}\widetilde{\omega}^2_{a,(i,t)} + \sigma^2_{(i,t)}\widetilde{\omega}^2_{a,(j,s)}\right] \\
&\geq \left(\frac{2}{C^3} - \frac{1}{C^6}\right)\left[\frac{1}{C^2} + \frac{1}{C^2}\right] \geq \frac{2}{C^5} > 0 \quad a.s.n.
\end{aligned}$$

Summing over $1 \leq (j,s) < (i,t) \leq m_n$, we obtain $K_{2,n}^{-1}\sum_{(i,t)=2}^{m_n}\sum_{(j,s)=1}^{(i,t)-1} A^2_{(i,t),(j,s)} E\left[\left(\varepsilon_{(j,s)}\underline{u}_{a,(i,t)} + \varepsilon_{(i,t)}\underline{u}_{a,(j,s)}\right)^2 |\mathcal{F}_n^W\right]$ $\geq (2/C^5)\, K_{2,n}^{-1}\sum_{(i,t)=2}^{m_n}\sum_{(j,s)=1}^{(i,t)-1} A^2_{(i,t),(j,s)} = (1/C^5)\, K_{2,n}^{-1}\sum_{(i,t)=1}^{m_n}\sum_{(j,s)=1}^{m_n} A^2_{(i,t),(j,s)}$, where the last equality follows from the symmetry of A and by the fact that $A_{(i,t),(i,t)} = 0$ for $(i,t) = 1, ..., m_n$. Furthermore, by straightforward calculation, we obtain

$$\begin{aligned}
\frac{1}{K_{2,n}}\sum_{(i,t)=1}^{m_n}\sum_{(j,s)=1}^{m_n} A^2_{(i,t),(j,s)} &= \frac{1}{K_{2,n}}\sum_{(i,t)=1}^{m_n}\sum_{(j,s)=1}^{m_n}\left[e'_{(i,t)}P^{\perp}e_{(j,s)} - e'_{(i,t)}M^{(Z,Q)}D_{\widehat{\vartheta}}M^{(Z,Q)}e_{(j,s)}\right]^2 \\
&= 1 + \frac{1}{K_{2,n}}\sum_{(i,t)=1}^{m_n} e'_{(i,t)}M^{(Z,Q)}D_{\widehat{\vartheta}}M^{(Z,Q)}D_{\widehat{\vartheta}}M^{(Z,Q)}e_{(i,t)} \geq 1
\end{aligned}$$

Putting everything together, we have

$K_{2,n}^{-1}\sum_{(i,t)=2}^{m_n}\sum_{(j,s)=1}^{(i,t)-1}A_{(i,t),(j,s)}^2 E\left[\left(\varepsilon_{(j,s)}\underline{u}_{a,(i,t)}+\varepsilon_{(i,t)}\underline{u}_{a,(j,s)}\right)^2|\mathcal{F}_n^W\right]\geq$ $\left(1/C^5\right)K_{2,n}^{-1}\sum_{(i,t)=1}^{m_n}\sum_{(j,s)=1}^{m_n}A_{(i,t),(j,s)}^2\geq 1/C^5\geq\underline{C}>0$, by choosing $\underline{C}$ such that $0<\underline{C}\leq 1/C^5$. Since the above argument holds for any $a\in\mathbb{R}^d$ such that $\|a\|=1$, it further follows that $VC\left(\underline{U}'A\varepsilon/\sqrt{K_{2,n}}|\mathcal{F}_n^W\right)\geq\underline{C}I_d>0\ \ a.s.$, as required.

To show part (c), note first that, given Assumption 2(i), there exists a positive constant C such that

$$
\begin{aligned}
&\max_{1\leq(j,s)\leq m_n}\lambda_{\max}\left(E\left[\underline{U}_{(j,s)}\underline{U}_{(j,s)}'|\mathcal{F}_n^W\right]\right)\\
\leq\ &\max_{1\leq(j,s)\leq m_n} tr\left\{E\left[\underline{U}_{(j,s)}\underline{U}_{(j,s)}'|\mathcal{F}_n^W\right]\right\}\\
\leq\ &\max_{1\leq(j,s)\leq m_n}\left\{E\left[\left\|U_{(j,s)}\right\|_2^2|\mathcal{F}_n^W\right]+2E\left[\left|U_{(j,s)}'\rho\varepsilon_{(j,s)}\right||\mathcal{F}_n^W\right]+\rho'\rho E\left[\varepsilon_{(j,s)}^2|\mathcal{F}_n^W\right]\right\}\\
\leq\ &\max_{1\leq(j,s)\leq m_n}\Bigg\{E\left[\left\|U_{(j,s)}\right\|_2^2|\mathcal{F}_n^W\right]+2\left\|\rho\right\|_2\sqrt{E\left[\varepsilon_{(j,s)}^2|\mathcal{F}_n^W\right]}\sqrt{E\left[\left\|U_{(j,s)}\right\|_2^2|\mathcal{F}_n^W\right]}\\
&\qquad+\left\|\rho\right\|_2^2E\left[\varepsilon_{(j,s)}^2|\mathcal{F}_n^W\right]\Bigg\}\\
\leq\ &2\left\{\left(\max_{1\leq(j,s)\leq m_n}E\left[\left\|U_{(j,s)}\right\|_2^2|\mathcal{F}_n^W\right]\right)+\left\|\rho\right\|_2^2\left(\max_{1\leq(j,s)\leq m_n}E\left[\varepsilon_{(j,s)}^2|\mathcal{F}_n^W\right]\right)\right\}\\
\leq\ &C<\infty\ \ a.s. \qquad (22)
\end{aligned}
$$

where the third inequality above follows from applying the CS inequality while the fourth inequality stems in part from applying the inequality $|XY|\leq(1/2)X^2+(1/2)Y^2$. Now, for any $a\in\mathbb{R}^d$ such that $\|a\|_2=1$; we obtain by applying the triangle and CS inequalities, expression (22), as well as part (a) of Lemma S2-1 and Assumptions 2(i) and 8

$$
\begin{aligned}
&a'VC\left(\underline{U}'A\varepsilon/\sqrt{K_{2,n}}|\mathcal{F}_n^W\right)a\\
\leq\ &\frac{1}{K_{2,n}}\sum_{\substack{(i,t),(j,s)=1\\(i,t)\neq(j,s)}}^{m_n}A_{(i,t),(j,s)}^2\Bigg(E\left[\varepsilon_{(i,t)}^2|\mathcal{F}_n^W\right]a'E\left[\underline{U}_{(j,s)}\underline{U}_{(j,s)}'|\mathcal{F}_n^W\right]a\\
&+\sqrt{a'E\left[\underline{U}_{(i,t)}\underline{U}_{(i,t)}'|\mathcal{F}_n^W\right]a}\sqrt{E\left[\varepsilon_{(i,t)}^2|\mathcal{F}_n^W\right]}\sqrt{a'E\left[\underline{U}_{(j,s)}\underline{U}_{(j,s)}'|\mathcal{F}_n^W\right]a}\sqrt{E\left[\varepsilon_{(j,s)}^2|\mathcal{F}_n^W\right]}\Bigg)\\
\leq\ &\frac{1}{K_{2,n}}\sum_{\substack{(i,t),(j,s)=1\\(i,t)\neq(j,s)}}^{m_n}A_{(i,t),(j,s)}^2\Bigg\{E\left[\varepsilon_{(i,t)}^2|\mathcal{F}_n^W\right]\lambda_{\max}\left(E\left[\underline{U}_{(j,s)}\underline{U}_{(j,s)}'|\mathcal{F}_n^W\right]\right)\\
&+\sqrt{\lambda_{\max}\left(E\left[\underline{U}_{(j,s)}\underline{U}_{(j,s)}'|\mathcal{F}_n^W\right]\right)}\sqrt{E\left[\varepsilon_{(i,t)}^2|\mathcal{F}_n^W\right]}\sqrt{\lambda_{\max}\left(E\left[\underline{U}_{(j,s)}\underline{U}_{(j,s)}'|\mathcal{F}_n^W\right]\right)}\sqrt{E\left[\varepsilon_{(j,s)}^2|\mathcal{F}_n^W\right]}\Bigg\}\\
=\ &O_{a.s.}(1).
\end{aligned}
$$

From this, we deduce that $\lambda_{\max}\left[VC\left(\underline{U}'A\varepsilon/\sqrt{K_{2,n}}|\mathcal{F}_n^W\right)\right] = \max_{\|a\|=1} a'VC\left(\underline{U}'A\varepsilon/\sqrt{K_{2,n}}|\mathcal{F}_n^W\right)a \leq \overline{C}$ *a.s.n.* Moreover, by applying the law of iterated expectations and part (i) of Theorem 16.1 of Billingsley (1995) that there exists a constant $\overline{C} > 0$ such that $a'VC\left(\underline{U}'A\varepsilon/\sqrt{K_{2,n}}\right)a = E_{W_n}\left\{E\left[\left(a'\underline{U}'A\varepsilon\right)^2/K_{2,n}|\mathcal{F}_n^W\right]\right\} \leq \overline{C}$, for all $a \in \mathbb{R}^d$ such that $\|a\| = 1$, from which we further deduce the unconditional version of this inequality, i.e., $\lambda_{\max}\left(VC\left[\underline{U}'A\varepsilon/\sqrt{K_{2,n}}\right]\right) = \max_{\|a\|=1} a'VC\left(\underline{U}'A\varepsilon/\sqrt{K_{2,n}}\right)a \leq \overline{C} < \infty$, where $\underline{U} = U - \varepsilon\rho'$ and $\underline{U}_{(i,t)} = U_{(i,t)} - \rho\varepsilon_{(i,t)}$.

Furthermore, since $\underline{U} = U - \varepsilon\rho'$, we see that, by setting $\rho = 0$ in the argument given above, we can also show that there exists a constant $\overline{C}$ such that $\lambda_{\max}\left(VC\left[U'A\varepsilon/\sqrt{K_{2,n}}|\mathcal{F}_n^W\right]\right) \leq \overline{C} < \infty$ *a.s.n.* and $\lambda_{\max}\left(VC\left[U'A\varepsilon/\sqrt{K_{2,n}}\right]\right) \leq \overline{C} < \infty$ for all n sufficiently large.

Finally, to show part (d), note first that, by straightforward calculations, we get $\Sigma_n = VC\left(\mathcal{Y}_n|\mathcal{F}_n^W\right) = VC\left(\Gamma'M^{(Z_1,Q)}\varepsilon/\sqrt{n}|\mathcal{F}_n^W\right) + VC\left(D_\mu^{-1}\underline{U}'A\varepsilon|\mathcal{F}_n^W\right) = \Sigma_{1,n} + \Sigma_{2,n}$. It follows by part (a) of this lemma that there exists a positive constant $\underline{C}$ such that $\lambda_{\min}\left(\Sigma_n\right) \geq \lambda_{\min}\left[VC\left(\Gamma'M^{(Z_1,Q)}\varepsilon/\sqrt{n}|\mathcal{F}_n^W\right)\right] + \lambda_{\min}\left[VC\left(D_\mu^{-1}\underline{U}'A\varepsilon|\mathcal{F}_n^W\right)\right] \geq \lambda_{\min}\left[VC\left(\Gamma'M^{(Z_1,Q)}\varepsilon/\sqrt{n}|\mathcal{F}_n^W\right)\right] \geq \underline{C} > 0$ *a.s.n.*, so that Σ_n is positive definite *a.s.n.* Moreover, again by part (a) of this lemma, for any $a \in \mathbb{R}^d$ such that $\|a\| = 1$,

$$\begin{aligned} a'\Sigma_n^{-1}a &\leq \frac{1}{\lambda_{\min}\left\{VC\left(\Gamma'M^{(Z_1,Q)}\varepsilon/\sqrt{n}|\mathcal{F}_n^W\right) + VC\left(D_\mu^{-1}\underline{U}'A\varepsilon|\mathcal{F}_n^W\right)\right\}} \\ &\leq \frac{1}{\lambda_{\min}\left[VC\left(\Gamma'M^{(Z_1,Q)}\varepsilon/\sqrt{n}|\mathcal{F}_n^W\right)\right]} \leq \frac{1}{\underline{C}} \leq \overline{C} < \infty \ a.s.n. \end{aligned}$$

where $\overline{C}$ can be taken to be any finite, positive constant such that $\overline{C} \geq 1/\underline{C}$. □

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) $D_\mu^{-1}X'A\varepsilon = \frac{\Gamma'M^{(Z_1,Q)}\varepsilon}{\sqrt{n}} + D_\mu^{-1}U'A\varepsilon + O_p\left(K_{2,n}^{-\varrho_\gamma}\right) + O_p\left(K_{2,n}^{-\left(\varrho_\gamma-1\right)}n^{-1}\right) + O_p\left(\kappa_n^{\max}/\left(\mu_n^{\min}K_{1,n}^{\varrho_f}\right)\right) = O_p\left(\max\left\{1, \sqrt{K_{2,n}}/\left(\mu_n^{\min}\right)\right\}\right)$

Proof of Lemma S2-4:

To show part (a), first write $D_\mu^{-1}X'A\varphi_n = D_\mu^{-1}\left(\Upsilon_n + \Phi_n + Q\Xi + U\right)'A\varphi_n = D_\mu^{-1}\Upsilon_n'A\varphi_n + D_\mu^{-1}\Phi_n'A\varphi_n + D_\mu^{-1}U'A\varphi_n$, where the second equality above follows from the fact that $Q'A\varphi_n = 0$. We will analyze each term on the right-hand side of the expression above in turn. Consider first the term $D_\mu^{-1}\Upsilon_n'A\varphi_n$. For $b \in \mathbb{R}^d$ such that $\|b\|_2 = 1$, note that, making use of the triangle and CS inequalities and Assumptions 3(iii), 5, 7(i), and 7(iii); we get

$$\begin{aligned} \left|b'D_\mu^{-1}\Upsilon_n'A\varphi_n\right| &\leq \left|b'D_\mu^{-1}\Upsilon_n'P^{\perp}\varphi_n\right| + \left|b'D_\mu^{-1}\Upsilon_n'M^{(Z,Q)}D_{\widehat{\vartheta}}M^{(Z,Q)}\varphi_n\right| \\ &= \frac{\tau_n}{n}\left|b'\Gamma'P^{\perp}\left(g - Z_1\theta_{K_{1,n}}\right)\right| \\ &\quad + \frac{\tau_n}{n}\left|b'\left(\Gamma - Z_2\Pi^{K_{2,n}}\right)'M^{(Z,Q)}D_{\widehat{\vartheta}}M^{(Z,Q)}\left(g - Z_1\theta^{K_{1,n}}\right)\right| \end{aligned}$$

$$
\begin{aligned}
&\leq \tau_n \sqrt{\frac{b'\Gamma'\Gamma b}{n}} \sqrt{\frac{m_n}{n}} \left\| g(\cdot) - \theta^{K_{1,n}\prime} Z_1(\cdot) \right\|_\infty \\
&\quad + \tau_n \sqrt{\max_{1 \leq (i,t) \leq m_n} \left| \widehat{\vartheta}_{(i,t)} \right|^2} \left(\frac{m_n \sqrt{d}}{n} \right) \left\| \gamma(\cdot) - \Pi^{K_{2,n}\prime} Z_2(\cdot) \right\|_{\infty,d} \left\| g(\cdot) - \theta^{K_{1,n}\prime} Z_1(\cdot) \right\|_\infty \\
&= O_{a.s.}\left(\frac{\tau_n}{K_{1,n}^{\varrho_g}} \right) + O_{a.s.}\left(\frac{\tau_n}{n K_{2,n}^{(\varrho_\gamma - 1)} K_{1,n}^{\varrho_g}} \right) = O_{a.s.}\left(\frac{\tau_n}{K_{1,n}^{\varrho_g}} \right)
\end{aligned}
$$

where $g = \left(g\left(W_{1,(1,1)}\right), .., g\left(W_{1,(1,T_1)}\right), ..., g\left(W_{1,(n,1)}\right), .., g\left(W_{1,(n,T_n)}\right) \right)'$. Since the above argument holds for all $b \in \mathbb{R}^d$ such that $\|b\|_2 = 1$, we further deduce that $D_\mu^{-1} \Upsilon_n' A \varphi_n = O_{a.s.}\left(\tau_n K_{1,n}^{-\varrho_g} \right)$. Next, consider $D_\mu^{-1} \Phi_n' A \varphi_n$. Again, let $b \in \mathbb{R}^d$ such that $\|b\|_2 = 1$, note that, by applying the triangle and CS inequalities along with Assumptions 3(ii), 4(i), 5, 7(i), and 7(ii); we obtain

$$
\begin{aligned}
\left| b' D_\mu^{-1} \Phi_n' A \varphi_n \right| &\leq \frac{\tau_n}{n} \left| b' D_\mu^{-1} D_\kappa F' P^\perp g \right| + \frac{\tau_n}{n} \left| b' D_\mu^{-1} D_\kappa F' M^{(Z,Q)} D_{\widehat{\vartheta}} M^{(Z,Q)} g \right| \\
&\leq \frac{\tau_n \left(\kappa_n^{\max}\right)}{\left(\mu_n^{\min}\right)} \frac{m_n \sqrt{d}}{n} \left\| f(\cdot) - \Theta^{K_{1,n}\prime} Z_1(\cdot) \right\|_{\infty,d} \left\| g(\cdot) - \theta^{K_{1,n}\prime} Z_1(\cdot) \right\|_\infty \\
&\quad + \frac{\tau_n \left(\kappa_n^{\max}\right)}{\left(\mu_n^{\min}\right)} \sqrt{\max_{1 \leq (i,t) \leq m_n} \left| \widehat{\vartheta}_{(i,t)} \right|^2} \frac{m_n \sqrt{d}}{n} \left\{ \left\| f(\cdot) - \Theta^{K_{1,n}\prime} Z_1(\cdot) \right\|_{\infty,d} \right. \\
&\qquad \left. \times \left\| g(\cdot) - \theta^{K_{1,n}\prime} Z_1(\cdot) \right\|_\infty \right\} \\
&= O_{a.s.}\left(\frac{\tau_n \left(\kappa_n^{\max}\right)}{\left(\mu_n^{\min}\right) K_{1,n}^{\varrho_f + \varrho_g}} \right) = o_{a.s.}(1),
\end{aligned}
$$

where $F = \left(f\left(W_{1,(1,1)}\right), .., f\left(W_{1,(1,T_1)}\right), ..., f\left(W_{1,(n,1)}\right), .., f\left(W_{1,(n,T_n)}\right) \right)'$ and where g is as defined above. Since the above argument holds for all $b \in \mathbb{R}^d$ such that $\|b\|_2 = 1$, we further deduce that $D_\mu^{-1} \Phi_n' A \varphi_n = O_{a.s.}\left(\tau_n \kappa_n^{\max} \left(\mu_n^{\min}\right)^{-1} K_{1,n}^{-\varrho_f - \varrho_g} \right) = o_{a.s.}(1)$. Now, consider the term $D_\mu^{-1} U' A \varphi_n$. Let $u_b = U D_\mu^{-1} b$, for $b \in \mathbb{R}^d$ such that $\|b\|_2 = 1$. Making use of the conditional serial independence assumption given in Assumption 1, we deduce that $E\left[u_b u_b' | \mathcal{F}_n^W \right] = E\left[U D_\mu^{-1} b b' D_\mu^{-1} U' | \mathcal{F}_n^W \right] \leq \left(\mu_n^{\min}\right)^{-2} \max_{1 \leq (i,t) \leq m_n} E\left[\left\| U_{(i,t)} \right\|_2^2 | \mathcal{F}_n^W \right] I_{m_n}$, where we take $A \leq B$ for two square matrices A and B to mean that $A - B$ is negative semi-definite, or, alternatively, $B - A$ is positive semidefinite. Applying the above inequality and Assumptions 2(i), 3(ii), 4(ii), 5, and 7(i); we obtain

$$
E\left(\left[b' D_\mu^{-1} U' A \varphi_n \right]^2 | \mathcal{F}_n^W \right) = \frac{\tau_n^2}{n} \left(g - Z_1 \theta^{K_{1,n}} \right)' A E\left[u_b u_b' | \mathcal{F}_n^W \right] A \left(g - Z_1 \theta^{K_{1,n}} \right)
$$

$$
\begin{aligned}
&\leq \frac{\tau_n^2}{n}\frac{\max_{1\leq(i,t)\leq m_n} E\left[\left\|U_{(i,t)}\right\|_2^2|\mathcal{F}_n^W\right]\left(g - Z_1\theta^{K_{1,n}}\right)' A^2\left(g - Z_1\theta^{K_{1,n}}\right)}{\left(\mu_n^{\min}\right)^2} \\
&\leq \frac{m_n C\tau_n^2}{n\left(\mu_n^{\min}\right)^2}\left[1 + \max_{1\leq(i,t)\leq m_n}\left|\widehat{\vartheta}_{(i,t)}\right|^2\right]\left\|g(\cdot) - \theta^{K_{1,n}\prime}Z_1(\cdot)\right\|_\infty^2 \\
&= O_{a.s.}\left(\frac{\tau_n^2}{\left(\mu_n^{\min}\right)^2 K_{1,n}^{2\varrho_g}}\right)
\end{aligned}
$$

Hence, there exists a positive constant $\overline{C} < \infty$ such that for all n sufficiently large $E\left(\left(\mu_n^{\min}\right)^2 K_{1,n}^{2\varrho_g}\tau_n^{-2}\left[b'D_\mu^{-1}U'A\varphi_n\right]^2\right) = E_{W_n}\left\{\left(\mu_n^{\min}\right)^2 K_{1,n}^{2\varrho_g}\tau_n^{-2}E\left(\left[b'D_\mu^{-1}U'A\varphi_n\right]^2|\mathcal{F}_n^W\right)\right\}$ $\leq \overline{C}$. It follows by applying Markov's inequality that $b'D_\mu^{-1}U'A\varphi_n = O_p\left(\tau_n\left(\mu_n^{\min}\right)^{-1}K_{1,n}^{-\varrho_g}\right)$. Finally, it follows from these intermediate results that $D_\mu^{-1}X'A\varphi_n = D_\mu^{-1}\Upsilon_n' A\varphi_n + D_\mu^{-1}\Phi_n' A\varphi_n + D_\mu^{-1}U'A\varphi_n = O_p\left(\tau_n K_{1,n}^{-\varrho_g}\right)$, as required to show part (a).

Part (b) of this lemma can be proved by generalizing the argument for part (iv) of Lemma A5 in Chao et al (2012) to our setting here with cluster sampling, fixed effects, and possibly many covariates. For the sake of brevity, we will not include an explicit proof here, but the details of a proof can be obtained from the authors upon request. □

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) $D_\mu^{-1}X'M^{(Z_1,Q)}\varepsilon = O_p\left(n/\mu_n^{\min}\right)$.

Proof of Lemma S2-5: Note first that part (a) can be shown in a way similar to the proof of part (a) of Lemma S2-4 above. Hence, for the sake of brevity, we will not include an explicit proof here, but the details of a proof can be obtained from the authors upon request.

To show part (b), let $b \in \mathbb{R}^d$ such that $\|b\|_2 = 1$ and let $u_b = UD_\mu^{-1}b$, $u_{b,(i,t)} = U_{(i,t)}'D_\mu^{-1}b$, and $F = \left(f\left(W_{1,(1,1)}\right), .., f\left(W_{1,(1,T_1)}\right), ..., f\left(W_{1,(n,1)}\right), .., f\left(W_{1,(n,T_n)}\right)\right)'$. Now, write $b'D_\mu^{-1}X'M^{(Z_1,Q)}\varepsilon = b'\Gamma'M^{(Z_1,Q)}\varepsilon/\sqrt{n} + b'D_\mu^{-1}D_\kappa F'M^{(Z_1,Q)}\varepsilon/\sqrt{n} + u_b'M^{(Z_1,Q)}\varepsilon$. By straightforward calculations and by applying Assumptions 1, 2(i), 3, 4, 5(i), 5(iii), and 7(ii) as well as the CS and Markov's inequalities, we can show that $b'\Gamma'M^{(Z_1,Q)}\varepsilon/\sqrt{n} = O_p(1)$ and $b'D_\mu^{-1}D_\kappa F'M^{(Z_1,Q)}\varepsilon/\sqrt{n} = O_p\left(\kappa_n^{\max}\left(\mu_n^{\min}\right)^{-1}K_{1,n}^{-\varrho_f}\right)$. From the CS inequality, we also obtain $E\left[\left|u_b'M^{(Z_1,Q)}\varepsilon\right||\mathcal{F}_n^W\right] \leq \sqrt{E\left[u_b'M^{(Z_1,Q)}u_b|\mathcal{F}_n^W\right]}\sqrt{E\left[\varepsilon'M^{(Z_1,Q)}\varepsilon|\mathcal{F}_n^W\right]}$. Next, note that, by applying Assumptions 2(i) and 6(ii), we have $E\left[\varepsilon'M^{(Z_1,Q)}\varepsilon|\mathcal{F}_n^W\right] \leq E\left[\varepsilon'\varepsilon|\mathcal{F}_n^W\right] \leq n\overline{T}\left(\max_{1\leq(i,t)\leq m_n} E\left[\varepsilon_{(i,t)}^2|\mathcal{F}_n^W\right]\right) = O_{a.s.}(n)$. Similarly, by applying Assumptions 2(i), 3(ii), and 6(ii); we obtain $E\left[u_b'M^{(Z_1,Q)}u_b|\mathcal{F}_n^W\right] = E\left[b'D_\mu^{-1}U'M^{(Z_1,Q)}UD_\mu^{-1}b|\mathcal{F}_n^W\right]$ $\leq \overline{T}\left(\max_{1\leq(i,t)\leq m_n} E\left[\left\|U_{(i,t)}\right\|_2^2|\mathcal{F}_n^W\right]\right)\left[n/\left(\mu_n^{\min}\right)^2\right] = O_{a.s.}\left(n\left(\mu_n^{\min}\right)^{-2}\right)$. It follows from these results that $E\left[\left|u_b'M^{(Z_1,Q)}\varepsilon\right||\mathcal{F}_n^W\right] = O_{a.s.}\left(n\left(\mu_n^{\min}\right)^{-1}\right)$. Hence, by the law of iterated expectations and Theorem 16.1 of Billingsley (1995), there exists a constant $\overline{C} < \infty$ such that for all n sufficiently large $E\left[\left(\left(\mu_n^{\min}\right)/n\right)\left|u_b'M^{(Z_1,Q)}\varepsilon\right|\right] = E_W\left(\left(\left(\mu_n^{\min}\right)/n\right)E\left[\left|u_b'M^{(Z_1,Q)}\varepsilon\right||\mathcal{F}_n^W\right]\right)$

$\leq \overline{C}$. It follows by applying Markov's inequality that $b'D_{\mu}^{-1}U'M^{(Z_1,Q)}\varepsilon = O_p\left(n\left(\mu_n^{\min}\right)^{-1}\right)$. Putting everything together, we have

$$\begin{aligned} b'D_{\mu}^{-1}X'M^{(Z_1,Q)}\varepsilon &= \frac{b'\Gamma'M^{(Z_1,Q)}\varepsilon}{\sqrt{n}} + \frac{b'D_{\mu}^{-1}D_{\kappa}F'M^{(Z_1,Q)}\varepsilon}{\sqrt{n}} + b'D_{\mu}^{-1}U'M^{(Z_1,Q)}\varepsilon \\ &= O_p(1) + O_p\left(\frac{\kappa_n^{\max}}{\left(\mu_n^{\min}\right)K_{1,n}^{\varrho_f}}\right) + O_p\left(\frac{n}{\left(\mu_n^{\min}\right)}\right) = O_p\left(\frac{n}{\left(\mu_n^{\min}\right)}\right) \end{aligned}$$

Since the above argument holds for all $b \in \mathbb{R}^d$ such that $\|b\|_2 = 1$, we further deduce that $D_{\mu}^{-1}X'M^{(Z_1,Q)}\varepsilon = O_p\left(n\left(\mu_n^{\min}\right)^{-1}\right)$. $\square$

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 $\underline{U}_{(i,t)} = U_{(i,t)} - \rho\varepsilon_{(i,t)}$.

Proof of Lemma S2-6: Note that, for $1 \leq p \leq 8$ and for any $(i,t) \in \{1, ..., m_n\}$, there exists a positive constant C such that

$$\begin{aligned} E\left[\left\|\underline{U}_{(i,t)}\right\|_2^p |\mathcal{F}_n^W\right] &= E\left[\left(\left\|U_{(i,t)} - \rho\varepsilon_{(i,t)}\right\|_2\right)^p |\mathcal{F}_n^W\right] \\ &\leq 2^{p-1}\left\{\left(E\left[\left\|U_{(i,t)}\right\|_2^8 |\mathcal{F}_n^W\right]\right)^{p/8} + \|\rho\|_2^p\left(E\left[\left|\varepsilon_{(i,t)}\right|^8 |\mathcal{F}_n^W\right]\right)^{p/8}\right\} \\ &\leq C < \infty \ \ a.s., \end{aligned}$$

where the first inequality follows by applying the triangle inequality, Loève's c_r inequality, and Liapunov's inequality in sequence and where the second inequality follows from applying Assumption 2(i) and and from the fact that $\rho \in \mathcal{S}_{\rho}$, some compact subset of $\mathbb{R}^d$ as stated in Assumption 8. Since the upper bound above holds for all $(i,t) \in \{1, ..., m_n\}$, for all $\rho \in \mathcal{S}_{\rho}$, and for all n, it further follows 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.*, as required. $\square$

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) $\widehat{\ell}_{F,n} = o_p\left(\left[\mu_n^{\min}\right]^2/n\right)$.

Proof of Lemma S2-7: To proceed, first define

$$\overline{D}_n = \begin{pmatrix} \mu_n^{\min} & 0 \\ 0 & D_{\mu} \end{pmatrix} \text{ and } L_{\delta} = \begin{pmatrix} 1 & 0 \\ \delta_0 & I_d \end{pmatrix},$$

and note that

$$L_{\delta}^{-1} = \begin{pmatrix} 1 & 0 \\ -\delta_0 & I_d \end{pmatrix}$$

Now, for any $\beta \in \mathbb{R}^{d+1}$ such that $\|\beta\|_2 = 1$ and for $\overline{X} = \begin{bmatrix} y & X \end{bmatrix}$, we can write $\beta'\overline{X}'A\overline{X}\beta =$

$\beta' L_\delta' \overline{D}_n \left(\overline{D}_n^{-1} L_\delta'^{-1} \overline{X}' A \overline{X} L_\delta^{-1} \overline{D}_n^{-1} \right) \overline{D}_n L_\delta \beta$. Moreover, by direct multiplication,

$$\overline{D}_n^{-1} L_\delta'^{-1} \overline{X}' A \overline{X} L_\delta^{-1} \overline{D}_n^{-1} = \begin{pmatrix} (y - X\delta_0)' A (y - X\delta_0) / \left(\mu_n^{\min}\right)^2 & (y - X\delta_0)' A X D_\mu^{-1} / \left(\mu_n^{\min}\right) \\ D_\mu^{-1} X' A (y - X\delta_0) / \left(\mu_n^{\min}\right) & D_\mu^{-1} X' A X D_\mu^{-1} \end{pmatrix},$$

Now, by straightforward but tedious calculations and by applying Assumptions 1, 2(i), 3(ii)-(iii), 4, 5, 6(i), and 7(i)-(iii) as well as Lemmas S2-3(c) and S2-1(a); we can show that, under the rate condition $\sqrt{K_2} / \left(\mu_n^{\min}\right)^2 \to 0$,

$$\overline{D}_n^{-1} L_\delta'^{-1} \overline{X}' A \overline{X} L_\delta^{-1} \overline{D}_n^{-1} = \begin{pmatrix} 0 & 0 \\ 0 & \Gamma' M^{(Z_1, Q)} \Gamma / n \end{pmatrix} + o_p(1)^1,$$

where, in light of Assumption 3(iii), $\Gamma' M^{(Z_1,Q)} \Gamma / n = O_p(1)$ and $\Gamma' M^{(Z_1,Q)} \Gamma / n$ is positive definite for all n large sufficiently large. It follows that $\overline{D}_n^{-1} L_\delta'^{-1} \overline{X}' A \overline{X} L_\delta^{-1} \overline{D}_n^{-1}$ is positive semi-definite w.p.a.1, so that $\beta' \overline{X}' A \overline{X} \beta = \beta' L_\delta' \overline{D}_n \left(\overline{D}_n^{-1} L_\delta'^{-1} \overline{X}' A \overline{X} L_\delta^{-1} \overline{D}_n^{-1} \right) \overline{D}_n L_\delta \beta \geq 0$ w.p.a.1 for all $\beta \in \mathbb{R}^{d+1}$ such that $\|\beta\|_2 = 1$. Moreover, by straightforward calculations, we can show that

$$\frac{\beta' \overline{X}' M^{(Z_1,Q)} \overline{X} \beta}{n} = \frac{\beta' L_{\delta,2}' D_\mu \Gamma' M^{(Z_1,Q)} \Gamma D_\mu L_{\delta,2} \beta}{n^2} + \beta' L_\delta' E\left[\frac{V' M^Q V}{n} \right] L_\delta \beta + o_p(1),$$

where $V = \begin{bmatrix} \varepsilon & U \end{bmatrix}$ and $L_{\delta,2} = \begin{bmatrix} \delta_0 & I_d \end{bmatrix}$ and where $E\left[V' M^Q V / n\right]$ is positive definite for all n sufficiently large in light of Assumptions 2(ii) and 6(i). Since $L_\delta \beta \neq 0$ for all $\beta \in \mathbb{R}^{d+1}$ such that $\|\beta\|_2 = 1$, it follows that $\beta' \overline{X}' M^{(Z_1,Q)} \overline{X} \beta / n > 0$ w.p.a.1 for all $\beta \in \mathbb{R}^{d+1}$ such that $\|\beta\|_2 = 1$. Hence, with probability approaching one as $n \to \infty$,

$$R(\beta) = \frac{\beta' \overline{X}' A \overline{X} \beta}{\beta' \overline{X}' M^{(Z_1,Q)} \overline{X} \beta}$$

is a continuous function of β for all values of β such that $\|\beta\|_2 = 1$. The Weierstrass extreme value theorem then implies that there exists some $\widetilde{\beta}$ such that $\widetilde{\beta} = \arg\min_{\|\beta\|_2 = 1} R(\beta)$ w.p.a.1. Next, note that $\widehat{\ell}_{L,n}$ is the smallest root of the determinantal equation $\det\left\{ \overline{X}' A \overline{X} - \ell \overline{X}' M^{(Z_1,Q)} \overline{X} \right\} = 0$; and, thus, $\widehat{\ell}_{L,n}$ has the representation

$$\widehat{\ell}_{L,n} = R\left(\widetilde{\beta}\right) = \frac{\widetilde{\beta}' \overline{X}' A \overline{X} \widetilde{\beta}}{\widetilde{\beta}' \overline{X}' M^{(Z_1,Q)} \overline{X} \widetilde{\beta}} = \min_{\|\beta\|_2 = 1} \left(\frac{\beta' \overline{X}' A \overline{X} \beta}{\beta' \overline{X}' M^{(Z_1,Q)} \overline{X} \beta} \right).$$

Now, let $\delta_* = \begin{pmatrix} 1 & -\delta_0' \end{pmatrix}' / \left\| \begin{pmatrix} 1 & -\delta_0' \end{pmatrix}' \right\|_2$; and we have, with probability approaching one as $n \to \infty$,

$$
\begin{aligned}
0 &\leq \widehat{\ell}_{L,n} = \min_{\|\beta\|_2=1} \left(\frac{\beta' \overline{X}' A \overline{X} \beta}{\beta' \overline{X}' M^{(Z_1,Q)} \overline{X} \beta} \right) \\
&\leq \frac{\delta_*' \overline{X}' A \overline{X} \delta_*}{\delta_*' \overline{X}' M^{(Z_1,Q)} \overline{X} \delta_*} \\
&= \frac{\left(\mu_n^{\min}\right)^2}{n} \left\{ \frac{(y - X\delta_0)' A (y - X\delta_0) / \sqrt{K_{2,n}}}{(y - X\delta_0) M^{(Z_1,Q)} (y - X\delta_0) / n} \right\} \frac{\sqrt{K_{2,n}}}{\left(\mu_n^{\min}\right)^2} \\
&= O\left(\frac{\left[\mu_n^{\min}\right]^2}{n} \right) O_p(1) \, O_p\left(\frac{\sqrt{K_{2,n}}}{\left(\mu_n^{\min}\right)^2} \right) = o_p\left(\frac{\left[\mu_n^{\min}\right]^2}{n} \right),
\end{aligned}
$$

given the rate condition $\sqrt{K_2} / \left(\mu_n^{\min}\right)^2 \to 0$. This shows part (a).

For part (b), we use the result in part (a) above and the fact that $m_n/n \sim 1$ by Assumption 5(i) to obtain

$$
\widehat{\ell}_{F,n} = \frac{\widehat{\ell}_{L,n} - \left(1 - \widehat{\ell}_{L,n}\right) C/m_n}{1 - \left(1 - \widehat{\ell}_{L,n}\right) C/m_n} = \left[\widehat{\ell}_{L,n} + O_p\left(\frac{1}{n}\right) \right] \left[1 + O_p\left(\frac{1}{n}\right) \right] = o_p\left(\frac{\left[\mu_n^{\min}\right]^2}{n} \right). \ \square \tag{23}
$$

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, $\max_{1 \leq (i,t) \leq m_n} E\left(u_{(i,t),n}^4 | \mathcal{F}_n^W\right) \leq C$, $\max_{1 \leq (i,t) \leq m_n} E\left(\varepsilon_{(i,t),n}^4 | \mathcal{F}_n^W\right) \leq C$, 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) $K_{2,n}^{-1} \sum_{1 \leq (j,s) < (i,t) \leq m_n} A_{(i,t),(j,s)}^2 \phi_{(i,t),n} \left\{ u_{(j,s),n} \varepsilon_{(j,s),n} - \overline{\psi}_{(j,s),n} \right\} \xrightarrow{p} 0$;
(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) $K_{2,n}^{-1} \sum_{1 \leq (k,v) < (j,s) < (i,t) \leq m_n} A_{(i,t),(j,s)} A_{(i,t),(k,v)} \phi_{(i,t),n} u_{(j,s),n} u_{(k,v),n} \xrightarrow{p} 0$.

Proof of Lemma S2-8: This lemma can be proved by generalizing the argument for parts (i), (iv), (v), and (vi) of Lemma B4 of Chao et al (2012) to our setting here with cluster sampling, fixed effects, and possibly many covariates. For the sake of brevity, we will not include an explicit proof here, but the details of a proof can be obtained from the authors upon request. $\square$

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 $\rho = \lim_{n\to\infty} E\left[U' M^Q \varepsilon\right] / E\left[\varepsilon' M^Q \varepsilon\right]$.

Proof of Lemma S2-9: This lemma can be proved by generalizing the argument of Lemma A8 of Hausman et al (2012) to our setting here with cluster sampling, fixed effects, and possibly many covariates. For the sake of brevity, we will not include an explicit proof here, but the details of a proof can be obtained from the authors upon request.

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

$$D_\mu^{-1} X' \left[A - \widehat{\ell}\left(\overline{\delta}_n\right) M^{(Z_1,Q)}\right] X D_\mu^{-1} = H_n + o_p(1). \tag{24}$$

Proof of Lemma S2-10: This lemma can be proved by generalizing the argument of Lemma A7 of Hausman et al (2012) to our setting here with cluster sampling, fixed effects, and possibly many covariates. For the sake of brevity, we will not include an explicit proof here, but the details of a proof can be obtained from the authors upon request. □

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

$$\widehat{\ell}_L = \frac{\left(y - X\widehat{\delta}_L\right)' A \left(y - X\widehat{\delta}_L\right)}{\left(y - X\widehat{\delta}_L\right)' M^{(Z_1,Q)} \left(y - X\widehat{\delta}_L\right)}, \tag{25}$$

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 $\widehat{Q}_{FELIM}(\delta) = (y - X\delta)' A (y - X\delta) / \left[(y - X\delta)' M^{(Z_1,Q)} (y - X\delta)\right]$.

Proof of Lemma S2-11:

Note that the first-order condition for minimizing the objective function $Q(\beta)$ can be written as $\partial Q\left(\widetilde{\beta}\right)/\partial\beta = 2\overline{X}' A \overline{X}\widetilde{\beta} / \left(\widetilde{\beta}'\overline{X}' M^{(Z_1,Q)}\overline{X}\widetilde{\beta}\right)$ $-\widetilde{\beta}'\overline{X}' A \overline{X}\widetilde{\beta}\left(2\overline{X}' M^{(Z_1,Q)}\overline{X}\widetilde{\beta}\right) / \left[\widetilde{\beta}'\overline{X}' M^{(Z_1,Q)}\overline{X}\widetilde{\beta}\right]^2 = 0$. Pre-multiplying this first order condition by the factor $\frac{1}{2}\widetilde{\beta}'\overline{X}' M^{(Z_1,Q)}\overline{X}\widetilde{\beta}$, we then obtain

$$0 = \left[\overline{X}' A \overline{X} - \widehat{\ell}_L \overline{X}' M^{(Z_1,Q)}\overline{X}\right]\widetilde{\beta} \tag{26}$$

where $\widehat{\ell}_L = Q\left(\widetilde{\beta}\right) = \widetilde{\beta}'\overline{X}' A \overline{X}\widetilde{\beta} / \widetilde{\beta}'\overline{X}' M^{(Z_1,Q)}\overline{X}\widetilde{\beta}$. It is clear that in order for there to be a nontrivial solution, i.e., $\widetilde{\beta} \neq 0$ such that equation (26) is true, $\widehat{\ell}_L$ must be a root of the determinantal equation $\det\left[\overline{X}' A \overline{X} - \ell\overline{X}' M^{(Z_1,Q)}\overline{X}\right] = 0$. Moreover, since our goal is to minimize the value of the objective function $Q(\beta)$, this implies that we should choose $\widehat{\ell}_L$ to be the smallest root of this determinantal equation. Now, define $\widetilde{\delta} = -\widetilde{\beta}_2/\widetilde{\beta}_1$, and rewrite the first-order conditions given by expression (26) as $0 = \overline{X}' A \overline{X}\widetilde{\beta} - \widehat{\ell}_L\overline{X}' M^{(Z_1,Q)}\overline{X}\widetilde{\beta} = \widetilde{\beta}_1\left\{\overline{X}' A\left(y - X\widetilde{\delta}\right) - \widehat{\ell}_L\overline{X}' M^{(Z_1,Q)}\left(y - X\widetilde{\delta}\right)\right\}$, so that, given the condition that $\left|\widetilde{\beta}_1\right| \geq \underline{C} > 0$ $a.s.n.$ for some constant $\underline{C}$ (as stated in Assumption 9), we must have

$$\overline{X}' A\left(y - X\widetilde{\delta}\right) - \widehat{\ell}_L\overline{X}' M^{(Z_1,Q)}\left(y - X\widetilde{\delta}\right) = 0 \tag{27}$$

Since $\overline{X} = [y, X]$, we can partition (27) into two sets of equations

$$\begin{aligned} 0 &= y' A\left(y - X\widetilde{\delta}\right) - \widehat{\ell}_L y' M^{(Z_1,Q)}\left(y - X\widetilde{\delta}\right), & (28) \\ 0 &= X' A\left(y - X\widetilde{\delta}\right) - \widehat{\ell}_L X' M^{(Z_1,Q)}\left(y - X\widetilde{\delta}\right). & (29) \end{aligned}$$

Solving (29) for $\widetilde{\delta}$, we obtain $\widetilde{\delta} = \left(X'\left[A - \widehat{\ell}_L M^{(Z_1,Q)}\right]X\right)^{-1} X'\left[A - \widehat{\ell}_L M^{(Z_1,Q)}\right]y = \widehat{\delta}_L$, so that the FELIM estimator $\widehat{\delta}_L$ is a solution to the second set of equations given by (29). In addition , note that, under condition (11) in Assumption 9, we have

$$\widehat{\ell}_L = \frac{\widetilde{\beta}'\overline{X}' A \overline{X}\widetilde{\beta}}{\widetilde{\beta}'\overline{X}' M^{(Z_1,Q)}\overline{X}\widetilde{\beta}} = \frac{\widetilde{\beta}_1\left(y - X\widetilde{\delta}\right)' A\left(y - X\widetilde{\delta}\right)\widetilde{\beta}_1}{\widetilde{\beta}_1\left(y - X\widetilde{\delta}\right) M^{(Z_1,Q)}\left(y - X\widehat{\delta}\right)\widetilde{\beta}_1} = \frac{\left(y - X\widehat{\delta}_L\right)' A\left(y - X\widehat{\delta}_L\right)}{\left(y - X\widehat{\delta}_L\right)' M^{(Z_1,Q)}\left(y - X\widehat{\delta}_L\right)}.$$

which shows (25). Furthermore, note that $\widehat{\delta}_L$ also satisfies equation (28) since

$$\begin{aligned}
& y'A\left(y-X\widehat{\delta}_L\right)-\widehat{\ell}_L y' M^{(Z_1,Q)}\left(y-X\widehat{\delta}_L\right) \\
= & \left(y-X\widehat{\delta}_L\right)'\left[A-\widehat{\ell}_L M^{(Z_1,Q)}\right]\left(y-X\widehat{\delta}_L\right)+\widehat{\delta}_L' X'\left[A-\widehat{\ell}_L M^{(Z_1,Q)}\right]\left(y-X\widehat{\delta}_L\right) \\
= & \left(y-X\widehat{\delta}_L\right)' A\left(y-X\widehat{\delta}_L\right)-\frac{\left(y-X\widehat{\delta}_L\right)' A\left(y-X\widehat{\delta}_L\right)\left(y-X\widehat{\delta}_L\right)' M^{(Z_1,Q)}\left(y-X\widehat{\delta}_L\right)}{\left(y-X\widehat{\delta}_L\right) M^{(Z_1,Q)}\left(y-X\widehat{\delta}_L\right)} \\
& +\widehat{\delta}_L' X'\left[A-\widehat{\ell}_L M^{(Z_1,Q)}\right] y \\
& -\widehat{\delta}_L' X'\left[A-\widehat{\ell}_L M^{(Z_1,Q)}\right] X\left(X'\left[A-\widehat{\ell}_L M^{(Z_1,Q)}\right] X\right)^{-1} X'\left[A-\widehat{\ell}_L M^{(Z_1,Q)}\right] y \\
= & \ 0
\end{aligned}$$

from which we further deduce that $\widehat{\delta}_L$ is a solution of the complete set of first-order conditions given by (27). Finally, since $\widehat{\Delta}\left(\widehat{\delta}_L\right)=X'A\left(y-X\widehat{\delta}_L\right)-\widehat{\ell}_L X' M^{(Z_1,Q)}\left(y-X\widehat{\delta}_L\right)$, the fact that $\widehat{\delta}_L$ is a solution of (29) directly imply that $\widehat{\delta}_L$ satisfies the set of (normalized) first-order conditions for minimizing the variance ratio objective function. □

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 $\overline{X}=\left[\begin{array}{cc} y & X \end{array}\right]$.

Proof of Lemma S2-12: The result follows directly from applying part (b) of Lemma S2-7 and parts (a) and (b) of Lemma S2-2. □

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 $\rho=\lim_{n\to\infty} E\left[U' M^Q\varepsilon\right]/E\left[\varepsilon' M^Q\varepsilon\right]$.

Proof of Lemma S2-13:

Note that, by Lemma S2-11 above, $\widehat{\ell}_{L,n}$ has the representation

$$\widehat{\ell}_{L,n}=\frac{\left(y-X\widehat{\delta}_{L,n}\right)' A\left(y-X\widehat{\delta}_{L,n}\right)}{\left(y-X\widehat{\delta}_{L,n}\right)' M^{(Z_1,Q)}\left(y-X\widehat{\delta}_{L,n}\right)}.$$

Next, from expression (23), we have $\widehat{\ell}_{F,n}=\widehat{\ell}_{L,n}+O_p\left(n^{-1}\right)$. It then follows, by tedious but straightforward calculations[2] and by making use of Assumptions 1, 2(i), 3-6, 7(i)-(iii), and 8 that

$$
\begin{aligned}
& D_{\mu}^{-1} X' \left[A - \widehat{\ell}_{F,n} M^{(Z_1,Q)}\right](y - X\delta_0) \\
= \;& D_{\mu}^{-1} X' A (y - X\delta_0) - \widehat{\ell}_{L,n} D_{\mu}^{-1} X' M^{(Z_1,Q)} (y - X\delta_0) + O_p\left(\frac{1}{n}\right) O_p\left(\frac{n}{\mu_n^{\min}}\right) \\
= \;& \frac{\Gamma' M^{(Z_1,Q)} \varepsilon}{\sqrt{n}} + D_{\mu}^{-1} (U - \varepsilon\rho')' A\varepsilon \left[1 + o_p(1)\right] + O_p\left(\frac{1}{\mu_n^{\min}} \max\left\{\sqrt{\frac{K_{2,n}}{n}}, \frac{K_{1,n}\sqrt{K_{2,n}}}{n}\right\}\right) \\
& + O_p\left(\frac{\sqrt{K_{2,n}}}{n}\right) + O_p\left(\frac{1}{K_{2,n}^{\varrho_\gamma}}\right) + O_p\left(\frac{\kappa_n^{\max}}{\mu_n^{\min}} \frac{1}{K_{1,n}^{\varrho_f}}\right) + O_p\left(\frac{\tau_n}{K_{1,n}^{\varrho_g}}\right) + O_p\left(\frac{1}{\mu_n^{\min}}\right) \\
= \;& \frac{\Gamma' M^{(Z_1,Q)} \varepsilon}{\sqrt{n}} + D_{\mu}^{-1} \underline{U}' A\varepsilon \left[1 + o_p(1)\right], \text{ where } \underline{U} = U - \varepsilon\rho'. \; \Box
\end{aligned}
$$

Lemma S2-14: For any $a \in \mathbb{R}^d$ such that $\|a\| = 1$, define $b_{1n} = \Sigma_n^{-1/2} a$, $\underline{u}_{2,(i,t),n} = b_{2n}' \underline{U}_{(i,t)} = \sqrt{K_{2,n}} a' \Sigma_n^{-1/2} D_{\mu}^{-1} \underline{U}_{(i,t)}$, $\sigma^2_{(i,t),n} = E\left[\varepsilon^2_{(i,t)} | \mathcal{F}_n^W\right]$, $\widetilde{\psi}_{(i,t),n} = E\left[\underline{u}_{2,(i,t),n} \varepsilon_{(i,t)} | \mathcal{F}_n^W\right]$, and $\widetilde{\omega}^2_{(i,t)} = E\left[\underline{u}^2_{2,(i,t),n} | \mathcal{F}_n^W\right]$. Suppose that Assumptions 1-2 and 5-6 are satisfied. Then, the following statements are true.

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

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

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

Proof of Lemma S2-14:

To show part (a), first let $\mathfrak{W}_n = \sum_{(i,t)=2}^{m_n} \sum_{(j,s)=1}^{(i,t)-1} \left[b_{1n}' \Gamma' M^{(Z_1,Q)} e_{(i,t)} / \sqrt{n}\right] \left(A_{(i,t),(j,s)} / \sqrt{K_{2,n}}\right) \times \left\{\varepsilon_{(j,s)} \widetilde{\psi}_{(i,t),n} + \underline{u}_{2,(j,s),n} \sigma^2_{(i,t),n}\right\}$. By taking expectation and applying the triangle inequality, we obtain $E\left[\mathfrak{W}_n^2 | \mathcal{F}_n^W\right] \leq \mathcal{H}_1 + \mathcal{H}_2 + \mathcal{H}_3 + \mathcal{H}_4$, where $\mathcal{H}_1 = \sum_{(i,t),(k,v)=2}^{m_n} \Big| n^{-1} \left[b_{1n}' \Gamma' M^{(Z_1,Q)} e_{(i,t)}\right] \times \left[b_{1n}' \Gamma' M^{(Z_1,Q)} e_{(k,v)}\right] \widetilde{\psi}_{(i,t),n} \widetilde{\psi}_{(k,v),n} \sum_{(j,s)=1}^{\min\{(i,t),(k,v)\}-1} \left(A_{(i,t),(j,s)} A_{(k,v),(j,s)} \sigma^2_{(j,s),n} / K_{2,n}\right) \Big|$,
$\mathcal{H}_2 = \sum_{(i,t),(k,v)=2}^{m_n} \Big| \left(\left[b_{1n}' \Gamma' M^{(Z_1,Q)} e_{(i,t)}\right] \left[b_{1n}' \Gamma' M^{(Z_1,Q)} e_{(k,v)}\right] / n\right) \sigma^2_{(i,t),n} \widetilde{\psi}_{(k,v),n} \times \sum_{(j,s)=1}^{\min\{(i,t),(k,v)\}-1} \left(A_{(i,t),(j,s)} A_{(k,v),(j,s)} \widetilde{\psi}_{(j,s),n} / K_{2,n}\right) \Big|$, $\mathcal{H}_3 = \sum_{(i,t),(k,v)=2}^{m_n} \Big| n^{-1} \left[b_{1n}' \Gamma' M^{(Z_1,Q)} e_{(i,t)}\right]$

[2] Further details are available from the authors upon request.

$\times \left[b_{1n}'\Gamma' M^{(Z_1,Q)} e_{(k,\upsilon)}\right] \sigma^2_{(k,\upsilon),n} \widetilde{\psi}_{(i,t),n} \sum_{(j,s)=1}^{\min\{(i,t),(k,\upsilon)\}-1} \left(A_{(i,t),(j,s)} A_{(k,\upsilon),(j,s)} \widetilde{\psi}_{(j,s),n}/K_{2,n}\right)\Big|,$

$\mathcal{H}_4 = \sum_{(i,t),(k,\upsilon)=2}^{m_n} \Big| \left(\left[b_{1n}'\Gamma' M^{(Z_1,Q)} e_{(i,t)}\right] \left[b_{1n}'\Gamma' M^{(Z_1,Q)} e_{(k,\upsilon)}\right]/n\right) \sigma^2_{(i,t),n} \sigma^2_{(k,\upsilon),n}$

$\times \sum_{(j,s)=1}^{\min\{(i,t),(k,\upsilon)\}-1} \left(A_{(i,t),(j,s)} A_{(k,\upsilon),(j,s)} \widetilde{\omega}^2_{(j,s),n}/K_{2,n}\right)\Big|$. Focusing first on $\mathcal{H}_1$, we obtain, by applying the CS inequality,

$$
\begin{aligned}
&\mathcal{H}_1 \\
&\leq \sqrt{\sum_{(i,t),((k,\upsilon))=2}^{m_n} \frac{b_{1n}'\Gamma' M^{(Z_1,Q)} e_{(i,t)} e_{(i,t)}' M^{(Z_1,Q)} \Gamma b_{1n} \widetilde{\psi}^2_{(i,t),n}}{n} \frac{b_{1n}'\Gamma' M^{(Z_1,Q)} e_{(k,\upsilon)} e_{(k,\upsilon)}' M^{(Z_1,Q)} \Gamma b_{1n} \widetilde{\psi}^2_{(k,\upsilon),n}}{n}} \\
&\quad \times \sqrt{\sum_{(i,t)=2}^{m_n} \sum_{(k,\upsilon)=2}^{m_n} \left(\sum_{(j,s)=1}^{\min\{(i,t),(k,\upsilon)\}-1} \frac{A_{(i,t),(j,s)} A_{(k,\upsilon),(j,s)} \sigma^2_{(j,s),n}}{K_{2,n}}\right)^2}
\end{aligned}
$$

Applying the CS inequality, Assumption 2(i), part (d) of Lemma S2-3, and Lemma S2-6, we obtain

$$
\begin{aligned}
\max_{1\leq(i,t)\leq m_n} \left|\widetilde{\psi}_{(i,t),n}\right| &= \max_{1\leq(i,t)\leq m_n} \sqrt{K_{2,n}} E\left[\left|\varepsilon_{(i,t)} \underline{U}_{(i,t)}' D_\mu^{-1} \Sigma_n^{-1/2} a\right| |\mathcal{F}_n^W\right] \\
&\leq \frac{\sqrt{K_{2,n}}}{(\mu_n^{\min})} \sqrt{a'\Sigma_n^{-1} a} \sqrt{\max_{1\leq(i,t)\leq m_n} E\left[\varepsilon^2_{(i,t)} |\mathcal{F}_n^W\right]} \sqrt{\max_{1\leq(i,t)\leq m_n} E\left[\left\|\underline{U}_{(i,t)}\right\|_2^2 |\mathcal{F}_n^W\right]} \\
&= O_{a.s.}\left(\frac{\sqrt{K_{2,n}}}{(\mu_n^{\min})}\right) \qquad (30)
\end{aligned}
$$

Moreover, by direct calculations,
$\sum_{(i,t)=2}^{m_n} \sum_{(k,\upsilon)=2}^{m_n} \left(\sum_{(j,s)=1}^{\min\{(i,t),(k,\upsilon)\}-1} A_{(i,t),(j,s)} A_{(k,\upsilon),(j,s)} \sigma^2_{(j,s),n}/K_{2,n}\right)^2 = K_{2,n}^{-2} tr\left\{L D_{\sigma^2} L' L D_{\sigma^2} L'\right\}$,
where L is the lower triangular matrix such that $L_{(i,t),(j,s)} = A_{(i,t),(j,s)} \mathbb{I}\left\{(i,t) > (j,s)\right\}$ and $D_{\sigma^2} = diag\left(\sigma^2_{(1,1),n}, ..., \sigma^2_{(n,T_n),n}\right) = diag\left(\sigma^2_{1,n}, ..., \sigma^2_{m_n,n}\right)$ and where $\sigma^2_{(i,t),n} = E\left[\varepsilon^2_{(i,t)} |\mathcal{F}_n^W\right]$ for $(i,t) = 1, ..., m_n$. In addition, by generalizing the results of Lemma B3 of Chao et al (2012) to our setting here, we can show that $\left\|LL'\right\|_F = O_{a.s.}\left(\sqrt{K_{2,n}}\right)$. Using these results, we further deduce, by applying the CS inequality, Assumptions 2 and 3(iii), and part (d) of Lemma S2-3 that

$$
\begin{aligned}
\mathcal{H}_1 &\leq \left(\max_{1\leq(i,t)\leq m_n} \left|\widetilde{\psi}_{(i,t),n}\right|\right)^2 \left(\frac{b_{1n}'\Gamma'\Gamma b_{1n}}{n}\right) \frac{1}{K_{2,n}} \sqrt{tr\left\{L' L D_{\sigma^2} L' L D_{\sigma^2}\right\}} \\
&\leq \left(\max_{1\leq(i,t)\leq m_n} \left|\widetilde{\psi}_{(i,t),n}\right|\right)^2 \left(\frac{b_{1n}'\Gamma'\Gamma b_{1n}}{n}\right) \frac{1}{K_{2,n}} \left(tr\left\{L' L D^2_{\sigma^2} L' L\right\}\right)^{1/4} \left(tr\left\{D_{\sigma^2} L' L L' L D_{\sigma^2}\right\}\right)^{1/4}
\end{aligned}
$$

$$
\begin{aligned}
&\leq \left(\max_{1 \leq (i,t) \leq m_n} \left| \widetilde{\psi}_{(i,t),n} \right| \right)^2 \left(\max_{1 \leq (i,t) \leq m_n} \sigma^2_{(i,t),n} \right) \left(\frac{b'_{1n} \Gamma' \Gamma b_{1n}}{n} \right) \frac{1}{K_{2,n}} \left\| LL' \right\|_F \\
&= O_{a.s.} \left(\frac{\sqrt{K_{2,n}}}{(\mu_n^{\min})^2} \right). \qquad (31)
\end{aligned}
$$

Similarly, let $D_{\widetilde{\psi}} = diag\left(\widetilde{\psi}_{(1,1),n}, ..., \widetilde{\psi}_{(n,T_n),n}\right) = diag\left(\widetilde{\psi}_{1,n}, ..., \widetilde{\psi}_{m_n,n}\right)$, we can also show

$$
\begin{aligned}
\mathcal{H}_2 &\leq \left(\max_{1 \leq (i,t) \leq m_n} \left| \widetilde{\psi}_{(i,t),n} \right| \right) \left(\max_{1 \leq (i,t) \leq m_n} \sigma^2_{(i,t),n} \right) \left(\frac{b'_{1n} \Gamma' \Gamma b_{1n}}{n} \right) \frac{1}{K_{2,n}} \sqrt{tr\left\{ L'LD_{\widetilde{\psi}} L'LD_{\widetilde{\psi}} \right\}} \\
&\leq \left(\max_{1 \leq (i,t) \leq m_n} \left| \widetilde{\psi}_{(i,t),n} \right| \right)^2 \left(\max_{1 \leq (i,t) \leq m_n} \sigma^2_{(i,t),n} \right) \left(\frac{b'_{1n} \Gamma' \Gamma b_{1n}}{n} \right) \frac{1}{K_{2,n}} \left\| LL' \right\|_F \\
&= O_{a.s.} \left(\frac{\sqrt{K_{2,n}}}{(\mu_n^{\min})^2} \right), \qquad (32)
\end{aligned}
$$

and

$$
\begin{aligned}
\mathcal{H}_3 &\leq \left(\max_{1 \leq (i,t) \leq m_n} \left| \widetilde{\psi}_{(i,t),n} \right| \right)^2 \left(\max_{1 \leq (i,t) \leq m_n} \sigma^2_{(i,t),n} \right) \left(\frac{b'_{1n} \Gamma' \Gamma b_{1n}}{n} \right) \frac{1}{K_{2,n}} \left\| LL' \right\|_F \\
&= O_{a.s.} \left(\frac{\sqrt{K_{2,n}}}{(\mu_n^{\min})^2} \right). \qquad (33)
\end{aligned}
$$

Moreover, let $D_{\widetilde{\omega}^2} = diag\left(\widetilde{\omega}^2_{(1,1),n}, ..., \widetilde{\omega}^2_{(n,T_n),n}\right) = diag\left(\widetilde{\omega}^2_{1,n}, ..., \widetilde{\omega}^2_{m_n,n}\right)$, and note that

$$
\begin{aligned}
\mathcal{H}_4 &\leq \left(\max_{1 \leq (i,t) \leq m_n} \sigma^2_{(i,t),n} \right)^2 \left(\frac{b'_{1n} \Gamma' \Gamma b_{1n}}{n} \right) \frac{1}{K_{2,n}} \sqrt{tr\left\{ L'LD_{\widetilde{\omega}^2} L'LD_{\widetilde{\omega}^2} \right\}} \\
&\leq \left(\max_{1 \leq (i,t) \leq m_n} \sigma^2_{(i,t),n} \right)^2 \left(\frac{b'_{1n} \Gamma' \Gamma b_{1n}}{n} \right) \left(\max_{1 \leq (i,t) \leq m_n} \widetilde{\omega}^2_{(i,t),n} \right) \frac{\left\| LL' \right\|_F}{K_{2,n}} \\
&= O_{a.s.} \left(\frac{\sqrt{K_{2,n}}}{(\mu_n^{\min})^2} \right), \qquad (34)
\end{aligned}
$$

where the order of magnitude is calculated by applying Assumptions 2 and 3(iii), part (d) of Lemma S2-3, and the fact that $\left\| LL' \right\|_F = O_{a.s.}\left(\sqrt{K_{2,n}}\right)$ and by making use of the result

$$
\begin{aligned}
\max_{1 \leq (i,t) \leq m_n} \widetilde{\omega}^2_{(i,t),n} &= \max_{1 \leq (i,t) \leq m_n} K_{2,n} a' \Sigma_n^{-1/2} D_\mu^{-1} E\left[\underline{U}_{(i,t)} \underline{U}'_{(i,t)} | \mathcal{F}_n^W \right] D_\mu^{-1} \Sigma_n^{-1/2} a \\
&\leq \left(\max_{1 \leq (i,t) \leq m_n} E\left[\left\| \underline{U}_{(i,t)} \right\|_2^2 | \mathcal{F}_n^W \right] \right) \frac{K_{2,n} a' \Sigma_n^{-1} a}{(\mu_n^{\min})^2}
\end{aligned}
$$

$$= O_{a.s.}\left(\frac{K_{2,n}}{(\mu_n^{\min})^2}\right)$$

which can be easily deduced from part (d) of Lemma S2-3 and Lemma S2-6. Combining (31)-(34), we obtain $E\left[\mathfrak{W}_n^2|\mathcal{F}_n^W\right] = O_{a.s.}\left(\sqrt{K_{2,n}}\left(\mu_n^{\min}\right)^{-2}\right)$. Hence, by the law of iterated expectations and by Theorem 16.1 of Billingsley (1995), there exists a constant $\overline{C} < \infty$ such that, for all n sufficiently large, $E\left(\left[\left(\mu_n^{\min}\right)^2/\sqrt{K_{2,n}}\right]\mathfrak{W}_n^2\right) = E_{W_n}\left(E\left\{\left[\left(\mu_n^{\min}\right)^2/\sqrt{K_{2,n}}\right]\mathfrak{W}_n^2|\mathcal{F}_n^W\right\}\right)$ $\leq \overline{C}$. It follows from the Markov's inequality that
$\mathfrak{W}_n = \sum_{(i,t)=2}^{m_n}\sum_{(j,s)=1}^{(i,t)-1}\left[b_{1n}'\Gamma' M^{(Z_1,Q)}e_{(i,t)}/\sqrt{n}\right]\left(A_{(i,t),(j,s)}/\sqrt{K_{2,n}}\right)\left\{\varepsilon_{(j,s)}\widetilde{\psi}_{(i,t),n} + \underline{u}_{2,(j,s),n}\sigma_{(i,t),n}^2\right\}$
$= O_p\left(K_{2,n}^{1/4}/\left(\mu_n^{\min}\right)\right) = o_p(1)$.

Parts (b) and (c) can be proved by generalizing the argument for parts (ii) and (iii) of Lemma B4 of Chao et al (2012) to our setting here with cluster sampling, fixed effects, and possibly many covariates. For the sake of brevity, we will not include an explicit proof here, but the details of a proof can be obtained from the authors upon request. □

Lemma S2-15: Let $\{X_{i,n}, \mathcal{F}_{i,n}, 1 \leq i \leq k_n, n \geq 1\}$ be a square integrable martingale difference array. Suppose that for all $\epsilon > 0$

$$\sum_{i=1}^{k_n} E\left[X_{i,n}^2 \mathbb{I}\left\{|X_{i,n}| > \epsilon\right\}|\mathcal{F}_{i-1,n}\right] \xrightarrow{p} 0 \tag{35}$$

and

$$\sum_{i=1}^{k_n} E\left[X_{i,n}^2|\mathcal{F}_{i-1,n}\right] \xrightarrow{p} 1. \tag{36}$$

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

Proof of Lemma S2-15: The proof of this central limit theorem for square integrable martingale difference array is given in Gänsler and Stute (1977). See also Corollary 3.1 in Hall and Heyde (1980).

Remark: Note that a sufficient condition for condition (35), which we will verify in lieu of (35) in the proof of Theorems 2 and 3 in Appendix S1, is the following

$$\sum_{i=1}^{k_n} E\left[|X_{i,n}|^{2+\delta}\right] \xrightarrow{p} 0, \text{ for some } \delta > 0. \tag{37}$$

Lemma S2-16: Let $\widetilde{L}_n$ be a sequence of $l \times d$, nonrandom matrices (with $l \leq d$) such that $\left\|\widetilde{L}_n\right\|_F^2 \leq \overline{C} < \infty$ for some constant $\overline{C}$, and let $\Sigma_{2,n} = VC\left(D_\mu^{-1}\underline{U}'A\varepsilon|\mathcal{F}_n^W\right)$

$= D_\mu^{-1} V C \left(\underline{U}' A \varepsilon | \mathcal{F}_n^W \right) D_\mu^{-1}$. Suppose that there exists a positive constant $\underline{C}$ such that $\lambda_{\min} \left(\left(\mu_n^{\min} \right)^2 \widetilde{L}_n H_n^{-1} \Sigma_{2,n} H_n^{-1} \widetilde{L}_n' / K_{2,n} \right) \geq \underline{C} > 0$ $a.s.n.$ Furthermore, let $a \in \mathbb{R}^d$ such that $\|a\|_2 = 1$ and let $\underline{u}_{a,(i,t),n} = a' \left(\left(\mu_n^{\min} \right)^2 \widetilde{L}_n H_n^{-1} \Sigma_{2,n} H_n^{-1} \widetilde{L}_n' / K_{2,n} \right)^{-1/2} \widetilde{L}_n D_\mu^{-1} \underline{U}_{(i,t)}$. Suppose that Assumptions 1-2 and 5-6 are satisfied and that $\left(\mu_n^{\min} \right)^2 / K_{2,n} = o(1)$ but $\sqrt{K_{2,n}} / \left(\mu_n^{\min} \right)^2 \to 0$. Under these conditions, the following statements are true:

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

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

Proof of Lemma S2-16:

To proceed, note first that Lemma S2-6 along with the assumptions on $\lambda_{\min} \left(\left(\mu_n^{\min} \right)^2 \widetilde{L}_n H_n^{-1} \Sigma_{2,n} H_n^{-1} \widetilde{L}_n' / K_{2,n} \right)$ and $\left\| \widetilde{L}_n \right\|_F^2$ together imply that

$$
\begin{aligned}
E\left[\underline{u}_{a,(i,t),n}^2 | \mathcal{F}_n^W \right] &= a' \left(\frac{\left(\mu_n^{\min} \right)^2 \widetilde{L}_n H_n^{-1} \Sigma_{2,n} H_n^{-1} \widetilde{L}_n'}{K_{2,n}} \right)^{-1/2} \widetilde{L}_n D_\mu^{-1} E\left[\underline{U}_{(i,t)} \underline{U}_{(i,t)}' | \mathcal{F}_n^W \right] \\
&\quad D_\mu^{-1} \widetilde{L}_n' \left(\frac{\left(\mu_n^{\min} \right)^2 \widetilde{L}_n H_n^{-1} \Sigma_{2,n} H_n^{-1} \widetilde{L}_n'}{K_{2,n}} \right)^{-1/2} a \\
&\leq \frac{1}{\left(\mu_n^{\min} \right)^2} \frac{\max_{1 \leq (i,t) \leq m_n} E\left[\left\| \underline{U}_{(i,t)} \right\|_2^2 | \mathcal{F}_n^W \right] \left\| \widetilde{L}_n \right\|_F^2}{\lambda_{\min} \left(\left(\mu_n^{\min} \right)^2 \widetilde{L}_n H_n^{-1} \Sigma_{2,n} H_n^{-1} \widetilde{L}_n' / K_{2,n} \right)} \\
&= O_{a.s.} \left(\frac{1}{\left(\mu_n^{\min} \right)^2} \right) \qquad (38)
\end{aligned}
$$

and that

$$
\begin{aligned}
E\left[\underline{u}_{a,(i,t),n}^4 | \mathcal{F}_n^W \right] &= E\Bigg[\Bigg(a' \left(\frac{\left(\mu_n^{\min} \right)^2 \widetilde{L}_n H_n^{-1} \Sigma_{2,n} H_n^{-1} \widetilde{L}_n'}{K_{2,n}} \right)^{-1/2} \widetilde{L}_n D_\mu^{-1} \underline{U}_{(i,t)} \underline{U}_{(i,t)}' D_\mu^{-1} \widetilde{L}_n' \\
&\quad \times \left(\frac{\left(\mu_n^{\min} \right)^2 \widetilde{L}_n H_n^{-1} \Sigma_{2,n} H_n^{-1} \widetilde{L}_n'}{K_{2,n}} \right)^{-1/2} a \Bigg)^2 | \mathcal{F}_n^W \Bigg]
\end{aligned}
$$

$$\leq \frac{1}{\left(\mu_n^{\min}\right)^4} \frac{\max_{1\leq(i,t)\leq m_n} E\left[\left\|\underline{U}_{(i,t)}\right\|_2^4 |\mathcal{F}_n^W\right] \left\|\widetilde{L}_n\right\|_F^4}{\left[\lambda_{\min}\left(\left(\mu_n^{\min}\right)^2 \widetilde{L}_n H_n^{-1}\Sigma_{2,n} H_n^{-1}\widetilde{L}_n'/K_{2,n}\right)\right]^2}$$

$$= O_{a.s.}\left(\frac{1}{\left(\mu_n^{\min}\right)^4}\right) \tag{39}$$

For part (a), define

$\mathfrak{z}_n = \left[\left(\mu_n^{\min}\right)^2/K_{2,n}\right] \sum_{(i,t)=2}^{m_n} \sum_{(j,s)=1}^{(i,t)-1} A_{(i,t),(j,s)}^2 \left(\underline{u}_{a,(j,s),n}^2 - E\left[\underline{u}_{a,(j,s),n}^2 |\mathcal{F}_n^W\right]\right) E\left[\varepsilon_{(i,t)}^2|\mathcal{F}_n^W\right]$, and note that we can apply Assumption 2(i), part (c) of Lemma S2-1, and the upper bounds given by (38) and (39) above to obtain

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

Hence, the law of iterated expectations and Theorem 16.1 of Billingsley (1995) imply that there exists a positive constant $\overline{C} < \infty$ such that, for all n sufficiently large, $E\left(n\mathfrak{z}_n^2\right) = E_{W_n}\left(nE\left[\mathfrak{z}_n^2|\mathcal{F}_n^W\right]\right) \leq \overline{C}$. Application of the Markov's inequality then implies that $\mathfrak{z}_n = O_p\left(n^{-1/2}\right) = o_p(1)$, which shows part (a).

Part (b) can be shown in a manner similar to part (a). For the sake of brevity, we will not include an explicit proof here, but the details of a proof can be obtained from the authors upon request. □

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

Proof of Lemma S2-17: This lemma can be proved by generalizing an argument similar to Lemma A8 of Chao et al (2012) to our setting here with cluster sampling, fixed effects, and possibly many covariates. For the sake of brevity, we will not include an explicit proof here, but the details of a proof can be obtained from the authors upon request. □

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 $\sqrt{K_{2,n}}/\left(\mu_n^{\min}\right)^2 \to 0$. Also, define the following notations: let $\widehat{\varepsilon} = M^{(Z,Q)}\left(y - X\widehat{\delta}_n\right)$, $J = \left[M^Q\circ M^Q\right]^{-1}$, $S_1 = X'AD\left(J\left[\widehat{\varepsilon}\circ\widehat{\varepsilon}\right]\right)AX$, $S_2 = \left(\widehat{\varepsilon}\circ\widehat{\varepsilon}\right)' J\left(A\circ A\right) J\left(\widehat{\varepsilon}\iota_d'\circ M^{(Z,Q)}X\right)$,

$\underline{S}_2 = (\widehat{\varepsilon} \circ \widehat{\varepsilon})' J (A \circ A) J \left(\widehat{\varepsilon}\iota_d' \circ \underline{\widehat{U}}\right)$ with $\underline{\widehat{U}} = M^{(Z,Q)}X - \widehat{\varepsilon}\widehat{\rho}_n'$, $S_3 = (\widehat{\varepsilon} \circ \widehat{\varepsilon})' J (A \circ A) J (\widehat{\varepsilon} \circ \widehat{\varepsilon})$, $S_4 = \left(\widehat{\varepsilon}\iota_d' \circ M^{(Z,Q)}X\right)' J (A \circ A) J \left(\widehat{\varepsilon}\iota_d' \circ M^{(Z,Q)}X\right)$, $\underline{S}_4 = \left(\widehat{\varepsilon}\iota_d' \circ \underline{\widehat{U}}\right)' J (A \circ A) J \left(\widehat{\varepsilon}\iota_d' \circ \underline{\widehat{U}}\right)$, and $\Sigma_{1,n} = \Gamma' M^{(Z_1,Q)} D_{\sigma^2} M^{(Z_1,Q)} \Gamma / n$. In addition, define $\sigma^2_{(i,t)} = E\left[\varepsilon^2_{(i,t)} | \mathcal{F}_n^W\right]$, $D_{\sigma^2} = diag\left(\sigma^2_{(1,1)}, ..., \sigma^2_{(n,T_n)}\right)$, $\phi_{(i,t)} = E\left[U_{(i,t)}\varepsilon_{(i,t)} | \mathcal{F}_n^W\right]$, $\Psi_{(i,t)} = E\left[U_{(i,t)}U'_{(i,t)} | \mathcal{F}_n^W\right]$,
$\underline{\phi}_{(i,t)} = E\left[\underline{U}_{(i,t)}\varepsilon_{(i,t)} | \mathcal{F}_n^W\right]$, 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 $\sigma^2_{(i,t)}$, $\phi_{(i,t)}$, $\Psi_{(i,t)}$, $\underline{\phi}_{(i,t)}$, and $\underline{\Psi}_{(i,t)}$ on $\mathcal{F}_n^W = \sigma(W_n)$. Then, under the above conditions, the following statements are true.

(a) $D_\mu^{-1} S_1 D_\mu^{-1} = \Sigma_{1,n} + \sum_{(i,t),(j,s)=1,(i,t)\neq(j,s)}^{m_n} A^2_{(i,t),(j,s)} \sigma^2_{(i,t)} D_\mu^{-1} \Psi_{(j,s)} D_\mu^{-1}$
$+ o_p\left(\max\left\{1, K_{2,n}\left(\mu_n^{\min}\right)^{-2}\right\}\right)$.

(b) $S_3/K_{2,n} - K_{2,n}^{-1} \sum_{(i,t),(j,s)=1,(i,t)\neq(j,s)}^{m_n} A^2_{(i,t),(j,s)} \sigma^2_{(i,t)} \sigma^2_{(j,s)} = o_p(1)$.

(c) $D_\mu^{-1} S_4 D_\mu^{-1} - \sum_{(i,t),(j,s)=1,(i,t)\neq(j,s)}^{m_n} A^2_{(i,t),(j,s)} D_\mu^{-1} \phi_{(i,t)} \phi'_{(j,s)} D_\mu^{-1} = o_p\left(K_{2,n}\left(\mu_n^{\min}\right)^{-2}\right)$.

(d) $\left(\mu_n^{\min}/K_{2,n}\right) S_2 D_\mu^{-1} - \left(\mu_n^{\min}/K_{2,n}\right) \sum_{(i,t),(j,s)=1,(i,t)\neq(j,s)}^{m_n} A^2_{(i,t),(j,s)} \sigma^2_{(i,t)} \phi'_{(j,s)} D_\mu^{-1} = o_p(1)$.

(e) $D_\mu^{-1}\widehat{\rho}_n = O_p\left(\left(\mu_n^{\min}\right)^{-1}\right)$ and $D_\mu^{-1}(\widehat{\rho}_n - \rho) = o_p\left(\left(\mu_n^{\min}\right)^{-1}\right)$, 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) $D_\mu^{-1} \underline{S}_4 D_\mu^{-1} - \sum_{(i,t),(j,s)=1,(i,t)\neq(j,s)}^{m_n} A^2_{(i,t),(j,s)} D_\mu^{-1} \underline{\phi}_{(i,t)} \underline{\phi}'_{(j,s)} D_\mu^{-1} = o_p\left(K_{2,n}\left(\mu_n^{\min}\right)^{-2}\right)$.

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

Proof of Lemma S2-18:

To show part (a), note first that, by making use of the decomposition $M^{(Z,Q)} = M^Q - P^{Z^\perp}$, where $P^{Z^\perp} = M^Q Z \left(Z' M^Q Z\right)^{-1} Z' M^Q$, we can, after some straightforward algebra, write $\widehat{\varepsilon} = -M^{(Z,Q)} X\left(\widehat{\delta}_n - \delta_0\right) + M^{(Z,Q)}\varphi_n + M^Q\varepsilon - P^{Z^\perp}\varepsilon$, from which we obtain

$$\begin{aligned} J\left[\widehat{\varepsilon} \circ \widehat{\varepsilon}\right] &= J\left[M^{(Z,Q)} X\left(\widehat{\delta}_n - \delta_0\right) \circ M^{(Z,Q)} X\left(\widehat{\delta}_n - \delta_0\right)\right] + J\left[M^{(Z,Q)}\varphi_n \circ M^{(Z,Q)}\varphi_n\right] \\ &\quad + J\left[M^Q\varepsilon \circ M^Q\varepsilon\right] + J\left[P^{Z^\perp}\varepsilon \circ P^{Z^\perp}\varepsilon\right] - 2J\left[M^{(Z,Q)}\varphi_n \circ M^{(Z,Q)} X\left(\widehat{\delta}_n - \delta_0\right)\right] \\ &\quad - 2J\left[M^Q\varepsilon \circ M^{(Z,Q)} X\left(\widehat{\delta}_n - \delta_0\right)\right] + 2J\left[P^{Z^\perp}\varepsilon \circ M^{(Z,Q)} X\left(\widehat{\delta}_n - \delta_0\right)\right] \\ &\quad + 2J\left[M^Q\varepsilon \circ M^{(Z,Q)}\varphi_n\right] - 2J\left[P^{Z^\perp}\varepsilon \circ M^{(Z,Q)}\varphi_n\right] - 2J\left[M^Q\varepsilon \circ P^{Z^\perp}\varepsilon\right] \qquad (40) \end{aligned}$$

where $J = \left[M^{Q} \circ M^{Q}\right]^{-1}$. Substituting the right-hand side of (40) into covariance matrix estimator $D_{\mu}^{-1} X' A D \left(J \left[\widehat{\varepsilon} \circ \widehat{\varepsilon}\right]\right) A X D_{\mu}^{-1}$, we get

$$
\begin{aligned}
& D_{\mu}^{-1} S_1 D_{\mu}^{-1} - D_{\mu}^{-1} X' A D \left(\varepsilon \circ \varepsilon\right) A X D_{\mu}^{-1} \\
= \ & \mathcal{T}_{1,n} + \mathcal{T}_{2,n} + \mathcal{T}_{3,n} + \mathcal{T}_{4,n} + \mathcal{T}_{5,n} + \mathcal{T}_{6,n} + \mathcal{T}_{7,n} + \mathcal{T}_{8,n} + \mathcal{T}_{9,n} + \mathcal{T}_{10,n},
\end{aligned}
$$

where $\mathcal{T}_{1,n} = D_{\mu}^{-1} X' A D \left(J \left[M^{(Z,Q)} X \left(\widehat{\delta}_n - \delta_0\right) \circ M^{(Z,Q)} X \left(\widehat{\delta}_n - \delta_0\right)\right]\right) A X D_{\mu}^{-1}$,
$\mathcal{T}_{2,n} = D_{\mu}^{-1} X' A D \left(J \left[M^{(Z,Q)} \varphi_n \circ M^{(Z,Q)} \varphi_n\right]\right) A X D_{\mu}^{-1}$,
$\mathcal{T}_{3,n} = D_{\mu}^{-1} X' A D \left(J \left[M^{Q} \varepsilon \circ M^{Q} \varepsilon\right]\right) A X D_{\mu}^{-1} - D_{\mu}^{-1} X' A D \left(\varepsilon \circ \varepsilon\right) A X D_{\mu}^{-1}$,
$\mathcal{T}_{4,n} = D_{\mu}^{-1} X' A D \left(J \left[P^{Z^{\perp}} \varepsilon \circ P^{Z^{\perp}} \varepsilon\right]\right) A X D_{\mu}^{-1}$,
$\mathcal{T}_{5,n} = -2 D_{\mu}^{-1} X' A D \left(J \left[M^{(Z,Q)} \varphi_n \circ M^{(Z,Q)} X \left(\widehat{\delta}_n - \delta_0\right)\right]\right) A X D_{\mu}^{-1}$,
$\mathcal{T}_{6,n} = -2 D_{\mu}^{-1} X' A D \left(J \left[M^{Q} \varepsilon \circ M^{(Z,Q)} X \left(\widehat{\delta}_n - \delta_0\right)\right]\right) A X D_{\mu}^{-1}$,
$\mathcal{T}_{7,n} = 2 D_{\mu}^{-1} X' A D \left(J \left[P^{Z^{\perp}} \varepsilon \circ M^{(Z,Q)} X \left(\widehat{\delta}_n - \delta_0\right)\right]\right) A X D_{\mu}^{-1}$,
$\mathcal{T}_{8,n} = 2 D_{\mu}^{-1} X' A D \left(J \left[M^{Q} \varepsilon \circ M^{(Z,Q)} \varphi_n\right]\right) A X D_{\mu}^{-1}$,
$\mathcal{T}_{9,n} = -2 D_{\mu}^{-1} X' A D \left(J \left[P^{Z^{\perp}} \varepsilon \circ M^{(Z,Q)} \varphi_n\right]\right) A X D_{\mu}^{-1}$, and
$\mathcal{T}_{10,n} = -2 D_{\mu}^{-1} X' A D \left(J \left[M^{Q} \varepsilon \circ P^{Z^{\perp}} \varepsilon\right]\right) A X D_{\mu}^{-1}$.

Consider the term $\mathcal{T}_{1,n}$. Let $\mathcal{G}_i = \{(\ell, h) : \ell = i \text{ and } h = 1, ..., T_i\}$; and note that for any $a, b \in \mathbb{R}^d$ such that $\|a\|_2 = \|b\|_2 = 1$, we have

$$
\begin{aligned}
|a' \mathcal{T}_{1,n} b| = & \left| \sum_{(i,t)=1}^{m_n} \sum_{(j,s) \neq (i,t)} \sum_{(k,v) \neq (i,t)} \sum_{(p,q)=1}^{m_n} a' D_{\mu}^{-1} X' e_{(j,s)} A_{(i,t),(j,s)} J_{(i,t),(p,q)} \right. \\
& \left. \times e'_{(p,q)} M^{(Z,Q)} X \left(\widehat{\delta}_n - \delta_0\right) \left(\widehat{\delta}_n - \delta_0\right)' X' M^{(Z,Q)} e_{(p,q)} A_{(i,t),(k,v)} e'_{(k,v)} X D_{\mu}^{-1} b \right| \\
= & \left| \sum_{(i,t)=1}^{m_n} \sum_{(j,s) \neq (i,t)} a' D_{\mu}^{-1} X' e_{(j,s)} A_{(i,t),(j,s)} \right. \\
& \times \left(\sum_{(p,q)=1}^{m_n} J_{(i,t),(p,q)} \mathbb{I}\left\{(p,q) \in \mathcal{G}_i\right\} e'_{(p,q)} M^{(Z,Q)} X \left(\widehat{\delta}_n - \delta_0\right) \left(\widehat{\delta}_n - \delta_0\right)' X' M^{(Z,Q)} e_{(p,q)} \right) \\
& \left. \times \sum_{(k,v) \neq (i,t)} A_{(i,t),(k,v)} e'_{(k,v)} X D_{\mu}^{-1} b \right|
\end{aligned}
$$

where $J_{(i,t),(p,q)}$ is the element in the $(i,t)^{th}$ row and the $(p,q)^{th}$ column of the matrix J for $(i,t), (p,q) \in \{1, ..., m_n\}$, where $\mathbb{I}\{\cdot\}$ denotes an indicator function, where $e_{(j,s)}$ is an $m_n \times 1$ elementary vector whose $(j,s)^{th}$ component is 1 and all other components are 0, and where the second equality above follows from the fact that $J_{(i,t),(p,q)} = 0$ for $p \neq i$ due to the sparsity (or block diagonal nature) of J. Now, let $D^{\Sigma J}\left(M^{(Z,Q)}XX'M^{(Z,Q)}\right)$ be a diagonal matrix whose $(i,t)^{th}$ diagonal element is given by $\sum_{q=1}^{T_i}\left|J_{(i,t),(i,q)}\right| e'_{(i,q)}M^{(Z,Q)}XX'M^{(Z,Q)}e_{(i,q)}$. Applying the triangle inequality and the inequality $|XY| \leq (1/2)\,X^2 + (1/2)\,Y^2$, we then obtain

$$\begin{aligned}
|a'\mathcal{T}_{1,n}b| &= \Bigg|\sum_{(i,t)=1}^{m_n} a'D_\mu^{-1}X'Ae_{(i,t)}\sum_{q=1}^{T_i} J_{(i,t),(i,q)}e'_{(i,q)}M^{(Z,Q)}X\left(\widehat{\delta}_n-\delta_0\right)\left(\widehat{\delta}_n-\delta_0\right)' \\
&\quad \times X'M^{(Z,Q)}e_{(i,q)}e'_{(i,t)}AXD_\mu^{-1}b\Bigg| \\
&\leq \frac{1}{2}\sum_{(i,t)=1}^{m_n}\sum_{q=1}^{T_i}\Big\{\left|J_{(i,t),(i,q)}\right| e'_{(i,q)}M^{(Z,Q)}X\left(\widehat{\delta}_n-\delta_0\right)\left(\widehat{\delta}_n-\delta_0\right)'X'M^{(Z,Q)}e_{(i,q)} \\
&\qquad \times a'D_\mu^{-1}X'Ae_{(i,t)}e'_{(i,t)}AXD_\mu^{-1}a\Big\} \\
&\quad +\frac{1}{2}\sum_{(i,t)=1}^{m_n}\sum_{q=1}^{T_i}\Big\{\left|J_{(i,t),(i,q)}\right| e'_{(i,q)}M^{(Z,Q)}X\left(\widehat{\delta}_n-\delta_0\right)\left(\widehat{\delta}_n-\delta_0\right)'X'M^{(Z,Q)}e_{(i,q)} \\
&\qquad \times b'D_\mu^{-1}X'Ae_{(i,t)}e'_{(i,t)}AXD_\mu^{-1}b\Big\} \\
&\leq \frac{1}{2}\left\|\widehat{\delta}_n-\delta_0\right\|_2^2 a'D_\mu^{-1}X'AD^{\Sigma J}\left(M^{(Z,Q)}XX'M^{(Z,Q)}\right)AXD_\mu^{-1}a \\
&\quad +\frac{1}{2}\left\|\widehat{\delta}_n-\delta_0\right\|_2^2 b'D_\mu^{-1}X'AD^{\Sigma J}\left(M^{(Z,Q)}XX'M^{(Z,Q)}\right)AXD_\mu^{-1}b \qquad (41)
\end{aligned}$$

By tedious but straightforward calculations, we can show that $D_\mu^{-1}X'AD^{\Sigma J}\left(M^{(Z,Q)}XX'M^{(Z,Q)}\right)AXD_\mu^{-1} = O_p\left(\max\left\{1, K_{2,n}\left(\mu_n^{\min}\right)^{-2}\right\}\right)$[3]. Hence, given the assumption that $\left\|\widehat{\delta}_n-\delta_0\right\|_2^2 \xrightarrow{p} 0$, it follows from expression (41) that $|a'\mathcal{T}_{1,n}b| = o_p\left(\max\left\{1, K_{2,n}\left(\mu_n^{\min}\right)^{-2}\right\}\right)$. Since the above argument holds for all $a, b \in \mathbb{R}^d$ such that $\|a\|_2 = \|b\|_2 = 1$, we further deduce that $\mathcal{T}_{1,n} = o_p\left(\max\left\{1, K_{2,n}\left(\mu_n^{\min}\right)^{-2}\right\}\right)$. By following a similar method of proof, we can also show that $\mathcal{T}_{k,n} = o_p\left(\max\left\{1, K_{2,n}\left(\mu_n^{\min}\right)^{-2}\right\}\right)$ for $k = 2, ..., 10$. For the sake of brevity, we will not provide detailed proofs for these other terms. Detailed argument for these other terms can be obtained from the authors upon request.

The fact that $\mathcal{T}_{k,n} = o_p\left(\max\left\{1, K_{2,n}\left(\mu_n^{\min}\right)^{-2}\right\}\right)$ for each $k = \{1, ..., 10\}$ further implies

[3] Details are available from the authors upon request.

that

$$D_{\mu}^{-1}S_1D_{\mu}^{-1}-D_{\mu}^{-1}X'AD\left(\varepsilon\circ\varepsilon\right)AXD_{\mu}^{-1}=\sum_{k=1}^{10}\mathcal{T}_{k,n}=o_p\left(\max\left\{1,\frac{K_{2,n}}{(\mu_n^{\min})^2}\right\}\right)\tag{42}$$

Moreover, by the result of Lemma S2-17,

$$D_{\mu}^{-1}X'AD\left(\varepsilon\circ\varepsilon\right)AXD_{\mu}^{-1}=\Sigma_{1,n}+\sum_{\substack{(i,t),(j,s)=1\\(i,t)\neq(j,s)}}^{m_n}A_{(i,t),(j,s)}^2\sigma_{(i,t)}^2D_{\mu}^{-1}\Psi_{(j,s)}D_{\mu}^{-1}+o_p(1).\tag{43}$$

Combining (42) and (43), we further obtain
$D_{\mu}^{-1}S_1D_{\mu}^{-1}=\Sigma_{1,n}+\sum_{(i,t),(j,s)=1,(i,t)\neq(j,s)}^{m_n}A_{(i,t),(j,s)}^2\sigma_{(i,t)}^2D_{\mu}^{-1}\Psi_{(j,s)}D_{\mu}^{-1}+o_p\left(\max\left\{1,K_{2,n}\left(\mu_n^{\min}\right)^{-2}\right\}\right)$, which shows part (a).

To show part (b), write $S_3/K_{2,n}-K_{2,n}^{-1}\sum_{(i,t),(j,s)=1,(i,t)\neq(j,s)}^{m_n}A_{(i,t),(j,s)}^2\sigma_{(i,t)}^2\sigma_{(j,s)}^2=\mathcal{A}+\mathfrak{A}$, where $\mathcal{A}=S_3/K_{2,n}-K_{2,n}^{-1}\sum_{(i,t),(j,s)=1,(i,t)\neq(j,s)}^{m_n}A_{(i,t),(j,s)}^2\varepsilon_{(i,t)}^2\varepsilon_{(j,s)}^2$ and where $\mathfrak{A}=K_{2,n}^{-1}\sum_{(i,t),(j,s)=1,(i,t)\neq(j,s)}^{m_n}A_{(i,t),(j,s)}^2\left(\varepsilon_{(i,t)}^2\varepsilon_{(j,s)}^2-\sigma_{(i,t)}^2\sigma_{(j,s)}^2\right)$. To analyze the term $\mathcal{A}$, note that, by direct calculation, we can obtain the following decomposition

$$\begin{aligned}\mathcal{A}&=\frac{1}{K_{2,n}}\sum_{\substack{(i,t),(j,s)=1\\(i,t)\neq(j,s)}}^{m_n}A_{(i,t),(j,s)}^2\left\{\sum_{h=1}^{T_i}J_{(i,t),(i,h)}\left[e_{(i,h)}'M^{(Z,Q)}\left(y-X\widehat{\delta}_n\right)\right]^2\right.\\&\quad\left.\times\sum_{v=1}^{T_j}J_{(j,s),(j,v)}\left[\left(y-X\widehat{\delta}_n\right)'M^{(Z,Q)}e_{(j,v)}\right]^2\right\}-\frac{1}{K_{2,n}}\sum_{\substack{(i,t),(j,s)=1\\(i,t)\neq(j,s)}}^{m_n}A_{(i,t),(j,s)}^2\varepsilon_{(i,t)}^2\varepsilon_{(j,s)}^2\\&=\sum_{\ell=1}^{6}\mathcal{A}_\ell\end{aligned}\tag{44}$$

where $\mathcal{A}_1=4K_{2,n}^{-1}\sum_{(i,t),(j,s)=1,(i,t)\neq(j,s)}^{m_n}A_{(i,t),(j,s)}^2$
$\times\sum_{h=1}^{T_i}J_{(i,t),(i,h)}\left(e_{(i,h)}'\left[P^{Z^\perp}\varepsilon-M^{(Z,Q)}\varphi_n\right]+e_{(i,h)}'M^{(Z,Q)}X\left[\widehat{\delta}_n-\delta_0\right]\right)\left(e_{(i,h)}'M^Q\varepsilon\right)$
$\times\sum_{v=1}^{T_j}J_{(j,s),(j,v)}\left(e_{(j,v)}'\left[P^{Z^\perp}\varepsilon-M^{(Z,Q)}\varphi_n\right]+e_{(j,v)}'M^{(Z,Q)}X\left[\widehat{\delta}_n-\delta_0\right]\right)\left(e_{(j,v)}'M^Q\varepsilon\right)$,
$\mathcal{A}_2=K_{2,n}^{-1}\sum_{(i,t),(j,s)=1,(i,t)\neq(j,s)}^{m_n}A_{(i,t),(j,s)}^2$
$\times\sum_{h=1}^{T_i}J_{(i,t),(i,h)}\left(e_{(i,h)}'\left[P^{Z^\perp}\varepsilon-M^{(Z,Q)}\varphi_n\right]+e_{(i,h)}'M^{(Z,Q)}X\left[\widehat{\delta}_n-\delta_0\right]\right)^2$
$\times\sum_{v=1}^{T_j}J_{(j,s),(j,v)}\left(e_{(j,v)}'\left[P^{Z^\perp}\varepsilon-M^{(Z,Q)}\varphi_n\right]+e_{(j,v)}'M^{(Z,Q)}X\left[\widehat{\delta}_n-\delta_0\right]\right)^2$,

$\mathcal{A}_3 = -4K_{2,n}^{-1} \sum_{(i,t),(j,s)=1,(i,t)\neq(j,s)}^{m_n} A_{(i,t),(j,s)}^2 \sum_{v=1}^{T_j} J_{(j,s),(j,v)} \left(e'_{(j,v)} M^Q \varepsilon\right)^2$
$\times \sum_{h=1}^{T_i} J_{(i,t),(i,h)} \left(e'_{(i,h)} \left[P^{Z^\perp}\varepsilon - M^{(Z,Q)}\varphi_n\right] + e'_{(i,h)} M^{(Z,Q)} X \left[\widehat{\delta}_n - \delta_0\right]\right) \left(e'_{(i,h)} M^Q \varepsilon\right)$,
$\mathcal{A}_4 = 2K_{2,n}^{-1} \sum_{(i,t),(j,s)=1,(i,t)\neq(j,s)}^{m_n} A_{(i,t),(j,s)}^2 \sum_{v=1}^{T_j} J_{(j,s),(j,v)} \left(e'_{(j,v)} M^Q \varepsilon\right)^2$
$\times \sum_{h=1}^{T_i} J_{(i,t),(i,h)} \left(e'_{(i,h)} \left[P^{Z^\perp}\varepsilon - M^{(Z,Q)}\varphi_n\right] + e'_{(i,h)} M^{(Z,Q)} X \left[\widehat{\delta}_n - \delta_0\right]\right)^2$
, $\mathcal{A}_5 = -4K_{2,n}^{-1} \sum_{(i,t),(j,s)=1,(i,t)\neq(j,s)}^{m_n} A_{(i,t),(j,s)}^2$
$\times \sum_{h=1}^{T_i} J_{(i,t),(i,h)} \left(e'_{(i,h)} \left[P^{Z^\perp}\varepsilon - M^{(Z,Q)}\varphi_n\right] + e'_{(i,h)} M^{(Z,Q)} X \left[\widehat{\delta}_n - \delta_0\right]\right)^2$
$\times \sum_{v=1}^{T_j} J_{(j,s),(j,v)} \left(e'_{(j,v)} \left[P^{Z^\perp}\varepsilon - M^{(Z,Q)}\varphi_n\right] + e'_{(j,v)} M^{(Z,Q)} X \left[\widehat{\delta}_n - \delta_0\right]\right) \left(e'_{(j,v)} M^Q \varepsilon\right)$,
and $\mathcal{A}_6 = K_{2,n}^{-1} \sum_{(i,t),(j,s)=1,(i,t)\neq(j,s)}^{m_n} \Big\{ A_{(i,t),(j,s)}^2$
$\times \sum_{h=1}^{T_i} J_{(i,t),(i,h)} \left(e'_{(i,h)} M^Q \varepsilon\right)^2 \sum_{v=1}^{T_j} J_{(j,s),(j,v)} \left(e'_{(j,v)} M^Q \varepsilon\right)^2 \Big\}$
$-K_{2,n}^{-1} \sum_{(i,t),(j,s)=1,(i,t)\neq(j,s)}^{m_n} A_{(i,t),(j,s)}^2 \varepsilon_{(i,t)}^2 \varepsilon_{(j,s)}^2$.

Consider now the $\mathcal{A}_1$ term. By applying the triangle inequality and the inequality $|XY| \leq (1/2)X^2 + (1/2)Y^2$ and collecting like terms, we get

$$\begin{aligned}
|\mathcal{A}_1| &\leq \frac{12}{K_{2,n}} \sum_{\substack{(i,t),(j,s)=1 \\ (i,t)\neq(j,s)}}^{m_n} A_{(i,t),(j,s)}^2 \sum_{h=1}^{T_i}\sum_{v=1}^{T_i} \left|J_{(i,t),(i,h)}\right| \left|J_{(j,s),(j,v)}\right| \left(e'_{(j,v)} M^Q \varepsilon\right)^2 \left(e'_{(i,h)} P^{Z^\perp} \varepsilon\right)^2 \\
&\quad + \frac{12}{K_{2,n}} \sum_{\substack{(i,t),(j,s)=1 \\ (i,t)\neq(j,s)}}^{m_n} A_{(i,t),(j,s)}^2 \sum_{h=1}^{T_i}\sum_{v=1}^{T_i} \left|J_{(i,t),(i,h)}\right| \left|J_{(j,s),(j,v)}\right| \Big\{ \left(e'_{(i,h)} M^Q \varepsilon\right)^2 \\
&\qquad\qquad \times \left(e'_{(j,v)} M^{(Z,Q)} X \left[\widehat{\delta}_n - \delta_0\right]\right)^2 \Big\} \\
&\quad + \frac{12}{K_{2,n}} \sum_{\substack{(i,t),(j,s)=1 \\ (i,t)\neq(j,s)}}^{m_n} A_{(i,t),(j,s)}^2 \sum_{h=1}^{T_i}\sum_{v=1}^{T_i} \left|J_{(i,t),(i,h)}\right| \left|J_{(j,s),(j,v)}\right| \left(e'_{(i,h)} M^Q \varepsilon\right)^2 \left(e'_{(j,v)} M^{(Z,Q)} \varphi_n\right)^2 \\
&= \mathcal{A}_{1,1} + \mathcal{A}_{1,2} + \mathcal{A}_{1,3}, \ (say).
\end{aligned}$$

Clearly $\mathcal{A}_{1,k} \geq 0$ for $k = 1, 2, 3$. Next, note that, after some straightforward moment calculations and after applying the triangle inequality as well as Assumptions 1, 2(i), 5, and

6 and part (a) of Lemma S2-1; we obtain

$$
\begin{aligned}
E\left[\mathcal{A}_{1,1}|\mathcal{F}_n^W\right] &= \frac{12}{K_{2,n}} \sum_{\substack{(i,t),(j,s)=1 \\ (i,t)\neq(j,s)}}^{m_n} A_{(i,t),(j,s)}^2 \sum_{h=1}^{T_i} \sum_{v=1}^{T_j} \left|J_{(i,t),(i,h)}\right| \left|J_{(j,s),(j,v)}\right| \sum_{q=1}^{T_j} M_{(j,v),(j,q)}^Q \\
&\quad \times \sum_{g=1}^{T_j} M_{(j,v),(j,g)}^Q \sum_{(l,r)=1}^{m_n} P_{(i,h),(l,r)}^{Z^\perp} \sum_{(k,c)=1}^{m_n} P_{(i,h),(k,c)}^{Z^\perp} E\left[\varepsilon_{(l,r)}\varepsilon_{(k,c)}\varepsilon_{(j,q)}\varepsilon_{(j,g)}|\mathcal{F}_n^W\right] \\
&\leq \frac{12\overline{T}^3}{K_{2,n}} \left(\max_{1\leq(i,t),(j,s)\leq m_n} \left|M_{(i,t),(j,s)}^Q\right|\right)^2 \left(\max_{1\leq(j,s)\leq m_n} E\left[\varepsilon_{(j,s)}^4|\mathcal{F}_n^W\right]\right) \\
&\quad \times \left(\max_{1\leq(i,t)\leq m_n} P_{(i,t),(i,t)}^{Z^\perp}\right)^2 \left(\max_{1\leq(i,t),(j,s)\leq m_n} \left|J_{(i,t),(j,s)}\right|\right)^2 \sum_{\substack{(i,t),(j,s)=1 \\ (i,t)\neq(j,s)}}^{m_n} A_{(i,t),(j,s)}^2 \\
&\quad + \frac{12\overline{T}^3}{K_{2,n}} \left(\max_{1\leq(i,t),(j,s)\leq m_n} \left|M_{(i,t),(j,s)}^Q\right|\right)^2 \left(\max_{1\leq(i,t),(j,s)\leq m_n} \left|J_{(i,t),(j,s)}\right|\right)^2 \\
&\quad \times \left(\max_{1\leq(i,t)\leq m_n} E\left[\varepsilon_{(i,t)}^2|\mathcal{F}_n^W\right]\right)^2 \left(\max_{1\leq(i,t)\leq m_n} P_{(i,t),(i,t)}^{Z^\perp}\right) \sum_{\substack{(i,t),(j,s)=1 \\ (i,t)\neq(j,s)}}^{m_n} A_{(i,t),(j,s)}^2 \\
&\quad + 24\overline{T}^4 \left(\max_{1\leq(i,t),(j,s)\leq m_n} \left|J_{(i,t),(j,s)}\right|\right)^2 \left(\max_{1\leq(i,t),(j,s)\leq m_n} \left|M_{(i,t),(j,s)}^Q\right|\right)^2 \\
&\qquad \times \left(\max_{1\leq(i,t)\leq m_n} E\left[\varepsilon_{(i,t)}^2|\mathcal{F}_n^W\right]\right)^2 \left(\max_{1\leq(i,t)\leq m_n} P_{(i,t),(i,t)}^{Z^\perp}\right)^2 \frac{1}{K_{2,n}} \sum_{\substack{(i,t),(j,s)=1 \\ (i,t)\neq(j,s)}}^{m_n} A_{(i,t),(j,s)}^2 \\
&= O_{a.s.}\left(\frac{K_n^2}{n^2}\right) + O_{a.s.}\left(\frac{K_n}{n}\right) + O_{a.s.}\left(\frac{K_n^2}{n^2}\right) = O_{a.s.}\left(\frac{K_n}{n}\right).
\end{aligned} \tag{45}
$$

It follows from the law of iterated expectations and Theorem 16.1 of Billingsley (1995) that there exists a constant $\overline{C} > 0$ such that for all n sufficiently large $E\left[(n/K_n)\left|\mathcal{A}_{1,1}\right|\right] = E_W\left(\frac{n}{K_n} E\left[\left|\mathcal{A}_{1,1}\right| |\mathcal{F}_n^W\right]\right) \leq \overline{C} < \infty$. Application of Markov's inequality then allows us to further deduce that

$$
\mathcal{A}_{1,1} = O_p\left(\frac{K_n}{n}\right) = o_p\left(1\right). \tag{46}
$$

Consider next the subterm $\mathcal{A}_{1,2}$. Here, by applying the CS inequality, we have

$$\begin{aligned}
\mathcal{A}_{1,2} &= \frac{12}{K_{2,n}} \sum_{\substack{(i,t),(j,s)=1 \\ (i,t)\neq(j,s)}}^{m_n} A^2_{(i,t),(j,s)} \sum_{h=1}^{T_i} \sum_{v=1}^{T_i} \left|J_{(i,t),(i,h)}\right| \left|J_{(j,s),(j,v)}\right| \left(e'_{(i,h)} M^Q \varepsilon\right)^2 \left(e'_{(j,v)} M^{(Z,Q)} X \left[\widehat{\delta}_n - \delta_0\right]\right)^2 \\
&\leq \frac{12 \left\|\widehat{\delta}_n - \delta_0\right\|_2^2}{K_{2,n}} \sum_{\substack{(i,t),(j,s)=1 \\ (i,t)\neq(j,s)}}^{m_n} A^2_{(i,t),(j,s)} \sum_{h=1}^{T_i} \sum_{v=1}^{T_j} \left\{ \left|J_{(i,t),(i,h)}\right| \left|J_{(j,s),(j,v)}\right| \left(e'_{(i,h)} M^Q \varepsilon\right)^2 \right. \\
&\qquad \left. \times e'_{(j,v)} M^{(Z,Q)} X X' M^{(Z,Q)} e_{(j,v)} \right\} \\
&\leq \left\|\widehat{\delta}_n - \delta_0\right\|_2^2 \mathcal{A}^*_{1,2},
\end{aligned}$$

where $\mathcal{A}^*_{1,2} = 12 K_{2,n}^{-1} \sum_{(i,t),(j,s)=1,(i,t)\neq(j,s)}^{m_n} A^2_{(i,t),(j,s)} \sum_{h=1}^{T_i} \left|J_{(i,t),(i,h)}\right| \left(e'_{(i,h)} M^Q \varepsilon\right)^2$ $\times \sum_{v=1}^{T_j} \left|J_{(j,s),(j,v)}\right| e'_{(j,v)} M^{(Z,Q)} X X' M^{(Z,Q)} e_{(j,v)}$. Next, by applying the inequality $(x+y)'(x+y) \leq 2x'x + 2y'y$ twice, we obtain

$$\begin{aligned}
0 &\leq \mathcal{A}^*_{1,2} \\
&\leq \frac{24}{K_{2,n}} \sum_{\substack{(i,t),(j,s)=1 \\ (i,t)\neq(j,s)}}^{m_n} A^2_{(i,t),(j,s)} \sum_{h=1}^{T_i} \sum_{v=1}^{T_j} \left|J_{(i,t),(i,h)}\right| \left|J_{(j,s),(j,v)}\right| \left(e'_{(i,h)} M^Q \varepsilon\right)^2 \\
&\quad \times e'_{(j,v)} M^{(Z,Q)} X_1 X_1' M^{(Z,Q)} e_{(j,v)} \\
&\quad + \frac{48}{K_{2,n}} \sum_{\substack{(i,t),(j,s)=1 \\ (i,t)\neq(j,s)}}^{m_n} A^2_{(i,t),(j,s)} \sum_{h=1}^{T_i} \sum_{v=1}^{T_j} \left|J_{(i,t),(i,h)}\right| \left|J_{(j,s),(j,v)}\right| \left(e'_{(i,h)} M^Q \varepsilon\right)^2 e'_{(j,v)} M^Q U U' M^Q e_{(j,v)} \\
&\quad + \frac{48}{K_{2,n}} \sum_{\substack{(i,t),(j,s)=1 \\ (i,t)\neq(j,s)}}^{m_n} A^2_{(i,t),(j,s)} \sum_{h=1}^{T_i} \sum_{v=1}^{T_j} \left|J_{(i,t),(i,h)}\right| \left|J_{(j,s),(j,v)}\right| \left(e'_{(i,h)} M^Q \varepsilon\right)^2 e'_{(j,v)} P^{Z^\perp} U U' P^{Z^\perp} e_{(j,v)} \\
&= \mathcal{A}^*_{1,2,1} + \mathcal{A}^*_{1,2,2} + \mathcal{A}^*_{1,2,3}, \ \ (say),
\end{aligned}$$

where $X_1 = \Gamma D_\mu/\sqrt{n} + F D_\kappa/\sqrt{n}$, $M^{(Z,Q)} = M^Q - P^{Z^\perp}$, $P^{Z^\perp} = M^Q Z \left(Z' M^Q Z\right)^{-1} Z' M^Q$, $M^Q = I_{m_n} - Q\left(Q'Q\right)^{-1} Q'$, and $F = \left(f\left(W_{1,(1,1)}\right), .., f\left(W_{1,(1,T_1)}\right), ..., f\left(W_{1,(n,1)}\right), .., f\left(W_{1,(n,T_n)}\right)\right)'$. By applying Assumptions 3(i), 4(i), 5, 7(ii), and 7(iii) and part (a) of Lemma S2-1, it is straightforward to show that $\sum_{(i,t),(j,s)=1,(i,t)\neq(j,s)}^{m_n} A^2_{(i,t),(j,s)}$ $\times \sum_{v=1}^{T_j} e'_{(j,v)} M^{(Z,Q)} X_1 X_1' M^{(Z,Q)} e_{(j,v)} = O_{a.s.}\left(K_{2,n} \max\left\{\left(\mu_n^{\max}\right)^2 K_{2,n}^{-2\varrho_\gamma}, \left(\kappa_n^{\max}\right)^2 K_{1,n}^{-2\varrho_f}\right\}\right)$[4].

[4]Details are available from the authors upon request.

Making use of this result as well as Assumptions 1, 2(i), 5, and 6; we then obtain

$$\begin{aligned}
0 &\leq E\left[\mathcal{A}^*_{1,2,1}|\mathcal{F}^W_n\right] \\
&\leq \frac{24}{K_{2,n}} \sum_{\substack{(i,t),(j,s)=1 \\ (i,t)\neq(j,s)}}^{m_n} A^2_{(i,t),(j,s)} \sum_{h=1}^{T_i}\sum_{v=1}^{T_j} \left|J_{(i,t),(i,h)}\right| \left|J_{(j,s),(j,v)}\right| \\
&\quad \times e'_{(j,v)} M^{(Z,Q)} X_1 X_1' M^{(Z,Q)} e_{(j,v)} e'_{(i,h)} M^Q E\left[\varepsilon\varepsilon'|\mathcal{F}^W_n\right] M^Q e_{(i,h)} \\
&\leq 24\overline{T}\left(\max_{1\leq(i,t),(j,s)\leq m_n}\left|J_{(i,t),(j,s)}\right|\right)^2 \left(\max_{1\leq(i,t)\leq m_n}\left|M^Q_{(i,t),(i,t)}\right|\right)\left(\max_{1\leq(i,t)\leq m_n} E\left[\varepsilon^2_{(i,t)}|\mathcal{F}^W_n\right]\right) \\
&\quad \times \frac{1}{K_{2,n}} \sum_{\substack{(i,t),(j,s)=1 \\ (i,t)\neq(j,s)}}^{m_n} A^2_{(i,t),(j,s)} \sum_{v=1}^{T_j} e'_{(j,v)} M^{(Z,Q)} X_1 X_1' M^{(Z,Q)} e_{(j,v)} \\
&= O_{a.s.}\left(\max\left\{\frac{(\mu^{\max}_n)^2}{K^{2\varrho_\gamma}_{2,n}}, \frac{(\kappa^{\max}_n)^2}{K^{2\varrho_f}_{1,n}}\right\}\right) = O_{a.s.}(1) \text{ (given Assumption 7(ii) and (iii)).}
\end{aligned}$$

Moreover, applying part (a) of Lemma S2-1 and Assumptions 2(i), 5, and 6; we get

$$\begin{aligned}
0 &\leq E\left[\mathcal{A}^*_{1,2,2}|\mathcal{F}^W_n\right] \\
&\leq \frac{48}{K_{2,n}} \sum_{\substack{(i,t),(j,s)=1 \\ (i,t)\neq(j,s)}}^{m_n} A^2_{(i,t),(j,s)} \sum_{h=1}^{T_i}\sum_{v=1}^{T_j} \left|J_{(i,t),(i,h)}\right| \left|J_{(j,s),(j,v)}\right| \sum_{g=1}^{T_i}\sum_{r=1}^{T_i} \left|M^Q_{(i,h),(i,g)}\right| \left|M^Q_{(i,h),(i,r)}\right| \\
&\quad \times \sum_{q=1}^{T_j}\sum_{c=1}^{T_j} \left|M^Q_{(j,v),(j,q)}\right| \left|M^Q_{(j,v),(j,c)}\right| E\left[\left|\varepsilon_{(i,g)}\varepsilon_{(i,r)} U'_{(j,q)} U_{(j,c)}\right| |\mathcal{F}^W_n\right] \\
&\leq 48\overline{T}^6\left(\max_{1\leq(i,t),(j,s)\leq m_n}\left|J_{(i,t),(j,s)}\right|\right)^2 \left(\max_{1\leq(i,t),(j,s)\leq m_n}\left|M^Q_{(i,t),(j,s)}\right|\right)^4 \left(\max_{1\leq(i,t)\leq m_n} E\left[\varepsilon^4_{(i,t)}|\mathcal{F}^W_n\right]\right)^{1/2} \\
&\quad \times \left(\max_{1\leq(i,t)\leq m_n} E\left[\left\|U_{(i,t)}\right\|_2^4 |\mathcal{F}^W_n\right]\right)^{1/2} \frac{1}{K_{2,n}} \sum_{\substack{(i,t),(j,s)=1 \\ (i,t)\neq(j,s)}}^{m_n} A^2_{(i,t),(j,s)} \\
&= O_{a.s.}(1).
\end{aligned}$$

In addition, for $E\left[\mathcal{A}^*_{1,2,3}|\mathcal{F}^W_n\right]$, we have, by straightforward calculations using Assumption 1 and the triangle inequality,

$$\begin{aligned}
0 &\leq E\left[\mathcal{A}^*_{1,2,3}|\mathcal{F}^W_n\right] \\
&\leq \frac{48}{K_{2,n}} \sum_{\substack{(i,t),(j,s)=1 \\ (i,t)\neq(j,s)}}^{m_n} A^2_{(i,t),(j,s)} \left\{ \sum_{h=1}^{T_i} \sum_{v=1}^{T_j} \left|J_{(i,t),(i,h)}\right| \left|J_{(j,s),(j,v)}\right| \right. \\
&\qquad \left. \times \sum_{g=1}^{T_i} \left(M^Q_{(i,h),(i,g)}\right)^2 \left(P^{Z^\perp}_{(j,v),(i,g)}\right)^2 E\left[\varepsilon^2_{(i,g)} U'_{(i,g)} U_{(i,g)}|\mathcal{F}^W_n\right] \right\} \\
&\quad + \frac{48}{K_{2,n}} \sum_{\substack{(i,t),(j,s)=1 \\ (i,t)\neq(j,s)}}^{m_n} A^2_{(i,t),(j,s)} \left\{ \sum_{h=1}^{T_i} \sum_{v=1}^{T_j} \left|J_{(i,t),(i,h)}\right| \left|J_{(j,s),(j,v)}\right| \sum_{g=1}^{T_i} \left(M^Q_{(i,h),(i,g)}\right)^2 \right. \\
&\qquad \left. \times \sum_{(k,q)=1}^{m_n} \left(P^{Z^\perp}_{(j,v),(k,q)}\right)^2 E\left[\varepsilon^2_{(i,g)}|\mathcal{F}^W_n\right] E\left[U'_{(k,q)} U_{(k,q)}|\mathcal{F}^W_n\right] \right\} \\
&\quad + \frac{96}{K_{2,n}} \sum_{\substack{(i,t),(j,s)=1 \\ (i,t)\neq(j,s)}}^{m_n} A^2_{(i,t),(j,s)} \left\{ \sum_{h=1}^{T_i} \left|J_{(i,t),(i,h)}\right| \sum_{g=1}^{T_i} \sum_{r=1}^{T_i} \left|M^Q_{(i,h),(i,g)}\right| \left|M^Q_{(i,h),(i,r)}\right| \right. \\
&\qquad \left. \times \sum_{v=1}^{T_j} \left|J_{(j,s),(j,v)}\right| \left|P^{Z^\perp}_{(j,v),(i,g)}\right| \left|P^{Z^\perp}_{(j,v),(i,r)}\right| E\left[\left|\varepsilon_{(i,g)}\varepsilon_{(i,r)} U'_{(i,g)} U_{(i,r)}\right| |\mathcal{F}^W_n\right] \right\} \quad (47)
\end{aligned}$$

Applying the CS inequality to (47) and making use of part (a) of Lemma S2-1 as well as Assumptions 2(i), 5, and 6; we further obtain

$$\begin{aligned}
& E\left[\mathcal{A}^*_{1,2,3}|\mathcal{F}^W_n\right] \\
\leq\ & 48\overline{T}^3 \left(\max_{1\leq(i,t),(j,s)\leq m_n} \left|J_{(i,t),(j,s)}\right|\right)^2 \left(\max_{1\leq(i,t),(j,s)\leq m_n} \left|M^Q_{(i,t),(j,s)}\right|\right)^2 \left(\max_{1\leq(i,t)\leq m_n} E\left[\varepsilon^4_{(i,t)}|\mathcal{F}^W_n\right]\right)^{1/2} \\
& \times \left(\max_{1\leq(i,t)\leq m_n} E\left[\left\|U_{(i,t)}\right\|_2^4 |\mathcal{F}^W_n\right]\right)^{1/2} \left(\max_{1\leq(i,t)\leq m_n} P^{Z^\perp}_{(i,t),(i,t)}\right)^2 \frac{1}{K_{2,n}} \sum_{\substack{(i,t),(j,s)=1 \\ (i,t)\neq(j,s)}}^{m_n} A^2_{(i,t),(j,s)} \\
& + 48\overline{T}^3 \left(\max_{1\leq(i,t),(j,s)\leq m_n} \left|J_{(i,t),(j,s)}\right|\right)^2 \left(\max_{1\leq(i,t),(j,s)\leq m_n} \left|M^Q_{(i,t),(j,s)}\right|\right)^2 \left(\max_{1\leq(i,t)\leq m_n} E\left[\varepsilon^2_{(i,t)}|\mathcal{F}^W_n\right]\right) \\
& \times \left(\max_{1\leq(i,t)\leq m_n} E\left[\left\|U_{(i,t)}\right\|_2^2 |\mathcal{F}^W_n\right]\right) \left(\max_{1\leq(i,t)\leq m_n} P^{Z^\perp}_{(i,t),(i,t)}\right) \frac{1}{K_{2,n}} \sum_{\substack{(i,t),(j,s)=1 \\ (i,t)\neq(j,s)}}^{m_n} A^2_{(i,t),(j,s)}
\end{aligned}$$

$$+96\overline{T}^4\left(\max_{1\leq(i,t),(j,s)\leq m_n}\left|J_{(i,t),(j,s)}\right|\right)^2\left(\max_{1\leq(i,t),(j,s)\leq m_n}\left|M^Q_{(i,t),(j,s)}\right|\right)^2\left(\max_{1\leq(i,t)\leq m_n}E\left[\varepsilon^4_{(i,t)}|\mathcal{F}^W_n\right]\right)^{1/2}$$
$$\times\left(\max_{1\leq(i,t)\leq m_n}E\left[\left\|U_{(i,t)}\right\|_2^4|\mathcal{F}^W_n\right]\right)^{1/2}\left(\max_{1\leq(i,t)\leq m_n}P^{Z^\perp}_{(i,t),(i,t)}\right)^2\frac{1}{K_{2,n}}\sum_{\substack{(i,t),(j,s)=1\\(i,t)\neq(j,s)}}^{m_n}A^2_{(i,t),(j,s)}$$
$$=\quad O_{a.s.}\left(\frac{K_n^2}{n^2}\right)+O_{a.s.}\left(\frac{K_n}{n}\right)+O_{a.s.}\left(\frac{K_n^2}{n^2}\right)=O_{a.s.}\left(\frac{K_n}{n}\right)=o_{a.s.}\left(1\right).$$

These calculations imply that $0\leq E\left[\mathcal{A}^*_{1,2}|\mathcal{F}^W_n\right]\leq E\left[\mathcal{A}^*_{1,2,1}|\mathcal{F}^W_n\right]+E\left[\mathcal{A}^*_{1,2,2}|\mathcal{F}^W_n\right]+E\left[\mathcal{A}^*_{1,2,3}|\mathcal{F}^W_n\right]$ $=O_{a.s.}\left(1\right)+O_{a.s.}\left(1\right)+o_{a.s.}\left(1\right)=O_{a.s.}\left(1\right)$. It follows by the law of iterated expectations and Theorem 16.1 of Billingsley (1995) that there exists a constant $\overline{C}<\infty$ such that for all n sufficiently large $E\left[\mathcal{A}^*_{1,2}\right]=E_{W_n}\left(E\left[\mathcal{A}^*_{1,2,}|\mathcal{F}^W_n\right]\right)\leq\overline{C}$. Markov's inequality then implies that

$$\mathcal{A}^*_{1,2}=O_p\left(1\right). \tag{48}$$

from which we further deduce that

$$\mathcal{A}_{1,2}\leq\left\|\widehat{\delta}_n-\delta_0\right\|_2^2\mathcal{A}^*_{1,2}=o_p\left(1\right)O_p\left(1\right)=o_p\left(1\right). \tag{49}$$

Turning our attention to $\mathcal{A}_{1,3}$, observe that

$$\begin{aligned}
\mathcal{A}_{1,3} &= \frac{12}{K_{2,n}}\sum_{\substack{(i,t),(j,s)=1\\(i,t)\neq(j,s)}}^{m_n}A^2_{(i,t),(j,s)}\sum_{h=1}^{T_i}\sum_{v=1}^{T_i}\left|J_{(i,t),(i,h)}\right|\left|J_{(j,s),(j,v)}\right|\left(e'_{(i,h)}M^Q\varepsilon\right)^2\left[e'_{(j,v)}M^{(Z,Q)}\varphi_n\right]^2\\
&= \frac{12}{K_{2,n}}\sum_{\substack{(i,t),(j,s)=1\\(i,t)\neq(j,s)}}^{m_n}A^2_{(i,t),(j,s)}\sum_{h=1}^{T_i}\sum_{v=1}^{T_i}\left|J_{(i,t),(i,h)}\right|\left|J_{(j,s),(j,v)}\right|\left(e'_{(i,h)}M^Q\varepsilon\right)^2\\
&\qquad\times\left[\frac{\tau_n}{\sqrt{n}}e'_{(j,v)}M^{(Z,Q)}\left(g-Z_1\theta^{K_{1,n}}\right)\right]^2\\
&\leq \frac{m_n\tau_n^2}{n}\left\|g\left(\cdot\right)-\theta^{K_{1,n}\prime}Z_1\left(\cdot\right)\right\|_\infty^2\mathcal{A}^*_{1,3}
\end{aligned}$$

where $\mathcal{A}^*_{1,3}=12K^{-1}_{2,n}\sum_{(i,t),(j,s)=1,(i,t)\neq(j,s)}^{m_n}A^2_{(i,t),(j,s)}\sum_{h=1}^{T_i}\sum_{v=1}^{T_j}\left|J_{(i,t),(i,h)}\right|\left|J_{(j,s),(j,v)}\right|\left(e'_{(i,h)}M^Q\varepsilon\right)^2$ and where $g=\left(g\left(W_{1,(1,1)}\right),..,g\left(W_{1,(1,T_1)}\right),....,g\left(W_{1,(n,1)}\right),..,g\left(W_{1,(n,T_n)}\right)\right)'$. Next, note that, by Lemma S2-1(a) and Assumptions 1, 2(i), 5, and 6,

$$
\begin{aligned}
E\left[\mathcal{A}_{1,3}^{*}|\mathcal{F}_n^W\right] &= \frac{12}{K_{2,n}} \sum_{\substack{(i,t),(j,s)=1 \\ (i,t)\neq(j,s)}}^{m_n} A_{(i,t),(j,s)}^2 \sum_{h=1}^{T_i}\sum_{v=1}^{T_i} \left|J_{(i,t),(i,h)}\right| \left|J_{(j,s),(j,v)}\right| E\left[\left(e_{(i,h)}' M^Q \varepsilon\right)^2 |\mathcal{F}_n^W\right] \\
&= \frac{12}{K_{2,n}} \sum_{\substack{(i,t),(j,s)=1 \\ (i,t)\neq(j,s)}}^{m_n} A_{(i,t),(j,s)}^2 \sum_{h=1}^{T_i}\sum_{v=1}^{T_j} \left|J_{(i,t),(i,h)}\right| \left|J_{(j,s),(j,v)}\right| e_{(i,h)}' M^Q E\left[\varepsilon\varepsilon'|\mathcal{F}_n^W\right] M^Q e_{(i,h)} \\
&\leq 12\overline{T}^2 \left(\max_{1\leq(i,t),(j,s)\leq m_n} \left|J_{(i,t),(j,s)}\right|\right)^2 \left(\max_{1\leq(i,t)\leq m_n} M_{(i,t),(i,t)}^Q\right) \\
&\quad \times \left(\max_{1\leq(i,t)\leq m_n} E\left[\varepsilon_{(i,t)}^2|\mathcal{F}_n^W\right]\right) \frac{1}{K_{2,n}} \sum_{\substack{(i,t),(j,s)=1 \\ (i,t)\neq(j,s)}}^{m_n} A_{(i,t),(j,s)}^2 \\
&= O_{a.s.}(1) \qquad (50)
\end{aligned}
$$

It follows by the law of iterated expectations and Theorem 16.1 of Billingsley (1995) that there exists a constant $\overline{C} < \infty$ such that for all n sufficiently large $E\left[\mathcal{A}_{1,3}^{*}\right] = E_{W_n}\left(E\left[\mathcal{A}_{1,3}^{*}|\mathcal{F}_n^W\right]\right) \leq \overline{C}$. Markov's inequality then implies that $\mathcal{A}_{1,3}^{*} = O_p(1)$, from which we further deduce using Assumption 7(i) that

$$
\mathcal{A}_{1,3} \leq \frac{m_n}{n}\tau_n^2 \left\|g(\cdot) - \theta^{K_{1,n}\prime} Z_1(\cdot)\right\|_\infty^2 \mathcal{A}_{1,3}^{*} = O_p\left(\frac{\tau_n^2}{K_{1,n}^{2\varrho_g}}\right) = o_p(1). \qquad (51)
$$

Putting things together, we note that the results given in expressions (46), (49), and (51) imply that

$$
\mathcal{A}_1 = \mathcal{A}_{1,1} + \mathcal{A}_{1,2} + \mathcal{A}_{1,3} = O_p\left(\frac{K_n}{n}\right) + o_p(1) + o_p(1) = o_p(1).
$$

By a similar method of proof, we can also show that $\mathcal{A}_k = o_p(1)$ for $k = 2, ..., 6$. For the sake of brevity, we will not provide detailed proofs for these other terms. Detailed argument for these other terms can be obtained from the authors upon request. It then follows from equation (44) that

$$
\begin{aligned}
\mathcal{A} &= \frac{S_3}{K_{2,n}} - \frac{1}{K_{2,n}} \sum_{\substack{(i,t),(j,s)=1 \\ (i,t)\neq(j,s)}}^{m_n} A_{(i,t),(j,s)}^2 \varepsilon_{(i,t)}^2 \varepsilon_{(j,s)}^2 \\
&= \mathcal{A}_1 + \mathcal{A}_2 + \mathcal{A}_3 + \mathcal{A}_4 + \mathcal{A}_5 + \mathcal{A}_6 = o_p(1). \qquad (52)
\end{aligned}
$$

Now, for the term $\mathfrak{A} = K_{2,n}^{-1} \sum_{(i,t),(j,s)=1,(i,t)\neq(j,s)}^{m_n} A_{(i,t),(j,s)}^2 \left(\varepsilon_{(i,t)}^2 \varepsilon_{(j,s)}^2 - \sigma_{(i,t)}^2 \sigma_{(j,s)}^2\right)$, it can be shown by straightforward calculation and by using Assumptions 1, 2(i), 5 and 6 and Lemma S2-1(a) that $E\left[\mathfrak{A}^2|\mathcal{F}_n^W\right]=O_{a.s.}\left(n^{-1}\right)$.[5] It then follows by application of the law of iterated expectations, Theorem 16.1 of Billingsley (1995), and the Markov's inequality that

$$\mathfrak{A}=O_p\left(\frac{1}{\sqrt{n}}\right). \tag{53}$$

Combining (52) and (53), we get

$$\frac{S_3}{K_{2,n}}-\frac{1}{K_{2,n}}\sum_{\substack{(i,t),(j,s)=1\\(i,t)\neq(j,s)}}^{m_n}A_{(i,t),(j,s)}^2\sigma_{(i,t)}^2\sigma_{(j,s)}^2=\mathcal{A}+\mathfrak{A}=o_p(1)+O_p\left(\frac{1}{\sqrt{n}}\right)=o_p(1),$$

which shows part (b).

The proofs for parts (c) and (d) are similar to that of part (b). Hence, for the sake of brevity, we will not provide detailed proofs of these parts here. Detailed proofs of these parts can be obtained from the authors upon request.

Turning our attention to part (e), note first that, we can write

$$\begin{aligned}\widehat{\rho}_n-\rho &= \frac{X'M^{(Z,Q)}\left(y-X\widehat{\delta}_n\right)/n}{\left(y-X\widehat{\delta}_n\right)'M^{(Z,Q)}\left(y-X\widehat{\delta}_n\right)/n}-\rho\\ &= \frac{X'M^{(Z,Q)}\left(y-X\widehat{\delta}_n\right)/n-U'M^{(Z_1,Q)}\varepsilon/n+U'M^{(Z_1,Q)}\varepsilon/n}{\left(y-X\widehat{\delta}_n\right)'M^{(Z,Q)}\left(y-X\widehat{\delta}_n\right)/n-\varepsilon'M^{Q}\varepsilon/n+\varepsilon'M^{Q}\varepsilon/n}-\rho\end{aligned}$$

where $\rho=\lim_{n\to\infty}\rho_n=\lim_{n\to\infty}\left(E\left[U'M^Q\varepsilon\right]/n\right)/\left(E\left[\varepsilon'M^Q\varepsilon\right]/n\right)$. By straightforward asymptotic analysis[6], we can show that

$$\begin{aligned}\frac{X'M^{(Z,Q)}\left(y-X\widehat{\delta}_n\right)}{n}-\frac{U'M^{(Z_1,Q)}\varepsilon}{n} &= o_p(1),\\ \frac{\left(y-X\widehat{\delta}_n\right)'M^{(Z,Q)}\left(y-X\widehat{\delta}_n\right)}{n}-\frac{\varepsilon'M^{(Z_1,Q)}\varepsilon}{n} &= o_p(1),\\ \frac{\varepsilon'M^{(Z_1,Q)}\varepsilon}{n}-\frac{\varepsilon'M^{Q}\varepsilon}{n} &= o_p(1),\end{aligned}$$

$$\frac{U'M^{(Z_1,Q)}\varepsilon}{\varepsilon'M^{Q}\varepsilon}-\rho=O_p\left(\max\left\{\frac{1}{\sqrt{n}},\frac{K_{1,n}}{n}\right\}\right)=o_p(1).$$

[5] For the sake of brevity, we have not provided the details of all the calculations here. A detailed argument can be obtained from the authors upon request.

[6] Further details are available from the authors upon request.

Next, note that, under Assumption 6(i), $T_i \geq 3$ for all i, so that $\frac{T_i-1}{T_i} \geq \frac{2}{3}$ for all i. Hence, by Assumption 2(ii), there exists a positive constant $\underline{C}$ such that $E\left[\varepsilon_{(i,t)}^2 | \mathcal{F}_n^W\right] \geq \underline{C} > 0$ $a.s.$, so that

$$\begin{aligned}
\frac{E\left[\varepsilon' M^Q \varepsilon\right]}{n} &= \frac{1}{n} E_{W_n}\left\{\sum_{i=1}^{n}\sum_{t=1}^{T_i}\left[\left(\frac{T_i-1}{T_i}\right)\right] E\left[\varepsilon_{(i,t)}^2 | \mathcal{F}_n^W\right]\right\} \\
&\geq \frac{2}{3n} E_{W_n}\left\{\sum_{i=1}^{n}\sum_{t=1}^{T_i} E\left[\varepsilon_{(i,t)}^2 | \mathcal{F}_n^W\right]\right\} \\
&\geq \frac{2}{3n} E_{W_n}\left[\sum_{i=1}^{n}\sum_{t=1}^{T_i} \underline{C}\right] = \frac{2}{3}\frac{m_n}{n}\underline{C} \quad \left(\text{since } m_n = \sum_{i=1}^{n} T_i\right) \\
&\geq \frac{2}{3}\underline{C} > 0
\end{aligned}$$

for all n sufficiently large. It follows from these results that $\widehat{\rho}_n - \rho = \left[U' M^{(Z_1,Q)}\varepsilon / \left(\varepsilon' M^Q \varepsilon\right)\right] - \rho + o_p(1) = o_p(1)$ and $\|\widehat{\rho}_n\|_2 \leq \|\widehat{\rho}_n - \rho\|_2 + \|\rho\|_2 = O_p(1)$.

Now, let $a \in \mathbb{R}^d$ such that $\|a\|_2 = 1$; and note that, by applying the CS inequality, we have

$$\begin{aligned}
\left|a' D_\mu^{-1} \widehat{\rho}_n\right| &\leq \frac{1}{\left(\mu_n^{\min}\right)} \|\widehat{\rho}_n\|_2 = O_p\left(\frac{1}{\left(\mu_n^{\min}\right)}\right), & (54) \\
\left|a' D_\mu^{-1}\left(\widehat{\rho}_n - \rho\right)\right| &\leq \frac{1}{\left(\mu_n^{\min}\right)} \|\widehat{\rho}_n - \rho\|_2 = o_p\left(\frac{1}{\left(\mu_n^{\min}\right)}\right), & (55)
\end{aligned}$$

Since the argument above holds for any $a \in \mathbb{R}^d$ such that $\|a\|_2 = 1$, we further deduce that $D_\mu^{-1}\widehat{\rho}_n = O_p\left(\left(\mu_n^{\min}\right)^{-1}\right)$ and $D_\mu^{-1}\left(\widehat{\rho}_n - \rho\right) = o_p\left(\left(\mu_n^{\min}\right)^{-1}\right)$, which shows part (e).

Part (f) can be shown by applying the results of parts (b), (c), (d), and (e) of this lemma as well as part (a) of Lemma S2-1 and Assumptions 2(i) and 3(ii). Part (g), on the other hand, can be proved by applying the results of parts (b), (d), and (e) of this lemma. For the sake of brevity, we will not provide detailed proofs of these parts here, but proofs of these parts can be obtained from the authors upon request. □

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.

[3] Hall, P. and C. Heyde (1980). *Martingale Limit Theory and Its Applications.* New York: Academic Press.