ONLINE APPENDIX: MULTIVARIATE FORECAST EVALUATION AND RATIONALITY TESTING

IVANA KOMUNJER AND MICHAEL T. OWYANG

This online appendix consists of three parts. Section 1 of the appendix contains additional information on the forecast data used in the empirical application (Section 6 of the paper). Section 2 of the appendix gives detailed proofs of Theorems 2 and 3 stated in the main text (Section 4 of the paper). Finally, the last part of the appendix contains additional panels of Table 1 discussed in the Monte Carlo experiment (Section 5.1 of the paper).

1. Information on the BCEI Forecast Data

In the empirical section of the paper, we use individual forecaster's forecasts from the Blue Chip Economic Indicators (BCEI). The data is proprietary. The survey reports monthly updates of forecasts from individual forecasters starting in 1976:08. Prior to 1984, firms reported current-year forecasts for the first five or six months of the year. In later months, they reported next-year forecasts. Starting in 1984, both current- and next-year forecasts were reported each month.

We use forecasts for three variables: output, inflation, and a short-term interest rate. The sample of output forecasts is split between GNP (1976:08 through 1991:12) and GDP (1992:01 through 2004:12). The BCEI began collecting CPI inflation forecasts in 1979:01 through the end of our sample in 2004:12. The short-term interest rate forecasts are split between the 3-month commercial paper (1976:08 through 1980:06), the 6-month commercial paper (1980:07 through 1981:12), and the 3-month T-bill (1982:01 through 2004:12) rates. For output and inflation, the target variable is the rate of change between the average of the levels for that year. This method is described by the BCEI in their monthly newsletter.

2. Supplemental Proofs

2.1. **Notation.** We first recall the notation.

For any real function $f : \mathbb{R}^n \to \mathbb{R}$ that is continuously differentiable to order $R \geqslant 2$ on $\mathbb{R}^n$, we let $\nabla_{\mathbf{u}} f(\mathbf{u})$ denote the gradient of $f(\cdot)$ with respect to $\mathbf{u}$, $\nabla_{\mathbf{u}} f(\mathbf{u}) \equiv (\partial f(\mathbf{u})/\partial u_i, \ldots, \partial f(\mathbf{u})/\partial u_n)'$, and use $\Delta_{\mathbf{uu}} f(\mathbf{u})$ to denote its Hessian matrix, $\Delta_{\mathbf{uu}} f(\mathbf{u}) \equiv (\partial^2 f(\mathbf{u})/\partial u_i \partial u_j)_{1\leqslant i,j\leqslant n}$.

For any scalar u, $u \in \mathbb{R}$, we let $\mathbb{I} : \mathbb{R} \to [0,1]$ be the indicator (or Heaviside) function, i.e., $\mathbb{I}(u) = 0$ if $u < 0$, $\mathbb{I}(u) = 1$ if $u > 0$, and $\mathbb{I}(0) = \frac{1}{2}$ (**?**). Similarly, we use $\operatorname{sgn} : \mathbb{R} \to \{-1, 0, 1\}$ to denote the sign function: $\operatorname{sgn}(u) = \mathbb{I}(u) - \mathbb{I}(-u) = 2\mathbb{I}(u) - 1$, and let $\delta : \mathbb{R} \to \mathbb{R}$ be the Dirac delta function. Note that the Heaviside function is the indefinite integral of the Dirac function, i.e., $\mathbb{I}(u) = \int_a^u d\delta$, where a is an arbitrary (possibly infinite) negative constant, $a \leqslant 0$.

For any n-vector $\mathbf{u}$, $\mathbf{u} = (u_1, \ldots, u_n)' \in \mathbb{R}^n$, we denote by $\|\mathbf{u}\|_p$ its l_p-norm, i.e., $\|\mathbf{u}\|_p = (|u_1|^p + \ldots + |u_n|^p)^{1/p}$ for $1 \leqslant p < \infty$, and $\|\mathbf{u}\|_\infty = \max_{1\leqslant i\leqslant n}(|u_i|)$. Similarly, for any $m \times n$-matrix $\mathbf{A} = (a_{ij})_{1\leqslant i\leqslant m, 1\leqslant j\leqslant n}$, we let $\|\mathbf{A}\|_\infty = \max_{1\leqslant i,j\leqslant n}(|a_{ij}|)$.

Hereafter, $\boldsymbol{\nu}_p(\mathbf{u})$, $\mathbf{V}_p(\mathbf{u})$ and $\mathbf{W}_p(\mathbf{u})$ are an n-vector and two $n \times n$-diagonal matrices defined as:

$$\begin{aligned}
\boldsymbol{\nu}_p(\mathbf{u}) &\equiv (\operatorname{sgn}(u_1)|u_1|^{p-1}, \ldots, \operatorname{sgn}(u_n)|u_n|^{p-1})' \\
\mathbf{V}_p(\mathbf{u}) &\equiv \operatorname{diag}(\delta(u_1)|u_1|^{p-1}, \ldots, \delta(u_n)|u_n|^{p-1}) \\
\mathbf{W}_p(\mathbf{u}) &\equiv \operatorname{diag}(|u_1|^{p-2}, \ldots, |u_n|^{p-2}),
\end{aligned}$$

respectively. Then, we have that:

$$\nabla_{\mathbf{u}} \|\mathbf{u}\|_p = \|\mathbf{u}\|_p^{1-p} \, \boldsymbol{\nu}_p(\mathbf{u})$$

and

$$\Delta_{\mathbf{uu}} \|\mathbf{u}\|_p = \|\mathbf{u}\|_p^{1-p} \left\{ 2\mathbf{V}_p(\mathbf{u}) + (p-1) \left[\mathbf{W}_p(\mathbf{u}) - \|\mathbf{u}\|_p^{-p} \, \boldsymbol{\nu}_p(\mathbf{u}) \boldsymbol{\nu}_p'(\mathbf{u}) \right] \right\},$$

which we shall often be using in what follows.

2.2. **Proofs of Theorems 2 and 3.**

Theorem 2. *Let Assumptions A1 through A8 hold. Then, given* p, $1 \leqslant p < \infty$, *we have* $\hat{\boldsymbol{\tau}}_P \xrightarrow{p} \boldsymbol{\tau}_0$ *as* $(R, P) \to \infty$.

Proof of Theorem 2. The minimizer $\boldsymbol{\tau}_0$ of $Q(\boldsymbol{\tau})$ can be written as:

$$\boldsymbol{\tau}_0 = -\{E[\mathbf{B}(p, \mathbf{e}^*_{t+1}, \mathbf{x}_t)]'\mathbf{S}^{-1}E[\mathbf{B}(p, \mathbf{e}^*_{t+1}, \mathbf{x}_t)]\}^{-1}E[\mathbf{B}(p, \mathbf{e}^*_{t+1}, \mathbf{x}_t)]'\mathbf{S}^{-1}E[\mathbf{a}(p, \mathbf{e}^*_{t+1}, \mathbf{x}_t)].$$

On the other hand, from Equation (4) we have $\hat{\boldsymbol{\tau}}_P \equiv -[\hat{\mathbf{B}}'_P\hat{\mathbf{S}}^{-1}\hat{\mathbf{B}}_P]^{-1}\hat{\mathbf{B}}'_P\hat{\mathbf{S}}^{-1}\hat{\mathbf{a}}_P$ with the $nd \times 1$ vector

$$\hat{\mathbf{a}}_P \equiv P^{-1}\sum_{t=R}^{T} p(\boldsymbol{\nu}_p(\hat{\mathbf{e}}_{t+1}) \otimes \mathbf{x}_t) \tag{S-1}$$

and the $nd \times n$ matrix

$$\hat{\mathbf{B}}_P \equiv P^{-1}\sum_{t=R}^{T} \|\hat{\mathbf{e}}_{t+1}\|_p^{p-1} (\mathbf{I}_n \otimes \mathbf{x}_t) + (p-1) \|\hat{\mathbf{e}}_{t+1}\|_p^{-1} (\boldsymbol{\nu}_p(\hat{\mathbf{e}}_{t+1}) \otimes \mathbf{x}_t)\hat{\mathbf{e}}'_{t+1}. \tag{S-2}$$

To show $\hat{\boldsymbol{\tau}}_P \xrightarrow{p} \boldsymbol{\tau}_0$, it is sufficient to show that *(i)* $\hat{\mathbf{a}}_P - E[\mathbf{a}(p, \mathbf{e}^*_{t+1}, \mathbf{x}_t)] \xrightarrow{p} \mathbf{0}$ and *(ii)* $\hat{\mathbf{B}}_P - E[\mathbf{B}(p, \mathbf{e}^*_{t+1}, \mathbf{x}_t)] \xrightarrow{p} \mathbf{0}$. Then, by using Lemma 5, the consistency of $\hat{\mathbf{S}}$, $\hat{\mathbf{S}} \xrightarrow{p} \mathbf{S}$, the positive definiteness of $\mathbf{S}$ (and thus of $\mathbf{S}^{-1}$) established in Lemma 6, and the continuity of the inverse function (away from zero), we have that $\hat{\boldsymbol{\tau}}_P \xrightarrow{p} \boldsymbol{\tau}_0$. By the triangle inequality we have $\left\|\hat{\mathbf{a}}_P - E[\mathbf{a}(p, \mathbf{e}^*_{t+1}, \mathbf{x}_t)]\right\|_1 \leqslant \|\hat{\mathbf{a}}_P - E[\mathbf{a}(p, \hat{\mathbf{e}}_{t+1}, \mathbf{x}_t)]\|_1 + \left\|E[\mathbf{a}(p, \hat{\mathbf{e}}_{t+1}, \mathbf{x}_t)] - E[\mathbf{a}(p, \mathbf{e}^*_{t+1}, \mathbf{x}_t)]\right\|_1$ and $\left\|\hat{\mathbf{B}}_P - E[\mathbf{B}(p, \mathbf{e}^*_{t+1}, \mathbf{x}_t)]\right\|_\infty \leqslant \left\|\hat{\mathbf{B}}_P - E[\mathbf{B}(p, \hat{\mathbf{e}}_{t+1}, \mathbf{x}_t)]\right\|_\infty + \left\|E[\mathbf{B}(p, \hat{\mathbf{e}}_{t+1}, \mathbf{x}_t)] - E[\mathbf{B}(p, \mathbf{e}^*_{t+1}, \mathbf{x}_t)]\right\|_\infty$. We first show that as $P \to \infty$, $\|\hat{\mathbf{a}}_P - E[\mathbf{a}(p, \hat{\mathbf{e}}_{t+1}, \mathbf{x}_t)]\|_1 \xrightarrow{p} \mathbf{0}$ and $\left\|\hat{\mathbf{B}}_P - E[\mathbf{B}(p, \hat{\mathbf{e}}_{t+1}, \mathbf{x}_t)]\right\|_\infty \xrightarrow{p} \mathbf{0}$ by using a law of large numbers (LLN) for α-mixing sequences [e.g., Corollary 3.48 in **?**]. From Theorems 3.35 and 3.49 **?** measurable functions of strictly stationary and mixing processes are strictly stationary and mixing of the same size. Hence, by A8 we have $\{p(\boldsymbol{\nu}_p(\hat{\mathbf{e}}_{t+1}) \otimes \mathbf{x}_t)\}$ and $\{\|\hat{\mathbf{e}}_{t+1}\|_p^{p-1} (\mathbf{I}_n \otimes \mathbf{x}_t) + (p-1) \|\hat{\mathbf{e}}_{t+1}\|_p^{-1} (\boldsymbol{\nu}_p(\hat{\mathbf{e}}_{t+1}) \otimes \mathbf{x}_t)\hat{\mathbf{e}}'_{t+1}\}$ strictly stationary and α-mixing of size $-r/(r-2)$ with $r>2$. Now let $\delta=\varepsilon/2>0$; we have

$$\begin{aligned} E[\|(\boldsymbol{\nu}_p(\hat{\mathbf{e}}_{t+1})\otimes\mathbf{x}_t)\|_1^{r+\delta}] &\leqslant nE[\|\hat{\mathbf{e}}_{t+1}\|_1^{(p-1)(r+\delta)}\|\mathbf{x}_t\|_1^{r+\delta}] \\ &\leqslant n\left\{E[\|\hat{\mathbf{e}}_{t+1}\|_1^{(p-1)(2r+2\delta)}]E[\|\mathbf{x}_t\|_1^{2r+2\delta}]\right\}^{1/2} \\ &\leqslant n\left\{\Delta_1\Delta_2\right\}^{1/2}<\infty, \end{aligned} \tag{S-3}$$

where the second inequality follows by Cauchy-Schwartz inequality and the third uses assumption A8. Hence, $\hat{\mathbf{a}}_P$ in Equation (S-1) satisfies the LLN and $\|\hat{\mathbf{a}}_P-E[\mathbf{a}(p,\hat{\mathbf{e}}_{t+1},\mathbf{x}_t)]\|_1\overset{p}{\to}\mathbf{0}$ as $P\to\infty$. Similarly, we have

$$\begin{aligned} E[\|\hat{\mathbf{e}}_{t+1}\|_p^{(p-1)(r+\delta)}\|(\mathbf{I}_n\otimes\mathbf{x}_t)\|_\infty^{r+\delta}] &\leqslant E[\|\hat{\mathbf{e}}_{t+1}\|_p^{(p-1)(r+\delta)}\|\mathbf{x}_t\|_1^{r+\delta}] \\ &\leqslant c^{(p-1)(r+\delta)}\left\{E[\|\hat{\mathbf{e}}_{t+1}\|_1^{(p-1)(2r+2\delta)}]E[\|\mathbf{x}_t\|_1^{2r+2\delta}]\right\}^{1/2} \\ &\leqslant c\left\{\Delta_1\Delta_2\right\}^{1/2}<\infty, \end{aligned} \tag{S-4}$$

where the second inequality uses the norm equivalence and Cauchy-Schwartz inequality, and the third inequality uses Assumption A8. In addition,

$$\begin{aligned} E[\|\hat{\mathbf{e}}_{t+1}\|_p^{-(r+\delta)}\left\|(\boldsymbol{\nu}_p(\hat{\mathbf{e}}_{t+1})\otimes\mathbf{x}_t)\hat{\mathbf{e}}_{t+1}'\right\|_\infty^{r+\delta}] &\leqslant E[\|\hat{\mathbf{e}}_{t+1}\|_p^{-(r+\delta)}\|(\boldsymbol{\nu}_p(\hat{\mathbf{e}}_{t+1})\otimes\mathbf{x}_t)\|_1^{r+\delta}\|\hat{\mathbf{e}}_{t+1}\|_1^{(r+\delta)}] \\ &\leqslant (1/d)^{r+\delta}E[\|(\boldsymbol{\nu}_p(\hat{\mathbf{e}}_{t+1})\otimes\mathbf{x}_t)\|_1^{r+\delta}]<\infty, \end{aligned} \tag{S-5}$$

where the second inequality uses again the norm equivalence and the third follows from Equation (S-3). Combining Equations (S-4) – (S-5) with triangular inequality and the fact that, for any $(a,b)\in\mathbb{R}$, there exists some $n_{r+\delta}>0$ such that $|a+b|^{r+\delta}\leqslant n_{r+\delta}[|a|^{r+\delta}+|b|^{r+\delta}]$, shows that $\hat{\mathbf{B}}_P$ in Equation (S-2) satisfies the LLN and so $\left\|\hat{\mathbf{B}}_P-E[\mathbf{B}(p,\hat{\mathbf{e}}_{t+1},\mathbf{x}_t)]\right\|_\infty\overset{p}{\to}\mathbf{0}$ as $P\to\infty$. Next we need to show that $\left\|E[\mathbf{a}(p,\hat{\mathbf{e}}_{t+1},\mathbf{x}_t)]-E[\mathbf{a}(p,\mathbf{e}_{t+1}^*,\mathbf{x}_t)]\right\|_1\to0$ and $\left\|E[\mathbf{B}(p,\hat{\mathbf{e}}_{t+1},\mathbf{x}_t)]-E[\mathbf{B}(p,\mathbf{e}_{t+1}^*,\mathbf{x}_t)]\right\|_\infty\to0$ as $P\to\infty$. We have

$$\begin{aligned} \left\|E[\mathbf{a}(p,\hat{\mathbf{e}}_{t+1},\mathbf{x}_t)-\mathbf{a}(p,\mathbf{e}_{t+1}^*,\mathbf{x}_t)]\right\|_1 &\leqslant E[\left\|\mathbf{a}(p,\hat{\mathbf{e}}_{t+1},\mathbf{x}_t)-\mathbf{a}(p,\mathbf{e}_{t+1}^*,\mathbf{x}_t)\right\|_1] \\ &=pE\{\left\|[\boldsymbol{\nu}_p(\hat{\mathbf{e}}_{t+1})-\boldsymbol{\nu}_p(\mathbf{e}_{t+1}^*)]\otimes\mathbf{x}_t\right\|_1\} \\ &\leqslant pnE[\left\|\hat{\mathbf{e}}_{t+1}-\mathbf{e}_{t+1}^*\right\|_1^{(p-1)}\|\mathbf{x}_t\|_1] \\ &\leqslant pn\{E[\left\|\hat{\mathbf{e}}_{t+1}-\mathbf{e}_{t+1}^*\right\|_1^{2(p-1)}]E[\|\mathbf{x}_t\|_1^2]\}^{1/2}\to0 \text{ as } t\to\infty, \end{aligned}$$

where the last statement follows by Assumptions A7 and A8. Similarly,

$$
\begin{aligned}
&\left\|E[\mathbf{B}(p,\hat{\mathbf{e}}_{t+1},\mathbf{x}_t)-\mathbf{B}(p,\mathbf{e}^*_{t+1},\mathbf{x}_t)]\right\|_\infty \\
&\leqslant E[\left\|\mathbf{B}(p,\hat{\mathbf{e}}_{t+1},\mathbf{x}_t)-\mathbf{B}(p,\mathbf{e}^*_{t+1},\mathbf{x}_t)\right\|_\infty] \\
&= E\left[\left\|[\|\hat{\mathbf{e}}_{t+1}\|_p^{p-1}-\left\|\mathbf{e}^*_{t+1}\right\|_p^{p-1}](\mathbf{I}_n\otimes\mathbf{x}_t)\right.\right. \\
&\left.\left.+(p-1)[\|\hat{\mathbf{e}}_{t+1}\|_p^{-1}(\boldsymbol{\nu}_p(\hat{\mathbf{e}}_{t+1})\otimes\mathbf{x}_t)\hat{\mathbf{e}}'_{t+1}-\left\|\mathbf{e}^*_{t+1}\right\|_p^{-1}(\boldsymbol{\nu}_p(\mathbf{e}^*_{t+1})\otimes\mathbf{x}_t)\mathbf{e}^{*\prime}_{t+1}]\right\|_\infty\right] \\
&\leqslant E\left[\left|\|\hat{\mathbf{e}}_{t+1}\|_p^{p-1}-\left\|\mathbf{e}^*_{t+1}\right\|_p^{p-1}\right|\|\mathbf{x}_t\|_1\right] \\
&+(p-1)E\left[\|\hat{\mathbf{e}}_{t+1}\|_p^{-1}\left\|(\boldsymbol{\nu}_p(\hat{\mathbf{e}}_{t+1})\otimes\mathbf{x}_t)(\hat{\mathbf{e}}'_{t+1}-\mathbf{e}^{*\prime}_{t+1})\right\|_\infty\right] \\
&+(p-1)E\left[\|\hat{\mathbf{e}}_{t+1}\|_p^{-1}\left\|\{[(\boldsymbol{\nu}_p(\hat{\mathbf{e}}_{t+1})-\boldsymbol{\nu}_p(\mathbf{e}^*_{t+1})]\otimes\mathbf{x}_t\}\mathbf{e}^{*\prime}_{t+1}\right\|_\infty\right] \\
&+(p-1)E\left[\left(\|\hat{\mathbf{e}}_{t+1}\|_p^{-1}-\left\|\mathbf{e}^*_{t+1}\right\|_p^{-1}\right)\left\|(\boldsymbol{\nu}_p(\mathbf{e}^*_{t+1})\otimes\mathbf{x}_t)\mathbf{e}^{*\prime}_{t+1}\right\|_\infty\right]\to 0 \text{ as } t\to\infty.
\end{aligned}
$$

Hence, as $R \to \infty$ we have $\left\|E[\mathbf{a}(p,\hat{\mathbf{e}}_{t+1},\mathbf{x}_t)-\mathbf{a}(p,\mathbf{e}^*_{t+1},\mathbf{x}_t)]\right\|_1 \to 0$ and $\left\|E[\mathbf{B}(p,\hat{\mathbf{e}}_{t+1},\mathbf{x}_t)-\mathbf{B}(p,\mathbf{e}^*_{t+1},\mathbf{x}_t)]\right\|_\infty \to 0$, so $\hat{\boldsymbol{\tau}}_P \xrightarrow{p} \boldsymbol{\tau}_0$ as $(R,P)\to\infty$. □

Theorem 3. *Let Assumptions A1-A3, A4', A5-A6, A7', A8-A10 hold. Then, given* p, $1\leqslant p<\infty$, *we have:* $\sqrt{P}(\hat{\boldsymbol{\tau}}_P-\boldsymbol{\tau}_0)\xrightarrow{d}\mathcal{N}(\mathbf{0},(\mathbf{B}^{*\prime}\mathbf{S}^{-1}\mathbf{B}^*)^{-1})$, *as* $R,P\to\infty$, *where* $\mathbf{S}=E[\mathbf{g}_p(\boldsymbol{\tau}_0;\mathbf{e}^*_{t+1},\mathbf{x}_t)\mathbf{g}_p(\boldsymbol{\tau}_0;\mathbf{e}^*_{t+1},\mathbf{x}_t)']$ *and* $\mathbf{B}^*\equiv E[\left\|\mathbf{e}^*_{t+1}\right\|_p^{p-1}(\mathbf{I}_n\otimes\mathbf{x}_t)+(p-1)\left\|\mathbf{e}^*_{t+1}\right\|_p^{-1}(\boldsymbol{\nu}_p(\mathbf{e}^*_{t+1})\otimes\mathbf{x}_t)\mathbf{e}^{*\prime}_{t+1}]$.

Proof of Theorem 3. To simplify the notation in this proof, let it be understood that $\sum_t$ denotes $\sum_{t=R}^T$ while $\sup_t$ stands for $\sup_{R\leqslant t\leqslant T}$. In order to show that $P^{1/2}(\hat{\boldsymbol{\tau}}_P-\boldsymbol{\tau}_0)$ is asymptotically normal, note that

$$
\begin{aligned}
\sqrt{P}(\hat{\boldsymbol{\tau}}_P-\boldsymbol{\tau}_0) &= -[\hat{\mathbf{B}}'_P\hat{\mathbf{S}}^{-1}\hat{\mathbf{B}}_P]^{-1}\hat{\mathbf{B}}'_P\hat{\mathbf{S}}^{-1}[\sqrt{P}(\hat{\mathbf{a}}_P+\hat{\mathbf{B}}_P\boldsymbol{\tau}_0)] \qquad \text{(S-6)}\\
&= -[\hat{\mathbf{B}}'_P\hat{\mathbf{S}}^{-1}\hat{\mathbf{B}}_P]^{-1}\hat{\mathbf{B}}'_P\hat{\mathbf{S}}^{-1}[\sqrt{P}\hat{\mathbf{m}}^*_P+\sqrt{P}\hat{\mathbf{m}}+\sqrt{P}(\hat{\mathbf{m}}_P-\hat{\mathbf{m}}-\hat{\mathbf{m}}^*_P)],
\end{aligned}
$$

where we have let $\hat{\mathbf{m}}\equiv E[\mathbf{a}(p,\hat{\mathbf{e}}_{t+1},\mathbf{x}_t)]+E[\mathbf{B}(p,\hat{\mathbf{e}}_{t+1},\mathbf{x}_t)]\boldsymbol{\tau}_0$, and

$$
\hat{\mathbf{m}}_P\equiv P^{-1}\sum_t\mathbf{g}_p(\boldsymbol{\tau}_0;\hat{\mathbf{e}}_{t+1},\mathbf{x}_t)=\hat{\mathbf{a}}_P+\hat{\mathbf{B}}_P\boldsymbol{\tau}_0,\ \text{and}\ \hat{\mathbf{m}}^*_P\equiv P^{-1}\sum_t\mathbf{g}_p(\boldsymbol{\tau}_0;\mathbf{e}^*_{t+1},\mathbf{x}_t). \qquad \text{(S-7)}
$$

The idea then is to show that the terms $\sqrt{P}\hat{\mathbf{m}}$ and $\sqrt{P}(\hat{\mathbf{m}}_P - \hat{\mathbf{m}} - \hat{\mathbf{m}}_P^*)$ on the right-hand side of Equation (S-6) are $o_p(\mathbf{1})$. We start by showing that the first term is $o(\mathbf{1})$. Let $\mathbf{m}^* \equiv E[\mathbf{a}(p, \mathbf{e}_{t+1}^*, \mathbf{x}_t)] + E[\mathbf{B}(p, \mathbf{e}_{t+1}^*, \mathbf{x}_t)]\boldsymbol{\tau}_0$. First, we show that $\nabla_{\mathbf{e}} E[\mathbf{g}_p(\boldsymbol{\tau}_0; \bar{\mathbf{e}}_{t+1}, \mathbf{x}_t)] = E[\nabla_{\mathbf{e}} \mathbf{g}_p(\boldsymbol{\tau}_0; \bar{\mathbf{e}}_{t+1}, \mathbf{x}_t)]$ for every $\bar{\mathbf{e}}_{t+1} \equiv c\hat{\mathbf{e}}_{t+1} + (1-c)\mathbf{e}_{t+1}^*$ with $c \in (0,1)$. Differentiating $\nabla_{\mathbf{e}} L_p(\boldsymbol{\tau}_0, \cdot)$ in Equation (9) we get that for any $\mathbf{e} \in \mathbb{R}^n \backslash \mathcal{E}$ ($\mathcal{E} = \{\mathbf{e} \in \mathbb{R}^n : e_i = 0 \text{ for some } i\}$),

$$\begin{aligned}\Delta_{\mathbf{ee}} L_p(\boldsymbol{\tau}_0, \mathbf{e}) &= 2p\mathbf{V}_p(\mathbf{e}) + p(p-1)\mathbf{W}_p(\mathbf{e}) \\ &\quad + (p-1)\left[2\frac{\boldsymbol{\tau}_0 \boldsymbol{\nu}_p'(\mathbf{e})}{\|\mathbf{e}\|_p} + \frac{\boldsymbol{\tau}_0'\mathbf{e}}{\|\mathbf{e}\|_p}\left((p-1)\mathbf{W}_p(\mathbf{e}) - \frac{\boldsymbol{\nu}_p(\mathbf{e})\boldsymbol{\nu}_p'(\mathbf{e})}{\|\mathbf{e}\|_p^p}\right)\right],\end{aligned} \tag{S-8}$$

where we have used the fact that for any $1 \leqslant p < \infty$, $\dfrac{\boldsymbol{\tau}_0'\mathbf{e}}{\|\mathbf{e}\|_p}\mathbf{V}_p(\mathbf{e}) = \mathbf{0}$ for all $\mathbf{e} \in \mathbb{R}^n \backslash \mathcal{E}$. Note that in the univariate case $n = 1$, the Hessian in Equation (S-8) reduces to $\Delta_{ee} L_p(\tau_0, e) = 2\{p\delta(e)|e|^{p-1} + p(p-1)[1 + \tau_0 \operatorname{sgn}(e)]|e|^{p-2}\}$ [see Equation (9) in EKT, p.1121]. Hence

$$\begin{aligned}\|\Delta_{\mathbf{ee}} L_p(\boldsymbol{\tau}_0, \bar{\mathbf{e}}_{t+1})\|_\infty &\leqslant 2p\,\|\mathbf{V}_p(\bar{\mathbf{e}}_{t+1})\| + p(p-1)c_3\,\|\bar{\mathbf{e}}_{t+1}\|_1^{p-2} \\ &\quad + (p-1)\left[2d_3\,\|\bar{\mathbf{e}}_{t+1}\|_1^{p-2} + (p-1)c_3\,\|\bar{\mathbf{e}}_{t+1}\|_1^{p-2} + c_3\,\|\bar{\mathbf{e}}_{t+1}\|_1^{p-2}\right] \\ &= 2p\,\|\mathbf{V}_p(\bar{\mathbf{e}}_{t+1})\| + 2(p-1)(pc_3 + d_3)\,\|\bar{\mathbf{e}}_{t+1}\|_1^{p-2},\end{aligned} \tag{S-9}$$

where we have used the norm equivalences: $c_1\|\bar{\mathbf{e}}_{t+1}\|_1 \leqslant \|\bar{\mathbf{e}}_{t+1}\|_{p-2} \leqslant c_2\|\bar{\mathbf{e}}_{t+1}\|_1$ for some $(c_1, c_2) > 0$ and $c_3 = c_2^{p-1}$ if $p \geq 2$ and $c_3 = c_1^{2-p}$ otherwise and, similarly, $d_1\|\bar{\mathbf{e}}_{t+1}\|_1 \leqslant \|\bar{\mathbf{e}}_{t+1}\|_p \leqslant d_2\|\bar{\mathbf{e}}_{t+1}\|_1$ for some $(d_1, d_2) > 0$ and $d_3 = d_2^{p-1}$ if $p \geq 2$ and $d_3 = d_1^{2-p}$ otherwise. Under A9, we have that $E[\sup_{c\in(0,1)}\left\|c\hat{\mathbf{e}}_{t+1} + (1-c)\mathbf{e}_{t+1}^*\right\|_1^{p-2}] < \infty$. Moreover, under A10, when $p = 1$ we have $E[\|\mathbf{V}_1(\bar{\mathbf{e}}_{t+1})\|_\infty] \leqslant M$ and when $p > 1$ we have $E[\|\mathbf{V}_1(\bar{\mathbf{e}}_{t+1})\|_\infty] = 0$, so the right-hand side of Equation (S-9) is bounded above by a quantity that is integrable; hence, we can apply Lebesgue's dominated convergence theorem to interchange the derivation and integration in $\nabla_{\mathbf{e}} E[\mathbf{g}_p(\boldsymbol{\tau}_0; \bar{\mathbf{e}}_{t+1}, \mathbf{x}_t)] = \nabla_{\mathbf{e}} E[\nabla_{\mathbf{e}} L_p(\boldsymbol{\tau}_0, \bar{\mathbf{e}}_{t+1}) \otimes \mathbf{x}_t] = E[\Delta_{\mathbf{ee}} L_p(\boldsymbol{\tau}_0, \bar{\mathbf{e}}_{t+1}) \otimes \mathbf{x}_t] = E[\nabla_{\mathbf{e}} \mathbf{g}_p(\boldsymbol{\tau}_0; \bar{\mathbf{e}}_{t+1}, \mathbf{x}_t)]$.

Second, we can use a mean value expansion around $\mathbf{e}_{t+1}^*$ that yields $\mathbf{0} = \sqrt{P}\mathbf{m}^* = \sqrt{P}\hat{\mathbf{m}} - E[P^{-1}\sum_t \nabla_{\mathbf{e}} \mathbf{g}_p(\boldsymbol{\tau}_0; \bar{\mathbf{e}}_{t+1}, \mathbf{x}_t)'\sqrt{P}(\hat{\mathbf{e}}_{t+1} - \mathbf{e}_{t+1}^*)]$, where for every t, $R \leqslant t \leqslant T$, we have $\bar{\mathbf{e}}_{t+1} \equiv c\hat{\mathbf{e}}_{t+1} + (1-c)\mathbf{e}_{t+1}^*$ with $c \in (0,1)$. We now show that

$P^{-1/2}\sum_t \nabla_{\mathbf{e}}\mathbf{g}_p(\boldsymbol{\tau}_0;\bar{\mathbf{e}}_{t+1},\mathbf{x}_t)'(\hat{\mathbf{e}}_{t+1}-\mathbf{e}^*_{t+1}) \xrightarrow{p} \mathbf{0}$ as $R\to\infty$ and $P\to\infty$. Consider ε from A7′(i) and note that we have

$$\begin{aligned}&\left\|P^{-1/2}\textstyle\sum_t \nabla_{\mathbf{e}}\mathbf{g}_p(\boldsymbol{\tau}_0;\bar{\mathbf{e}}_{t+1},\mathbf{x}_t)'(\hat{\mathbf{e}}_{t+1}-\mathbf{e}^*_{t+1})\right\|_1\\ &\quad=\left\|P^{-1/2}\textstyle\sum_t R^{-1/2+\varepsilon}\nabla_{\mathbf{e}}\mathbf{g}_p(\boldsymbol{\tau}_0;\bar{\mathbf{e}}_{t+1},\mathbf{x}_t)'R^{1/2-\varepsilon}(\hat{\mathbf{e}}_{t+1}-\mathbf{e}^*_{t+1})\right\|_1\\ &\quad\leqslant \sup_t\left\|R^{1/2-\varepsilon}(\hat{\mathbf{e}}_{t+1}-\mathbf{e}^*_{t+1})\right\|_1 P^{-1/2}\textstyle\sum_t \left\|\nabla_{\mathbf{e}}\mathbf{g}_p(\boldsymbol{\tau}_0;\bar{\mathbf{e}}_{t+1},\mathbf{x}_t)\right\|_\infty R^{-1/2+\varepsilon},\end{aligned}$$

so for any $\eta>0$ and $\delta>0$

$$\begin{aligned}&\lim_{R,P\to\infty}\Pr\left(\left\|P^{-1/2}\textstyle\sum_t \nabla_{\mathbf{e}}\mathbf{g}_p(\boldsymbol{\tau}_0;\bar{\mathbf{e}}_{t+1},\mathbf{x}_t)'(\hat{\mathbf{e}}_{t+1}-\mathbf{e}^*_{t+1})\right\|_1>\eta\right)\\ &\leqslant \lim_{R,P\to\infty}\Pr\left(\left\|P^{-1/2}\textstyle\sum_t \nabla_{\mathbf{e}}\mathbf{g}_p(\boldsymbol{\tau}_0;\bar{\mathbf{e}}_{t+1},\mathbf{x}_t)'(\hat{\mathbf{e}}_{t+1}-\mathbf{e}^*_{t+1})\right\|_1>\eta,\sup_t\left\|R^{1/2-\varepsilon}(\hat{\mathbf{e}}_{t+1}-\mathbf{e}^*_{t+1})\right\|_1\leqslant\delta\right)\\ &+\lim_{R,P\to\infty}\Pr\left(\sup_t\left\|R^{1/2-\varepsilon}(\hat{\mathbf{e}}_{t+1}-\mathbf{e}^*_{t+1})\right\|_1>\delta\right)\\ &\leqslant \lim_{R,P\to\infty}\Pr\left(\left\|P^{-1/2}\textstyle\sum_t \nabla_{\mathbf{e}}\mathbf{g}_p(\boldsymbol{\tau}_0;\bar{\mathbf{e}}_{t+1},\mathbf{x}_t)'(\hat{\mathbf{e}}_{t+1}-\mathbf{e}^*_{t+1})\right\|_1>\eta,\sup_t\left\|R^{1/2-\varepsilon}(\hat{\mathbf{e}}_{t+1}-\mathbf{e}^*_{t+1})\right\|_1\leqslant\delta\right)\\ &\leqslant \lim_{R,P\to\infty}\Pr\left(P^{-1/2}\textstyle\sum_t \left\|\nabla_{\mathbf{e}}\mathbf{g}_p(\boldsymbol{\tau}_0;\bar{\mathbf{e}}_{t+1},\mathbf{x}_t)\right\|_\infty R^{-1/2+\varepsilon}>\dfrac{\eta}{\delta},\sup_t\left\|R^{1/2-\varepsilon}(\hat{\mathbf{e}}_{t+1}-\mathbf{e}^*_{t+1})\right\|_1\leqslant\delta\right).\end{aligned}$$

where the first inequality uses A7′(ii). Now, using Markov's inequality we have

$$\begin{aligned}&\Pr\left(P^{-1/2}\textstyle\sum_t \left\|\nabla_{\mathbf{e}}\mathbf{g}_p(\boldsymbol{\tau}_0;\bar{\mathbf{e}}_{t+1},\mathbf{x}_t)\right\|_\infty R^{-1/2+\varepsilon}>\dfrac{\eta}{\delta},\sup_t\left\|R^{1/2-\varepsilon}(\hat{\mathbf{e}}_{t+1}-\mathbf{e}^*_{t+1})\right\|_1\leqslant\delta\right)\\ &\leqslant \frac{\delta}{\eta}E\left(P^{-1/2}\textstyle\sum_t \left\|\nabla_{\mathbf{e}}\mathbf{g}_p(\boldsymbol{\tau}_0;\bar{\mathbf{e}}_{t+1},\mathbf{x}_t)\right\|_\infty R^{-1/2+\varepsilon}\right).\end{aligned}$$

Moreover, $\left\|\nabla_{\mathbf{e}}\mathbf{g}_p(\boldsymbol{\tau}_0;\bar{\mathbf{e}}_{t+1},\mathbf{x}_t)\right\|_\infty \leqslant \left\|\Delta_{\mathbf{ee}}L_p\left(\boldsymbol{\tau}_0,\bar{\mathbf{e}}_{t+1}\right)\right\|_\infty\cdot\left\|\mathbf{x}_t\right\|_1$ so that under A9

$$\begin{aligned}&E\left(\left\|\nabla_{\mathbf{e}}\mathbf{g}_p(\boldsymbol{\tau}_0;\bar{\mathbf{e}}_{t+1},\mathbf{x}_t)\right\|_\infty\right)\\ &\leqslant E\left(\sup_{c\in(0,1)}\left\|\nabla_{\mathbf{e}}\mathbf{g}_p(\boldsymbol{\tau}_0;c\hat{\mathbf{e}}_{t+1}+(1-c)\mathbf{e}^*_{t+1},\mathbf{x}_t)\right\|_\infty\right)\\ &\leqslant 2(p-1)\left(pc_3+d_3\right)E\left(\left\|\mathbf{x}_t\right\|_1\sup_{c\in(0,1)}\left\|c\hat{\mathbf{e}}_{t+1}+(1-c)\mathbf{e}^*_{t+1}\right\|_1^{p-2}\right)<\infty.\end{aligned}$$

Now

$$\begin{aligned}E\left(P^{-1/2}\textstyle\sum_t \left\|\nabla_{\mathbf{e}}\mathbf{g}_p(\boldsymbol{\tau}_0;\bar{\mathbf{e}}_{t+1},\mathbf{x}_t)\right\|_\infty R^{-1/2+\varepsilon}\right)&\leqslant E\left(\left\|\nabla_{\mathbf{e}}\mathbf{g}_p(\boldsymbol{\tau}_0;\bar{\mathbf{e}}_{t+1},\mathbf{x}_t)\right\|_\infty\right)P^{-1/2}\textstyle\sum_t R^{-1/2+\varepsilon}\\ &\leqslant E\left(\left\|\nabla_{\mathbf{e}}\mathbf{g}_p(\boldsymbol{\tau}_0;\bar{\mathbf{e}}_{t+1},\mathbf{x}_t)\right\|_\infty\right)\left(\frac{P}{R^{1-2\varepsilon}}\right)^{1/2}\to 0\end{aligned}$$

as $R \to \infty$ and $P \to \infty$, where the last limit results uses A7$'$(i). Hence $\sqrt{P}\hat{\mathbf{m}} \to \mathbf{0}$ as $R \to \infty$ and $P \to \infty$. The term $\sqrt{P}(\hat{\mathbf{m}}_P - \hat{\mathbf{m}} - \hat{\mathbf{m}}_P^*)$ on the right-hand side of Equation (S-6) is $o_p(\mathbf{1})$ provided that $\mathbf{g}$ satisfies a certain Lipshitz condition (given below) and that A7$'$ holds. Using a reasoning similar to that above, we have any $\eta > 0$ and $\delta > 0$

$$\begin{aligned}&\lim_{R,P\to\infty} \Pr\left(P^{1/2}\left\|\hat{\mathbf{m}}_P - \hat{\mathbf{m}} - \hat{\mathbf{m}}_P^*\right\|_1 > \eta\right)\\ &\leqslant \lim_{R,P\to\infty} \Pr\left(P^{1/2}\left\|\hat{\mathbf{m}}_P - \hat{\mathbf{m}} - \hat{\mathbf{m}}_P^*\right\|_1 > \eta, \sup_t R^{1/2-\varepsilon}\left\|\hat{\mathbf{e}}_{t+1} - \mathbf{e}_{t+1}^*\right\|_1 \leqslant \delta\right).\end{aligned}$$

Now, let $r_P(\delta) \equiv \sup\{r_{t+1}(\hat{\mathbf{e}}_{t+1}) : \left\|\hat{\mathbf{e}}_{t+1} - \mathbf{e}_{t+1}^*\right\|_1 \leqslant \delta, R \leqslant t \leqslant T\}$, where we let

$$\begin{aligned}&r_{t+1}(\hat{\mathbf{e}}_{t+1})\\ &\equiv \frac{\left\|\mathbf{g}_p(\boldsymbol{\tau}_0; \hat{\mathbf{e}}_{t+1}, \mathbf{x}_t) - \mathbf{g}_p(\boldsymbol{\tau}_0; \mathbf{e}_{t+1}^*, \mathbf{x}_t) - [\Delta_{\mathbf{ee}} L_p\left(\boldsymbol{\tau}_0, \mathbf{e}_{t+1}^*\right) \otimes \mathbf{x}_t](\hat{\mathbf{e}}_{t+1} - \mathbf{e}_{t+1}^*)\right\|_1}{\left\|\hat{\mathbf{e}}_{t+1} - \mathbf{e}_{t+1}^*\right\|_1},\end{aligned} \tag{S-10}$$

where $\Delta_{\mathbf{ee}} L_p\left(\boldsymbol{\tau}_0, \mathbf{e}_{t+1}^*\right)$ is as defined in Equation (S-8). Then, by the definition of $r_{t+1}(\hat{\mathbf{e}}_{t+1})$,

$$\begin{aligned}&P^{1/2}\left\|\hat{\mathbf{m}}_P - \hat{\mathbf{m}} - \hat{\mathbf{m}}_P^*\right\|_1\\ &\leqslant P^{1/2}\Big\{\Big\|P^{-1}\textstyle\sum_t [\Delta_{\mathbf{ee}} L_p\left(\boldsymbol{\tau}_0, \mathbf{e}_{t+1}^*\right) \otimes \mathbf{x}_t](\hat{\mathbf{e}}_{t+1} - \mathbf{e}_{t+1}^*) - E\{[\Delta_{\mathbf{ee}} L_p\left(\boldsymbol{\tau}_0, \mathbf{e}_{t+1}^*\right) \otimes \mathbf{x}_t](\hat{\mathbf{e}}_{t+1} - \mathbf{e}_{t+1}^*)\}\Big\|_1\\ &\qquad + P^{-1}\textstyle\sum_t r_{t+1}(\hat{\mathbf{e}}_{t+1})\left\|\hat{\mathbf{e}}_{t+1} - \mathbf{e}_{t+1}^*\right\|_1 + E\left(r_{t+1}(\hat{\mathbf{e}}_{t+1})\left\|\hat{\mathbf{e}}_{t+1} - \mathbf{e}_{t+1}^*\right\|_1\right)\Big\}\\ &\leqslant P^{1/2}\Big\{P^{-1}\textstyle\sum_t \left\|[\Delta_{\mathbf{ee}} L_p\left(\boldsymbol{\tau}_0, \mathbf{e}_{t+1}^*\right) \otimes \mathbf{x}_t] - E\{[\Delta_{\mathbf{ee}} L_p\left(\boldsymbol{\tau}_0, \mathbf{e}_{t+1}^*\right) \otimes \mathbf{x}_t]\}\right\|_\infty \sup_t \left\|\hat{\mathbf{e}}_{t+1} - \mathbf{e}_{t+1}^*\right\|_1\\ &\qquad + [r_P(\delta_R) + E(r_P(\delta_R))]\sup_t \left\|\hat{\mathbf{e}}_{t+1} - \mathbf{e}_{t+1}^*\right\|_1\Big\}\end{aligned}$$

and so by the Markov inequality

$$\begin{aligned}&\Pr\left(P^{1/2}\left\|\hat{\mathbf{m}}_P - \hat{\mathbf{m}} - \hat{\mathbf{m}}_P^*\right\|_1 > \eta, \sup_t R^{1/2-\varepsilon}\left\|\hat{\mathbf{e}}_{t+1} - \mathbf{e}_{t+1}^*\right\|_1 \leqslant \delta\right)\\ &\leqslant \left(\frac{P}{R^{1-2\varepsilon}}\right)^{1/2}\frac{\delta}{\eta}\Big[E\Big(P^{-1}\sum_t \left\|[\Delta_{\mathbf{ee}} L_p\left(\boldsymbol{\tau}_0, \mathbf{e}_{t+1}^*\right) \otimes \mathbf{x}_t] - E\{[\Delta_{\mathbf{ee}} L_p\left(\boldsymbol{\tau}_0, \mathbf{e}_{t+1}^*\right) \otimes \mathbf{x}_t]\}\right\|_\infty\Big)\\ &\qquad + E\big(r_P(\delta_R) + E(r_P(\delta_R))\big)\Big].\end{aligned}$$

Using standard arguments for stochastic equicontinuity such as those given in **?**, we can show that $r_{t+1}(\hat{\mathbf{e}}_{t+1}) \to 0$ as $\Pr(\left\|\hat{\mathbf{e}}_{t+1} - \mathbf{e}_{t+1}^*\right\|_1 > \varepsilon) \to 0$ for any $\varepsilon > 0$, so that $r_P(\delta) \to 0$ with probability 1, which by the dominated convergence theorem ensures $E(r_P(\delta)) \to 0$.

Next, we show that the sample mean of $\{\Delta_{\mathbf{ee}} L_p\left(\boldsymbol{\tau}_0, \mathbf{e}^*_{t+1}\right) \otimes \mathbf{x}_t\}$ converges in probability to its expected value. By assumption A4′ we know that $\{\Delta_{\mathbf{ee}} L_p\left(\boldsymbol{\tau}_0, \mathbf{e}^*_{t+1}\right) \otimes \mathbf{x}_t\}$ is strictly stationary and α-mixing with α of size $-r/(r-2)$ with $r > 2$ [see Theorems 3.35 and 3.49 in **?**]. Moreover, for $\delta = \min\{\varepsilon/2, \epsilon/2\} > 0$ in assumptions A4′ and A8, we have

$$\begin{aligned}
&E[\left\|\Delta_{\mathbf{ee}} L_p\left(\boldsymbol{\tau}_0, \mathbf{e}^*_{t+1}\right) \otimes \mathbf{x}_t\right\|_\infty^{r+\delta}] \\
&\leqslant \{E[\left\|\Delta_{\mathbf{ee}} L_p\left(\boldsymbol{\tau}_0, \mathbf{e}^*_{t+1}\right)\right\|_\infty^{2r+2\delta}] E[\|\mathbf{x}_t\|_1^{2r+2\delta}]\}^{1/2} \\
&\leqslant \left(\max\{E[\left\|\Delta_{\mathbf{ee}} L_p\left(\boldsymbol{\tau}_0, \mathbf{e}^*_{t+1}\right)\right\|_\infty^{2r+\epsilon}], 1\}\right)^{1/2} \left(\max\{E[\|\mathbf{x}_t\|_1^{2r+\varepsilon}], 1\}\right)^{1/2} < \infty,
\end{aligned}$$

since from Equation (S-9) we know

$$\begin{aligned}
&\left\|\Delta_{\mathbf{ee}} L_p\left(\boldsymbol{\tau}_0, \mathbf{e}^*_{t+1}\right)\right\|_\infty^{2r+\epsilon} \\
&\leqslant n_r\{[2p]^{2r+\epsilon} \left\|\mathbf{V}_p(\mathbf{e}^*_{t+1})\right\|_\infty^{2r+\epsilon} + [2(p-1)\left(pc_3 + d_3\right)]^{2r+\epsilon} \left\|\mathbf{e}^*_{t+1}\right\|_1^{(p-2)(2r+\epsilon)}\},
\end{aligned}$$

where again n_r is such that for any $(a, b) > 0$ we have $(a+b)^{2r+\epsilon} \leqslant n_r(a^{2r+\epsilon} + b^{2r+\epsilon})$; and A10 and A4′ imply that $E[\left\|\mathbf{V}_1(\mathbf{e}^*_{t+1})\right\|_\infty^{2r+\epsilon}] \leqslant M$, $E[\left\|\mathbf{V}_p(\mathbf{e}^*_{t+1})\right\|_\infty^{2r+\epsilon}] = 0$ and $E[\left\|\mathbf{e}^*_{t+1}\right\|_1^{(p-2)(2r+\epsilon)}] < \infty$. Using the weak LLN for α-mixing sequences [e.g., Corollary 3.48 in **?**] then gives

$$P^{-1} \sum_{t=R}^{T} \Delta_{\mathbf{ee}} L_p\left(\boldsymbol{\tau}_0, \mathbf{e}^*_{t+1}\right) \otimes \mathbf{x}_t \xrightarrow{p} E[\Delta_{\mathbf{ee}} L_p\left(\boldsymbol{\tau}_0, \mathbf{e}^*_{t+1}\right) \otimes \mathbf{x}_t]$$

as $P \to \infty$. Then, combining all of the above with A7′(ii) gives

$$\lim_{R,P\to\infty} \Pr\left(\sqrt{P}\left\|\hat{\mathbf{m}}_P - \hat{\mathbf{m}} - \hat{\mathbf{m}}^*_P\right\|_1 > \eta, \sup_{R\leqslant t\leqslant T} R^{1/2-\varepsilon}\left\|\hat{\mathbf{e}}_{t+1} - \mathbf{e}^*_{t+1}\right\|_1 \leqslant \delta\right) = 0$$

and the term $\sqrt{P}(\hat{\mathbf{m}}_P - \hat{\mathbf{m}} - \hat{\mathbf{m}}^*_P)$ on the right-hand side of Equation (S-6) is $o_p(\mathbf{1})$ as $R \to \infty$ and $P \to \infty$. Finally, we use the central limit theorem (CLT) for strictly stationary and α-mixing sequences [e.g., Theorem 5.20 in **?**] to show that $\sqrt{P}\hat{\mathbf{m}}^*_P \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{S})$. Using Theorems 3.35 and 3.49 in **?**, which together show that time-invariant measurable functions of strictly stationary and mixing sequences are strictly stationary and mixing of the same size, we know by A4′ that $\{\mathbf{g}_p(\boldsymbol{\tau}_0; \mathbf{e}^*_{t+1}, \mathbf{x}_t)\}$ is strictly stationary and α-mixing with mixing coefficient of size $-r/(r-2)$, $r > 2$. In the proof of Theorem

2 we have moreover shown that $E[\left\|\mathbf{g}_p(\boldsymbol{\tau}_0;\mathbf{e}^*_{t+1},\mathbf{x}_t)\right\|_1^{r+\varepsilon}] < \infty$. The CLT [e.g., Theorem 5.20 in **?**] then ensures

$$\sqrt{P}\hat{\mathbf{m}}^*_P \xrightarrow{d} \mathcal{N}(\mathbf{0},\mathbf{S}). \tag{S-11}$$

The remainder of the asymptotic normality proof is similar to the standard case: the positive definiteness of $\mathbf{S}^{-1}$, $\hat{\mathbf{S}} \xrightarrow{p} \mathbf{S}$ and $\hat{\mathbf{B}}_P \xrightarrow{p} \mathbf{B}^* \equiv E[\mathbf{B}(p,\mathbf{e}^*_{t+1},\mathbf{x}_t)]$ as $R \to \infty$ and $P \to \infty$ ($\mathbf{B}$ was defined in Equation (12)) together with A5(ii) ensure that $(\mathbf{B}^{*\prime}\mathbf{S}^{-1}\mathbf{B}^*)^{-1}$ exists, so by using $\sqrt{P}(\hat{\boldsymbol{\tau}}_P-\boldsymbol{\tau}_0) = -[\hat{\mathbf{B}}'_P\hat{\mathbf{S}}^{-1}\hat{\mathbf{B}}_P]^{-1}\hat{\mathbf{B}}'_P\hat{\mathbf{S}}^{-1}[\sqrt{P}\hat{\mathbf{m}}^*_P+o_p(\mathbf{1})]$, the limit result in (S-11) and the Slutsky theorem we have $\sqrt{P}(\hat{\boldsymbol{\tau}}_P-\boldsymbol{\tau}_0) \xrightarrow{d} \mathcal{N}(\mathbf{0},(\mathbf{B}^{*\prime}\mathbf{S}^{-1}\mathbf{B}^*)^{-1})$, which completes the proof of asymptotic normality. □