# Supplementary Materials

## Simulation based Bias Correction Methods for Complex Models

## A Proof of Corollary 1

For $\hat{\boldsymbol{\theta}}$ and $\tilde{\boldsymbol{\theta}}$, the proof is directly obtained by verifying the conditions of Theorem 2.1 of Newey and McFadden (1994). First, $\hat{\boldsymbol{\theta}}$ satisfies

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta} \in \boldsymbol{\Theta}}{\text{argmin}} \ \hat{Q}\left(\boldsymbol{\theta}, n\right),$$

where $\hat{Q}\left(\boldsymbol{\theta}, n\right) = ||\hat{\boldsymbol{\pi}}(\boldsymbol{\theta}_0, n) - \bar{\boldsymbol{\pi}}(\boldsymbol{\theta}, n)||_2^2$. Let $Q\left(\boldsymbol{\theta}\right) = ||\boldsymbol{\pi}(\boldsymbol{\theta}_0) - \boldsymbol{\pi}(\boldsymbol{\theta})||_2^2$. By Assumption 1, $Q\left(\boldsymbol{\theta}\right)$ is a continuous function of $\boldsymbol{\theta}$ and is uniquely minimized at $\boldsymbol{\theta}_0$. Consider the absolute value of the difference between $\hat{Q}\left(\boldsymbol{\theta}, n\right)$ and $Q\left(\boldsymbol{\theta}\right)$, i.e.

$$
\begin{aligned}
\left|\hat{Q}\left(\boldsymbol{\theta},n\right)-Q\left(\boldsymbol{\theta}\right)\right| &= \big|\, ||\hat{\boldsymbol{\pi}}(\boldsymbol{\theta}_0,n)-\bar{\boldsymbol{\pi}}(\boldsymbol{\theta},n)||_2^2-||\boldsymbol{\pi}(\boldsymbol{\theta}_0)-\boldsymbol{\pi}(\boldsymbol{\theta})||_2^2\,\big| \\
&= \big|||\{\hat{\boldsymbol{\pi}}(\boldsymbol{\theta}_0,n)-\boldsymbol{\pi}(\boldsymbol{\theta}_0)\}+\{\boldsymbol{\pi}(\boldsymbol{\theta}_0)-\boldsymbol{\pi}(\boldsymbol{\theta})\} \\
&+\{\boldsymbol{\pi}(\boldsymbol{\theta})-\bar{\boldsymbol{\pi}}(\boldsymbol{\theta},n)\}||_2^2-||\boldsymbol{\pi}(\boldsymbol{\theta}_0)-\boldsymbol{\pi}(\boldsymbol{\theta})||_2^2\big| \\
&= \big|||\hat{\boldsymbol{\pi}}(\boldsymbol{\theta}_0,n)-\boldsymbol{\pi}(\boldsymbol{\theta}_0)||_2^2+||\bar{\boldsymbol{\pi}}(\boldsymbol{\theta},n)-\boldsymbol{\pi}(\boldsymbol{\theta})||_2^2 \\
&+2\{\hat{\boldsymbol{\pi}}(\boldsymbol{\theta}_0,n)-\boldsymbol{\pi}(\boldsymbol{\theta}_0)\}^T\{\boldsymbol{\pi}(\boldsymbol{\theta}_0)-\bar{\boldsymbol{\pi}}(\boldsymbol{\theta},n)\} \\
&+2\{\boldsymbol{\pi}(\boldsymbol{\theta})-\bar{\boldsymbol{\pi}}(\boldsymbol{\theta},n)\}^T\{\boldsymbol{\pi}(\boldsymbol{\theta}_0)-\boldsymbol{\pi}(\boldsymbol{\theta})\}\big| \\
&\leq ||\hat{\boldsymbol{\pi}}(\boldsymbol{\theta}_0,n)-\boldsymbol{\pi}(\boldsymbol{\theta}_0)||_2^2+||\bar{\boldsymbol{\pi}}(\boldsymbol{\theta},n)-\boldsymbol{\pi}(\boldsymbol{\theta})||_2^2 \\
&+2\|\hat{\boldsymbol{\pi}}(\boldsymbol{\theta}_0,n)-\boldsymbol{\pi}(\boldsymbol{\theta}_0)\|_2\|\boldsymbol{\pi}(\boldsymbol{\theta}_0)-\bar{\boldsymbol{\pi}}(\boldsymbol{\theta},n)\|_2 \\
&+2\|\boldsymbol{\pi}(\boldsymbol{\theta})-\bar{\boldsymbol{\pi}}(\boldsymbol{\theta},n)\|_2\|\boldsymbol{\pi}(\boldsymbol{\theta}_0)-\boldsymbol{\pi}(\boldsymbol{\theta})\|_2.
\end{aligned}
$$

Because of the uniform consistency of $\hat{\boldsymbol{\pi}}(\boldsymbol{\theta},n)$ as an estimator of $\pi(\boldsymbol{\theta})$ (see (1)), we have $\hat{\boldsymbol{\pi}}(\boldsymbol{\theta},n)\overset{p}{\to}\boldsymbol{\pi}(\boldsymbol{\theta})$ and hence

$$
\bar{\boldsymbol{\pi}}(\boldsymbol{\theta},n)=\frac{1}{H}\sum_{h=1}^{H}\hat{\boldsymbol{\pi}}_h(\boldsymbol{\theta},n)\overset{p}{\to}\boldsymbol{\pi}(\boldsymbol{\theta}),
$$

uniformly. The function $\boldsymbol{\pi}(\boldsymbol{\theta})$ is continuous and consequently we have that $||\boldsymbol{\pi}(\boldsymbol{\theta}_0)-\bar{\boldsymbol{\pi}}(\boldsymbol{\theta},n)||_2$ and $||\boldsymbol{\pi}(\boldsymbol{\theta}_0)-\boldsymbol{\pi}(\boldsymbol{\theta})||_2$ are bounded for all $\boldsymbol{\theta}\in\boldsymbol{\Theta}$ and that

$$
\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}\left|\hat{Q}\left(\boldsymbol{\theta},n\right)-Q\left(\boldsymbol{\theta}\right)\right|\overset{p}{\to}0.
$$

By combining the above results, the four condition of Theorem 2.1. of Newey and McFadden (1994) can be verified implying that $\hat{\boldsymbol{\theta}}\overset{p}{\to}\boldsymbol{\theta}_0$. Using the same argument and simply replacing $\bar{\boldsymbol{\pi}}(\boldsymbol{\theta}, n)$ by $\hat{\boldsymbol{\pi}}(\boldsymbol{\theta}, nH)$, we have that $\tilde{\boldsymbol{\theta}}$ defined in (5) is also a consistent estimator for $\boldsymbol{\theta}_0$.

Finally, the consistency of $\tilde{\boldsymbol{\theta}}_B$ is directly implied by the consistency of $\hat{\boldsymbol{\theta}}$ and Theorem 4 presented in Section 3 of the main document. $\square$

## B Proof of Theorem 1

From (3) we have that $\hat{\boldsymbol{\theta}}$ satisfies:

$$\hat{\boldsymbol{\pi}}(\boldsymbol{\theta}_0, n) = \bar{\boldsymbol{\pi}}(\hat{\boldsymbol{\theta}}, n).$$

Using (2) and (4), we have

$$\begin{aligned}\hat{\boldsymbol{\pi}}(\boldsymbol{\theta}_0, n) &= \boldsymbol{\theta}_0 + \mathbf{b}\left(\boldsymbol{\theta}_0, n\right) + \mathbf{c}(n) + \mathbf{v}\left(\boldsymbol{\theta}_{\mathbf{0}}, n\right) = \bar{\boldsymbol{\pi}}(\hat{\boldsymbol{\theta}}, n) \\ &= \hat{\boldsymbol{\theta}} + \mathbf{b}\left(\hat{\boldsymbol{\theta}}, n\right) + \mathbf{c}(n) + \frac{1}{H}\sum_{h=1}^{H} \mathbf{v}_h\left(\hat{\boldsymbol{\theta}}, n\right),\end{aligned}$$

with $\mathbf{v}_h\left(\hat{\boldsymbol{\theta}}, n\right)$ corresponding to the noise of the $h^{\text{th}}$ simulated sample. By rearranging the terms and defining

$$\tilde{\mathbf{v}} \equiv \mathbf{v}\left(\boldsymbol{\theta}_{\mathbf{0}}, n\right) - \frac{1}{H}\sum_{h=1}^{H} \mathbf{v}_h\left(\hat{\boldsymbol{\theta}}, n\right),$$

we obtain,

$$\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}_0 + \left\{\mathbf{b}\left(\boldsymbol{\theta}_0, n\right) - \mathbf{b}\left(\hat{\boldsymbol{\theta}}, n\right)\right\} + \tilde{\mathbf{v}}. \tag{1}$$

We now consider $\mathbb{E}\left[\hat{\theta}_j\right] - \theta_j$, with $\hat{\theta}_j$, respectively $\theta_j$, the $j^{\text{th}}$ element of $\hat{\boldsymbol{\theta}}$, respectively $\boldsymbol{\theta}_0$. Using (1), we get

$$\mathbb{E}\left[\hat{\theta}_j\right] - \theta_j = \mathbb{E}\left[\sum_{i=1}^{p} a_{i,j}\frac{\left(\theta_i - \hat{\theta}_i\right)}{n} + \sum_{k=1}^{p}\sum_{l=1}^{p} d_{k,l,j}\frac{\left(\theta_k\theta_l - \hat{\theta}_k\hat{\theta}_l\right)}{n^2} + \mathcal{O}_p\left(n^{-3}\right)\right].$$

Next, we define $\mathbf{u} \equiv [u_1 \ \ldots \ u_p]^T$, where

$$u_j \equiv \sum_{k=1}^{p} \sum_{l=1}^{p} d_{k,l,j} \frac{\left(\theta_k \theta_l - \hat{\theta}_k \hat{\theta}_l\right)}{n^2} + \mathcal{O}_p\left(n^{-3}\right),$$

and using this new quantity we can write

$$\mathbb{E}\left[\hat{\boldsymbol{\theta}}\right] - \boldsymbol{\theta}_0 = \frac{1}{n} \mathbf{A} \left(\boldsymbol{\theta}_0 - \mathbb{E}\left[\hat{\boldsymbol{\theta}}\right]\right) + \mathbb{E}\left[\mathbf{u}\right],$$

where $\mathbf{A} \equiv \left[a_{i,j}\right]_{i,j=1,\ldots,p}$. Therefore, we obtain

$$\left(\mathbf{I}_p + \frac{1}{n}\mathbf{A}\right)\left(\mathbb{E}\left[\hat{\boldsymbol{\theta}}\right] - \boldsymbol{\theta}_0\right) = \mathbb{E}\left[\mathbf{u}\right].$$

For sufficiently large $n$ the matrix $\left(\mathbf{I}_p + \frac{1}{n}\mathbf{A}\right)^{-1}$ is invertible and we obtain

$$\mathbb{E}\left[\hat{\boldsymbol{\theta}}\right] - \boldsymbol{\theta}_0 = \left(\mathbf{I}_p + \frac{1}{n}\mathbf{A}\right)^{-1} \mathbb{E}\left[\mathbf{u}\right]. \tag{2}$$

Because $\mathbf{u} = \mathcal{O}_p(n^{-2})$, a direct consequence of (2) is that

$$\mathbb{E}\left[\hat{\boldsymbol{\theta}}\right] - \boldsymbol{\theta}_0 = \mathcal{O}(n^{-2}). \tag{3}$$

We now consider the variance of $\hat{\boldsymbol{\theta}}$. Using (1), we get

$$\operatorname{var}\left(\hat{\boldsymbol{\theta}}\right) = \operatorname{var}\left\{\mathbf{b}\left(\hat{\boldsymbol{\theta}}, n\right)\right\} + \operatorname{var}\left(\tilde{\mathbf{v}}\right) - \left(\mathbf{W} + \mathbf{W}^T\right), \tag{4}$$

where $\mathbf{W} \equiv \text{cov}\{\mathbf{b}(\hat{\boldsymbol{\theta}}, n), \tilde{\mathbf{v}}\}$. We now investigate the three elements of (4) separately. For the second term, we have

$$\begin{aligned}\text{var}\left(\tilde{\mathbf{v}}\right) = &\text{var}\left\{\mathbf{v}\left(\boldsymbol{\theta}_{\mathbf{0}}, n\right)\right\} + \frac{1}{H}\text{var}\left\{\mathbf{v}_1\left(\hat{\boldsymbol{\theta}}, n\right)\right\} \\ &- \left[\text{cov}\left\{\mathbf{v}\left(\boldsymbol{\theta}_{\mathbf{0}}, n\right), \mathbf{v}_1\left(\hat{\boldsymbol{\theta}}, n\right)\right\} + \text{cov}^T\left\{\mathbf{v}\left(\boldsymbol{\theta}_{\mathbf{0}}, n\right), \mathbf{v}_1\left(\hat{\boldsymbol{\theta}}, n\right)\right\}\right].\end{aligned}$$

Moreover, we have

$$\begin{aligned}\text{cov}\left\{\mathbf{v}\left(\boldsymbol{\theta}_{\mathbf{0}}, n\right), \mathbf{v}_1\left(\hat{\boldsymbol{\theta}}, n\right)\right\} &= \mathbb{E}\left[\text{cov}\left\{\mathbf{v}\left(\boldsymbol{\theta}_{\mathbf{0}}, n\right), \mathbf{v}_1\left(\hat{\boldsymbol{\theta}}, n\right)|\hat{\boldsymbol{\theta}}\right\}\right] \\ &+ \text{cov}\left[\mathbb{E}\left\{\mathbf{v}\left(\boldsymbol{\theta}_{\mathbf{0}}, n\right)|\hat{\boldsymbol{\theta}}\right\}, \mathbb{E}\left\{\mathbf{v}_1\left(\hat{\boldsymbol{\theta}}, n\right)|\hat{\boldsymbol{\theta}}\right\}\right] \\ &= \mathbf{0}\end{aligned} \tag{5}$$

Thus, we get

$$\text{var}\left(\tilde{\mathbf{v}}\right) = n^{-\alpha}\mathbf{V}_{\boldsymbol{\theta}_0, n} + \frac{1}{H}\mathbb{E}\left[\text{var}\left\{\mathbf{v}_1\left(\hat{\boldsymbol{\theta}}, n\right)|\hat{\boldsymbol{\theta}}\right\}\right]. \tag{6}$$

Using (3), Assumption 4 and performing a MacLaurin expansion on the last term of (6) we obtain

$$\begin{aligned}&\mathbb{E}\left[\text{var}\left\{\mathbf{v}_1\left(\hat{\boldsymbol{\theta}}, n\right)|\hat{\boldsymbol{\theta}}\right\}\right] \\ &= \text{var}\left\{\mathbf{v}_1\left(\boldsymbol{\theta}_0, n\right)\right\} + \mathbb{E}\left[\mathbf{D}_1(\boldsymbol{\theta}^*, n)(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0), \ldots, \mathbf{D}_p(\boldsymbol{\theta}^*, n)(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)\right] \\ &= \text{var}\left\{\mathbf{v}_1\left(\boldsymbol{\theta}_0, n\right)\right\} + \mathcal{O}\left(n^{-2}\right),\end{aligned}$$

where $\boldsymbol{\theta}^* \in \boldsymbol{\Theta}$ is on the line connecting $\hat{\boldsymbol{\theta}}$ and $\boldsymbol{\theta}_0$. Therefore, we get

$$\text{var}\left(\tilde{\mathbf{v}}\right) = n^{-\alpha}\left(1 + \frac{1}{H}\right)\mathbf{V}_{\boldsymbol{\theta}_0, n} + \mathcal{O}(n^{-2}). \tag{7}$$

Note that the above results directly implies that $\text{var}(\tilde{\mathbf{v}}) = \mathcal{O}\left(n^{-\min(\alpha,2)}\right)$.

Next, we consider the first element of (4) and we study the variance of the $j^{\text{th}}$ element of the vector $\mathbf{b}(\hat{\boldsymbol{\theta}}, n)$. For simplicity, we define $r_j \equiv \sum_{k=1}^{p}\sum_{l=1}^{p} d_{k,l,j}\frac{\hat{\theta}_k\hat{\theta}_l}{n^2} + \mathcal{O}_p\left(n^{-3}\right)$, which is $\mathcal{O}_p(n^{-2})$. Then, using Cauchy-Schwarz inequality, we have

$$\begin{aligned}
\text{var}\left\{b_j\left(\hat{\boldsymbol{\theta}}, n\right)\right\} &= \text{var}\left(\sum_{i=1}^{p} a_{i,j}\frac{\hat{\theta}_i}{n} + r_j\right)\\
&= \text{var}\left(\sum_{i=1}^{p} a_{i,j}\frac{\hat{\theta}_i}{n}\right) + 2\,\text{cov}\left(\sum_{i=1}^{p} a_{i,j}\frac{\hat{\theta}_i}{n}, r_j\right) + \mathcal{O}\left(n^{-4}\right)\\
&= \frac{1}{n^2}\sum_{k=1}^{p}\sum_{l=1}^{p} a_{k,j}a_{l,j}\,\text{cov}\left(\hat{\theta}_k, \hat{\theta}_l\right) + \mathcal{O}\left(n^{-3}\right)\\
&\leq \frac{p^2}{n^2}\max_{i,j=1,\ldots,p}\left\{a_{k,j}a_{l,j}\,\text{cov}\left(\hat{\theta}_k, \hat{\theta}_l\right)\right\} + \mathcal{O}\left(n^{-3}\right)\\
&= \mathcal{O}\left(n^{-\min(3,2+\alpha)}\right),
\end{aligned}$$

since from (4) and (7) we have $\text{var}\left(\hat{\theta}_i\right) = \mathcal{O}\left(n^{-\min(\alpha,2)}\right)$.

Hence

$$\text{var}\left\{\mathbf{b}\left(\hat{\boldsymbol{\theta}}, n\right)\right\} = \mathcal{O}\left(n^{-\min(2+\alpha,3)}\right), \tag{8}$$

componentwise.

Considering the last term of (4) and using $w_{j,k}$ to denote the $(j,k)$ element of the matrix $\mathbf{W}$, we obtain by Cauchy-Schwarz inequality together with (8) and (6) that

$$w_{j,k}^2 \leq \text{var}\left\{b_j\left(\hat{\boldsymbol{\theta}}, n\right)\right\}\text{var}\left(\tilde{v}_k\right) = \mathcal{O}\left(n^{-\min(3+\alpha,2+2\alpha,5)}\right). \tag{9}$$

By combining the results of (8), (6) and (9) , we get

$$\mathrm{var}\left(\hat{\boldsymbol{\theta}}\right) = n^{-\alpha}\left(1+\frac{1}{H}\right)\mathbf{V}_{\boldsymbol{\theta}_0,n} + \mathcal{O}\left(n^{-\min(2,1+\alpha)}\right), \tag{10}$$

which verifies the second part of Theorem 1. Note that from (10) we also have that $\mathrm{var}\left(\hat{\boldsymbol{\theta}}\right) = \mathcal{O}\left(n^{-\min(2,\alpha)}\right)$.

Next, we return to (3) and study further the term $\mathbf{u}$. Indeed, using (3) we have that

$$\mathbb{E}\left[\hat{\theta}_k\right]\mathbb{E}\left[\hat{\theta}_l\right] = \left\{\theta_k + \mathcal{O}(n^{-2})\right\}\left\{\theta_l + \mathcal{O}(n^{-2})\right\} = \theta_k\theta_l + \mathcal{O}(n^{-2}),$$

and therefore, we obtain

$$\begin{aligned}
\mathbb{E}\left[u_j\right] &= \frac{1}{n^2}\sum_{k=1}^{p}\sum_{l=1}^{p} d_{k,l,j}\left(\theta_k\theta_l - \mathbb{E}\left[\hat{\theta}_k\hat{\theta}_l\right]\right) + \mathcal{O}\left(n^{-3}\right)\\
&= \frac{1}{n^2}\sum_{k=1}^{p}\sum_{l=1}^{p} d_{k,l,j}\left\{\theta_k\theta_l - \mathbb{E}\left[\hat{\theta}_k\right]\mathbb{E}\left[\hat{\theta}_l\right] - \mathrm{cov}\left(\hat{\theta}_k,\hat{\theta}_l\right)\right\} + \mathcal{O}\left(n^{-3}\right)\\
&= \frac{1}{n^2}\sum_{k=1}^{p}\sum_{l=1}^{p} d_{k,l,j}\left(\theta_k\theta_l - \mathbb{E}\left[\hat{\theta}_k\right]\mathbb{E}\left[\hat{\theta}_l\right]\right) + \mathcal{O}\left(n^{-\min(2+\alpha,3)}\right)\\
&= \mathcal{O}\left(n^{-\min(2+\alpha,3)}\right).
\end{aligned}$$

Using (2), we finally get

$$\mathbb{E}\left[\hat{\boldsymbol{\theta}}\right] - \boldsymbol{\theta}_0 = \left(\mathbf{I}_p + \frac{1}{n}\mathbf{A}\right)^{-1}\mathbb{E}\left[\mathbf{u}\right] = \mathcal{O}\left(n^{-\min(2+\alpha,3)}\right),$$

which verifies the first part of Theorem 1 and concludes the proof. □

# C Proof of Theorem 2

From (5) we have that $\tilde{\boldsymbol{\theta}}$ satisfies:

$$\hat{\boldsymbol{\pi}}(\boldsymbol{\theta}_0, n) = \hat{\boldsymbol{\pi}}(\tilde{\boldsymbol{\theta}}, nH).$$

Using (2), we have

$$\begin{aligned}\hat{\boldsymbol{\pi}}(\boldsymbol{\theta}_0, n) &= \boldsymbol{\theta}_0 + \mathbf{b}\left(\boldsymbol{\theta}_0, n\right) + \mathbf{c}(n) + \mathbf{v}\left(\boldsymbol{\theta}_{\mathbf{0}}, n\right) = \hat{\boldsymbol{\pi}}(\tilde{\boldsymbol{\theta}}, nH)\\ &= \tilde{\boldsymbol{\theta}} + \mathbf{b}\left(\tilde{\boldsymbol{\theta}}, nH\right) + \mathbf{c}(nH) + \mathbf{v}\left(\tilde{\boldsymbol{\theta}}, nH\right),\end{aligned}$$

By rearranging the terms and defining

$$\mathbf{v}^* \equiv \mathbf{v}\left(\boldsymbol{\theta}_{\mathbf{0}}, n\right) - \mathbf{v}\left(\tilde{\boldsymbol{\theta}}, nH\right),$$

we obtain,

$$\tilde{\boldsymbol{\theta}} = \boldsymbol{\theta}_0 + \left\{\mathbf{b}\left(\boldsymbol{\theta}_0, n\right) - \mathbf{b}\left(\tilde{\boldsymbol{\theta}}, nH\right)\right\} + \left\{\mathbf{c}(n) - \mathbf{c}(nH)\right\} + \mathbf{v}^*. \tag{11}$$

We now consider $\mathbb{E}\left[\hat{\theta}_j\right] - \theta_j$, with $\hat{\theta}_j$, respectively $\theta_j$, the $j^{\text{th}}$ element of $\hat{\boldsymbol{\theta}}$, respectively $\boldsymbol{\theta}_0$. Using (11) and, as in the proof of Theorem 1, we define $\mathbf{u} \equiv [u_1 \ \ldots \ u_p]^T$, where

$$u_j \equiv \sum_{k=1}^{p}\sum_{l=1}^{p} d_{k,l,j} \frac{\left(\theta_k\theta_l - \hat{\theta}_k\hat{\theta}_l\right)}{n^2} + \mathcal{O}_p\left(n^{-3}\right),$$

we get

$$\mathbb{E}\left[\hat{\theta}_j\right] - \theta_j = \mathbb{E}\left[\sum_{i=1}^{p} a_{i,j} \frac{\left(\theta_i - \hat{\theta}_i\right)}{n} + u_j\right] + \left[\mathbf{c}(n) - \mathbf{c}(nH)\right]_j,$$

Therefore, we have

$$\mathbb{E}\left[\hat{\boldsymbol{\theta}}\right] - \boldsymbol{\theta}_0 = \frac{1}{n}\mathbf{A}\left(\boldsymbol{\theta}_0 - \mathbb{E}\left[\hat{\boldsymbol{\theta}}\right]\right) + \mathbb{E}\left[\mathbf{u}\right] + \mathbf{c}(n) - \mathbf{c}(nH),$$

where $\mathbf{A} \equiv \left[a_{i,j}\right]_{i,j=1,\ldots,p}$. Then we can write

$$\left(\mathbf{I}_p + \frac{1}{n}\mathbf{A}\right)\left(\mathbb{E}\left[\hat{\boldsymbol{\theta}}\right] - \boldsymbol{\theta}_0\right) = \mathbb{E}\left[\mathbf{u}\right] + \mathbf{c}(n) - \mathbf{c}(nH).$$

For sufficiently large $n$ the matrix $\left(\mathbf{I}_p + \frac{1}{n}\mathbf{A}\right)^{-1}$ is invertible and we obtain

$$\mathbb{E}\left[\hat{\boldsymbol{\theta}}\right] - \boldsymbol{\theta}_0 = \left(\mathbf{I}_p + \frac{1}{n}\mathbf{A}\right)^{-1}\left\{\mathbb{E}\left[\mathbf{u}\right] + \mathbf{c}(n) - \mathbf{c}(nH)\right\}. \tag{12}$$

Because $\mathbf{u} = \mathcal{O}_p(n^{-2})$, a direct consequence of (12) is that

$$\mathbb{E}\left[\tilde{\boldsymbol{\theta}}\right] - \boldsymbol{\theta}_0 = \mathcal{O}\left(n^{-\min(2,\lambda)}\right), \tag{13}$$

which verifies the first part of Theorem 2. We now consider the variance of $\tilde{\boldsymbol{\theta}}$. Using (11), we get

$$\operatorname{var}\left(\tilde{\boldsymbol{\theta}}\right) = \operatorname{var}\left\{\mathbf{b}\left(\tilde{\boldsymbol{\theta}}, nH\right)\right\} + \operatorname{var}\left(\mathbf{v}^*\right) - \left(\mathbf{M} + \mathbf{M}^T\right). \tag{14}$$

where $\mathbf{M} \equiv \text{cov}\left\{\mathbf{b}\left(\tilde{\boldsymbol{\theta}}, nH\right), \mathbf{v}^*\right\}$. We now investigate the three elements of (14) separately. First, it is clear $b_j\left(\tilde{\boldsymbol{\theta}}, nH\right) = \mathcal{O}_p\left((nH)^{-1}\right)$ and therefore by Cauchy-Schwarz inequality we have that $\text{var}\left\{\mathbf{b}\left(\tilde{\boldsymbol{\theta}}, nH\right)\right\} = \mathcal{O}\left((nH)^{-2}\right)$ componentwise. Considering the second term of (14), we have

$$\begin{aligned}\text{var}\left(\mathbf{v}^*\right) &= \text{var}\left\{\mathbf{v}\left(\boldsymbol{\theta}_{\mathbf{0}}, n\right)\right\} + \text{var}\left\{\mathbf{v}\left(\tilde{\boldsymbol{\theta}}, nH\right)\right\} - \left(\mathbf{G} + \mathbf{G}^T\right) \\ &= n^{-\alpha}\mathbf{V}_{\boldsymbol{\theta}_0,n} + \text{var}\left\{\mathbf{v}\left(\tilde{\boldsymbol{\theta}}, nH\right)\right\} \\ &= n^{-\alpha}\mathbf{V}_{\boldsymbol{\theta}_0,n} + \mathbb{E}\left[\text{var}\left\{\mathbf{v}\left(\tilde{\boldsymbol{\theta}}, nH\right)|\tilde{\boldsymbol{\theta}}\right\}\right],\end{aligned} \tag{15}$$

where $\mathbf{G} \equiv \text{cov}\left\{\mathbf{v}\left(\boldsymbol{\theta}_{\mathbf{0}}, n\right), \mathbf{v}\left(\tilde{\boldsymbol{\theta}}, nH\right)\right\} = \mathbf{0}$ similarly to (5) in the proof of Theorem 1. Using (13), Assumption 4 and performing a MacLaurin expansion on the last term of (15) we obtain

$$\begin{aligned}&\mathbb{E}\left[\text{var}\left\{\mathbf{v}\left(\tilde{\boldsymbol{\theta}}, nH\right)|\tilde{\boldsymbol{\theta}}\right\}\right] \\ &= \text{var}\left\{\mathbf{v}\left(\boldsymbol{\theta}_0, nH\right)\right\} + \mathbb{E}\left[\mathbf{D}_1(\boldsymbol{\theta}^*, nH)(\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}_0), \ldots, \mathbf{D}_p(\boldsymbol{\theta}^*, nH)(\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)\right] \\ &= \text{var}\left\{\mathbf{v}\left(\boldsymbol{\theta}_0, nH\right)\right\} + \mathcal{O}\left(n^{-\min(2,\lambda)}\right),\end{aligned}$$

where $\boldsymbol{\theta}^* \in \boldsymbol{\Theta}$ is on the line connecting $\boldsymbol{\theta}_0$ and $\tilde{\boldsymbol{\theta}}$. Therefore, we get

$$\text{var}\left(\mathbf{v}^*\right) = n^{-\alpha}\mathbf{V}_{\boldsymbol{\theta}_0,n} + (nH)^{-\alpha}\mathbf{V}_{\boldsymbol{\theta}_0,nH} + \mathcal{O}\left(n^{-\min(2,\lambda)}\right). \tag{16}$$

Considering the last term of (14) and using $m_{j,k}$ to denote the $(j,k)$ element of the matrix $\mathbf{M}$, we obtain by Cauchy-Schwarz inequality that

$$m_{j,k}^2 \leq \text{var}\left\{b_j\left(\tilde{\boldsymbol{\theta}}, nH\right)\right\}\text{var}\left(v_k^*\right) = \mathcal{O}\left((nH)^{-2} n^{-\min(\alpha,2,\lambda)}\right).$$

By combining the previous results, we get

$$\operatorname{var}\left(\tilde{\boldsymbol{\theta}}\right) = n^{-\alpha}\mathbf{V}_{\boldsymbol{\theta}_0,n} + (nH)^{-\alpha}\,\mathbf{V}_{\boldsymbol{\theta}_0,nH} + \mathcal{O}\left(\max\left\{n^{-\min(2,\lambda)}, H^{-1}n^{-\frac{\min(\alpha,\lambda)}{2}-1}\right\}\right),$$

which verifies the second part of Theorem 2 and concludes the proof. □

## D Proof of Theorem 3

From (7), $\hat{\boldsymbol{\theta}}_B$ is defined as:

$$\hat{\boldsymbol{\theta}}_B = 2\hat{\boldsymbol{\pi}}(\boldsymbol{\theta}_0, n) - \bar{\boldsymbol{\pi}}\left\{\hat{\boldsymbol{\pi}}(\boldsymbol{\theta}_0, n), n\right\}.$$

To simplify the notation we write $\hat{\boldsymbol{\pi}}(\boldsymbol{\theta}_0, n)$ as $\hat{\boldsymbol{\pi}}$ for the rest proof. Using (2) and (4), we obtain

$$\begin{aligned}
\hat{\boldsymbol{\theta}}_B &= \hat{\boldsymbol{\pi}} + \left\{\boldsymbol{\theta}_0 + \mathbf{b}\left(\boldsymbol{\theta}_0, n\right) + \mathbf{c}(n) + \mathbf{v}\left(\boldsymbol{\theta}_\mathbf{0}, n\right)\right\} \\
&\quad - \left\{\hat{\boldsymbol{\pi}} + \mathbf{b}\left(\hat{\boldsymbol{\pi}}, n\right) + \mathbf{c}(n) + \frac{1}{H}\sum_{h=1}^{H}\mathbf{v}_h\left(\hat{\boldsymbol{\pi}}, n\right)\right\} \\
&= \boldsymbol{\theta}_0 + \left\{\mathbf{b}\left(\boldsymbol{\theta}_0, n\right) - \mathbf{b}\left(\hat{\boldsymbol{\pi}}, n\right)\right\} + \left\{\mathbf{v}\left(\boldsymbol{\theta}_\mathbf{0}, n\right) - \frac{1}{H}\sum_{h=1}^{H}\mathbf{v}_h\left(\hat{\boldsymbol{\pi}}, n\right)\right\}.
\end{aligned} \tag{17}$$

Next, we consider $\mathbb{E}\left[b_j\left(\boldsymbol{\theta}_0, n\right) - b_j\left(\hat{\boldsymbol{\pi}}, n\right)\right]$, with $b_j(\cdot,\cdot)$, respectively $\theta_j$ and $\hat{\pi}_j$, the $j^{\text{th}}$ element of $\mathbf{b}(\cdot,\cdot)$, respectively $\boldsymbol{\theta}_0$ and $\hat{\boldsymbol{\pi}}$. We obtain

$$\begin{aligned}
&\mathbb{E}\left[b_j\left(\boldsymbol{\theta}_0, n\right) - b_j\left(\hat{\boldsymbol{\pi}}, n\right)\right] \\
&= \mathbb{E}\left[\sum_{i=1}^{p} a_{i,j}\frac{\theta_i - \hat{\pi}_i}{n} + \mathcal{O}_p\left(n^{-2}\right)\right] \\
&= \mathbb{E}\left[\sum_{i=1}^{p} a_{i,j}\frac{\theta_i - \left\{\theta_i + b_i\left(\boldsymbol{\theta}_0, n\right) + c_i(n) + v_i\left(\boldsymbol{\theta}_\mathbf{0}, n\right)\right\}}{n}\right] + \mathcal{O}\left(n^{-2}\right) \\
&= \mathbb{E}\left[\sum_{i=1}^{p} -a_{i,j}\frac{b_i\left(\boldsymbol{\theta}_0, n\right) + c_i(n)}{n}\right] + \mathcal{O}\left(n^{-2}\right) = \mathcal{O}\left(n^{-\min(1,\beta)-1}\right).
\end{aligned}$$

Therefore, we obtain

$$\begin{aligned}\mathbb{E}\left[\hat{\boldsymbol{\theta}}_B\right] - \boldsymbol{\theta}_0 &= \mathbb{E}\left[\mathbf{b}\left(\boldsymbol{\theta}_0, n\right) - \mathbf{b}\left(\hat{\boldsymbol{\pi}}, n\right)\right] + \mathbb{E}\left[\mathbf{v}\left(\boldsymbol{\theta}_{\mathbf{0}}, n\right) - \frac{1}{H}\sum_{h=1}^{H}\mathbf{v}_h\left(\hat{\boldsymbol{\pi}}, n\right)\right] \\ &= \mathcal{O}\left(n^{-\min(1,\beta)-1}\right),\end{aligned}$$

which verifies the first part of Theorem 3. We now consider the variance of $\hat{\boldsymbol{\theta}}_B$. Using (17), we get

$$\begin{aligned}\operatorname{var}\left(\hat{\boldsymbol{\theta}}_B\right) = \operatorname{var}\left\{\mathbf{b}\left(\hat{\boldsymbol{\pi}}, n\right)\right\} + \operatorname{var}&\left\{\mathbf{v}\left(\boldsymbol{\theta}_{\mathbf{0}}, n\right) - \frac{1}{H}\sum_{h=1}^{H}\mathbf{v}_h\left(\hat{\boldsymbol{\pi}}, n\right)\right\} \\ &- \left(\mathbf{Q} + \mathbf{Q}^T\right),\end{aligned} \tag{18}$$

where $\mathbf{Q} \equiv \operatorname{cov}\left\{\mathbf{b}\left(\hat{\boldsymbol{\pi}}, n\right), \mathbf{v}\left(\boldsymbol{\theta}_{\mathbf{0}}, n\right) - \frac{1}{H}\sum_{h=1}^{H}\mathbf{v}_h\left(\hat{\boldsymbol{\pi}}, n\right)\right\}$. We now investigate the three elements of (18) separately. For the first element, it is clear that $b_j\left(\hat{\boldsymbol{\pi}}, n\right) = \mathcal{O}_p\left(n^{-1}\right)$ and therefore by Cauchy-Schwarz inequality we have that $\operatorname{var}\left\{\mathbf{b}\left(\hat{\boldsymbol{\pi}}, n\right)\right\} = \mathcal{O}\left(n^{-2}\right)$ componentwise. For the second term of (18), and using $\operatorname{cov}\{\mathbf{v}\left(\boldsymbol{\theta}_{\mathbf{0}}, n\right), \mathbf{v}_h\left(\hat{\boldsymbol{\pi}}, n\right)\} = \mathbf{0}$, $h = 1, \ldots, H$, similarly to (5) in the proof of Theorem 1, we have

$$\begin{aligned}&\operatorname{var}\left\{\mathbf{v}\left(\boldsymbol{\theta}_{\mathbf{0}}, n\right) - \frac{1}{H}\sum_{h=1}^{H}\mathbf{v}_h\left(\hat{\boldsymbol{\pi}}, n\right)\right\} \\ &= \operatorname{var}\left\{\mathbf{v}\left(\boldsymbol{\theta}_{\mathbf{0}}, n\right)\right\} + \operatorname{var}\left\{\frac{1}{H}\sum_{h=1}^{H}\mathbf{v}_h\left(\hat{\boldsymbol{\pi}}, n\right)\right\} \\ &= n^{-\alpha}\mathbf{V}_{\boldsymbol{\theta}_0, n} + \frac{1}{H}\mathbb{E}\left[\operatorname{var}\left\{\mathbf{v}_1\left(\hat{\boldsymbol{\pi}}, n\right) \middle| \hat{\boldsymbol{\pi}}\right\}\right],\end{aligned} \tag{19}$$

Using Assumption 4 and performing a MacLaurin expansion on the last term of (19) we obtain

$$
\begin{aligned}
&\frac{1}{H}\mathbb{E}\left[\operatorname{var}\left\{\mathbf{v}_1\left(\hat{\boldsymbol{\pi}}, n\right)\left|\hat{\boldsymbol{\pi}}\right.\right\}\right] \\
&= \frac{1}{H}\operatorname{var}\left\{\mathbf{v}\left(\boldsymbol{\theta}_0, n\right)\right\} + \frac{1}{H}\mathbb{E}\left[\mathbf{D}_1(\boldsymbol{\theta}^*, n)(\hat{\boldsymbol{\pi}} - \boldsymbol{\theta}_0), \dots, \mathbf{D}_p(\boldsymbol{\theta}^*, n)(\hat{\boldsymbol{\pi}} - \boldsymbol{\theta}_0)\right] \\
&= \frac{1}{H}n^{-\alpha}\mathbf{V}_{\boldsymbol{\theta}_0,n} + \mathcal{O}\left(H^{-1}n^{-\min(1,\beta)}\right),
\end{aligned}
$$

where $\boldsymbol{\theta}^* \in \boldsymbol{\Theta}$ is on the line connection $\hat{\boldsymbol{\pi}}$ and $\boldsymbol{\theta}_0$. Therefore, we get

$$
\begin{aligned}
&\operatorname{var}\left\{\mathbf{v}\left(\boldsymbol{\theta}_\mathbf{0}, n\right) - \frac{1}{H}\sum_{h=1}^{H}\mathbf{v}_h\left(\hat{\boldsymbol{\pi}}, n\right)\right\} \\
&= n^{-\alpha}\left(1 + \frac{1}{H}\right)\mathbf{V}_{\boldsymbol{\theta}_0,n} + \mathcal{O}\left(H^{-1}n^{-\min(1,\beta)}\right).
\end{aligned} \tag{20}
$$

Considering the last term of (18) and using $q_{j,k}$ to denote the $(j,k)$ element of the matrix $\mathbf{Q}$, we obtain by Cauchy-Schwarz inequality that

$$
\begin{aligned}
q_{j,k}^2 &\leq \operatorname{var}\left\{b_j\left(\hat{\boldsymbol{\pi}}, n\right)\right\}\left[\operatorname{var}\left\{\mathbf{v}\left(\boldsymbol{\theta}_\mathbf{0}, n\right) - \frac{1}{H}\sum_{h=1}^{H}\mathbf{v}_h\left(\hat{\boldsymbol{\pi}}, n\right)\right\}\right]_{k,k} \\
&= \mathcal{O}\left(\max\left[n^{-(2+\alpha)}, H^{-1}n^{-\min(1,\beta)-2}\right]\right)
\end{aligned}
$$

By combining the previous results, we get

$$
\begin{aligned}
\operatorname{var}\left(\hat{\boldsymbol{\theta}}_B\right) = n^{-\alpha}&\left(1 + \frac{1}{H}\right)\mathbf{V}_{\boldsymbol{\theta}_0,n} \\
&+ \mathcal{O}\left(\max\left\{n^{-\frac{\min(2,\alpha)}{2}-1}, H^{-1}n^{-\min(1,\beta)}, H^{-1/2}n^{-\frac{\min(1,\beta)}{2}-1}\right\}\right),
\end{aligned}
$$

which verifies the second part of Theorem 3 and concludes the proof. □

## E Proof of Theorem 4

For large $H$, $\bar{\boldsymbol{\pi}}(\boldsymbol{\theta}, n) = \boldsymbol{\pi}(\boldsymbol{\theta}, n) + O_p(H^{-1/2})$. Let $\mathbf{F}(\boldsymbol{\theta}) = \hat{\boldsymbol{\pi}}(\boldsymbol{\theta}_0, n) + \boldsymbol{\theta} - \bar{\boldsymbol{\pi}}(\boldsymbol{\theta}, n)$, then for $\boldsymbol{\theta_1} \neq \boldsymbol{\theta_2}$,

$$\begin{aligned}
& \|\mathbf{F}(\boldsymbol{\theta_1}) - \mathbf{F}(\boldsymbol{\theta_2})\| \\
= \; & \|(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2) - (\boldsymbol{\pi}(\boldsymbol{\theta}_1, n) - \boldsymbol{\pi}(\boldsymbol{\theta}_2, n)) + O_p(H^{-1/2})\| \\
= \; & \|\boldsymbol{d}(\boldsymbol{\theta}_1, n) - \boldsymbol{d}(\boldsymbol{\theta}_2, n)\| + O_p(H^{-1/2}) \\
= \; & \|\boldsymbol{b}(\boldsymbol{\theta}_1, n) - \boldsymbol{b}(\boldsymbol{\theta}_2, n)\| + O_p(H^{-1/2}) \\
= \; & \left\| \left[ \sum_{i=1}^{p} \frac{a_{i,j}}{n} (\theta_{1i} - \theta_{2i}) + \sum_{k=1}^{p} \sum_{l=1}^{p} \frac{d_{k,l,j}}{n^2} (\theta_{1k}\theta_{1l} - \theta_{2k}\theta_{2l}) + \mathcal{O}\left(n^{-3}\right) \right]_{j=1,\ldots,p} \right\| \\
& + O_p(H^{-1/2}) \\
< \; & \delta \|\boldsymbol{\theta_1} - \boldsymbol{\theta_2}\|.
\end{aligned}$$

with $0 \leq \delta < 1$ for sufficiently large $n$. Thus, taking into account that $\boldsymbol{\Theta}$ is compact, Banach fixed point theorem ensures that $\mathbf{F}(\tilde{\boldsymbol{\theta}}_B^{(k-1)})$ converges, i.e. $\tilde{\boldsymbol{\theta}}_B^{(k)}$ indeed converges.

At convergence of the iterative bootstrap procedure, we have that

$$\tilde{\boldsymbol{\theta}}_B^{(k-1)} = \tilde{\boldsymbol{\theta}}_B^{(k)} \equiv \tilde{\boldsymbol{\theta}}_B.$$

(8) then yields

$$\tilde{\boldsymbol{\theta}}_B = \hat{\boldsymbol{\pi}}(\boldsymbol{\theta}_0, n) + \left(\tilde{\boldsymbol{\theta}}_B - \bar{\boldsymbol{\pi}}\left(\tilde{\boldsymbol{\theta}}_B, n\right)\right),$$

which leads to

$$\hat{\boldsymbol{\pi}}(\boldsymbol{\theta}_0, n) - \bar{\boldsymbol{\pi}}\left(\tilde{\boldsymbol{\theta}}_B, n\right) = \mathbf{0}.$$

This concludes the proof. □

# F Generalized Linear Latent Variable Models Simulations

Parameters values for the GLLVM simulation study (ordinal case) are given in Table 1. The following algorithm has been used to produce the simulation based estimator:

1. Compute the starting estimator $\hat{\boldsymbol{\pi}}$ on the original data set
   (a) Using Factor Analysis on the data as if they were multivariate normal, compute the Bartlett scores (Bartlett, 1950) from the loadings obtained with the Factor Analysis
   (b) Compute the thresholds using the empirical cumulative log-odds, i.e. the one obtained at $\mathbf{z}_{(2)} = \mathbf{0}$ as $\hat{\lambda}_{0,s}^{(j)} = \log\left(\frac{\hat{\gamma}_{0s}^{(j)}}{1-\hat{\gamma}_{0s}^{(j)}}\right)$ with $\hat{\gamma}_{0s}^{(j)} = \frac{1}{n}\sum_{i=1}^{n} \iota(x_i^{(j)} \leq s)$
   (c) Using the Bartlett scores as covariates and fixing the thresholds to the values obtained in (1b) compute ordinal GLM estimates for the loadings
2. Fix the starting value for the parameters $\tilde{\boldsymbol{\theta}}_B^{(0)} = \hat{\boldsymbol{\pi}}$
3. Fix $k = 1$, set the seed
4. For each of the $H$ samples
   (a) Generate $q$ vectors of $\boldsymbol{z}^{(k)}, k = 1, \ldots, q$ of size $n$ from a standard normal distribution

| $\boldsymbol{\lambda}_{01}$ | $\boldsymbol{\lambda}_{02}$ | $\boldsymbol{\lambda}_{03}$ | $\boldsymbol{\lambda}_{04}$ | $\boldsymbol{\lambda}_{1}$ | $\boldsymbol{\lambda}_{2}$ | $\boldsymbol{\lambda}_{3}$ |
|---|---|---|---|---|---|---|
| -1.5 | 0.4 | 1.4 | 2.5 | 1.75 | 0 | -1.0 |
| -1.5 | 0.4 | 1.4 | 2.5 | 1.0 | 0 | 2.0 |
| -1.5 | 0.4 | 1.4 | 2.5 | -2.0 | 2.3 | 1.0 |
| -1.5 | 0.4 | 1.4 | 2.5 | 1.2 | 0 | -1.0 |
| -1.5 | 0.4 | 1.4 | 2.5 | -1.6 | -1.6 | 0.9 |
| -1.5 | 0.4 | 1.4 | 2.5 | 1.0 | 0 | -0.7 |
| -1.5 | 0.4 | 1.4 | 2.5 | 0 | -0.7 | 0 |
| -1.5 | 0.4 | 1.4 | 2.5 | -1.2 | 2.7 | 0 |
| -1.5 | 0.4 | 1.4 | 2.5 | 0 | 2.0 | 1.0 |
| -1.5 | 0.4 | 1.4 | 2.5 | 0 | 1.5 | 0 |
| -1.5 | 0.4 | 1.4 | 2.5 | 1.3 | -1.0 | -0.9 |
| -1.5 | 0.4 | 1.4 | 2.5 | -0.8 | 0.8 | 2.0 |
| -1.5 | 0.4 | 1.4 | 2.5 | 0 | -1.8 | 0 |
| -1.5 | 0.4 | 1.4 | 2.5 | 0 | 0 | 1.6 |
| -1.5 | 0.4 | 1.4 | 2.5 | 2.0 | 0 | -1.8 |

TABLE 1: *True parameter values used in the simulation study*

(b) Generate the probabilities $\gamma_1^{(j)}, \ldots, \gamma_m^{(j)}$ from (25) using $\tilde{\boldsymbol{\theta}}_B^{(k-1)}$

(c) Generate the responses $x_i^{(j)}$ based on the probabilities $\gamma_1^{(j)}, \ldots, \gamma_m^{(j)}$.

5. Get the inconsistent estimator $\hat{\boldsymbol{\pi}}_h(\tilde{\boldsymbol{\theta}}_B^{(k-1)})$ for each of the $H$ samples, using the procedure in (1)

6. Compute $\bar{\boldsymbol{\pi}}(\tilde{\boldsymbol{\theta}}_B^{(k-1)}, n) = \frac{1}{H}\sum_{h=1}^{H} \hat{\boldsymbol{\pi}}_h(\tilde{\boldsymbol{\theta}}_B^{(k-1)}, n)$

7. Update $\tilde{\boldsymbol{\theta}}_B^{(k)} = \hat{\boldsymbol{\pi}} + \left(\tilde{\boldsymbol{\theta}}_B^{(k-1)} - \bar{\boldsymbol{\pi}}(\tilde{\boldsymbol{\theta}}_B^{(k-1)}, n)\right)$

8. Update $k \leftarrow k+1$

9. Use a convergence criterion; if the conditions of convergence are met stop; otherwise start again from Step (4)

# References

Bartlett, M. S. Tests of significance in factor analysis. *British Journal of Statistical Psychology*, 3:77–85, 1950.

Newey, W. K. and McFadden, D. Large Sample Estimation and Hypothesis Testing. In R. F. Engle and D. McFadden, editors, *Handbook of Econometrics*, volume 4, pages 2111–2245. Elsevier, 1994.