Supplemental Appendix to Identification and Inference of Olley and Pakes' (1996) Estimator of Production Function

Jinyong Hahn
Department of Economics, UCLA

Zhipeng Liao
Department of Economics, UCLA

Geert Ridder
Department of Economics, USC

October 5, 2021

Abstract

This supplemental appendix contains additional technical details. Section SA provides detailed description of the three-step estimator mentioned in the main text. Section SB derives the asymptotic properties of the three-step estimator and provides consistent estimation of its asymptotic variance. The detailed proofs of the asymptotic properties of the three-step estimator and the consistency of the asymptotic variance estimator are included in Section SC.

SA The Three-step Series Estimator

In this section, we describe the three-step procedure on estimating $\beta_{k,0}$. The model can be rewritten as

$$y_{1,i} = l_{1,i}\beta_{l,0} + \phi\left(i_{1,i}, k_{1,i}\right) + \eta_{1,i},$$
$$y^*_{2,i} = k_{2,i}\beta_{k,0} + g(\omega_{1,i}) + u_{2,i},$$

where $y^*_{2,i} \equiv y_{2,i} - l_{2,i}\beta_{l,0}$ and $\omega_{1,i} = \phi\left(i_{1,i}, k_{1,i}\right) - k_{1,i}\beta_{k,0}$. The following restrictions are maintained throughout the appendix

$$\mathbb{E}\left[\eta_{1,i}\middle| i_{1,i}, k_{1,i}\right] = 0 \quad \text{and} \quad \mathbb{E}\left[u_{2,i}\middle| i_{1,i}, k_{1,i}\right] = 0. \tag{SA.1}$$

For any β_k, let

$$\omega_{1,i}(\beta_k) \equiv \phi\left(i_{1,i}, k_{1,i}\right) - k_{1,i}\beta_k \quad \text{and} \quad g(\omega_{1,i}(\beta_k); \beta_k) \equiv \mathbb{E}[y^*_{2,i} - \beta_k k_{2,i}|\omega_{1,i}(\beta_k)]. \tag{SA.2}$$

Then $\omega_{1,i} = \omega_{1,i}(\beta_{k,0})$ and $g(\omega_{1,i}) = g(\omega_{1,i}(\beta_{k,0}); \beta_{k,0})$ by definition. The unknown parameters are $\beta_{l,0}$, $\beta_{k,0}$, $\phi(\cdot)$ and $g(\cdot; \beta_k)$ for any β_k in Θ_k, where Θ_k is a compact subset of $\mathbb{R}$ which contains $\beta_{k,0}$ as an interior point.

Suppose that we have data $\{(y_{t,i}, i_{t,i}, k_{t,i}, l_{t,i})_{t=1,2}\}_{i=1}^n$ and a preliminary estimator $\hat{\beta}_l$ of $\beta_{l,0}$. The asymptotic theory established here allows for a generic estimator of $\beta_{l,0}$, as long as certain regularity conditions (i.e., Assumptions SC1(iii) and SC4(i) in Section SC) hold. For example, $\hat{\beta}_l$ may be obtained from the partially linear regression proposed in Olley and Pakes (1996), or from the GMM estimation proposed in Ackerberg, Caves, and Frazer (2015). The unknown parameters $\beta_{k,0}$, $\phi(\cdot)$ and $g(\cdot; \beta_k)$ for any $\beta_k \in \Theta_k$ are estimated through the following three-step estimation procedure.

Step 1. Estimating $\phi(\cdot)$. Let $P_1(x_{1,i}) \equiv (p_{1,1}(x_{1,i}), \ldots, p_{1,m_1}(x_{1,i}))'$ be an m_1-dimensional approximating functions of $x_{1,i}$ where $x_{1,i} \equiv (i_{1,i}, k_{1,i})$. Define $\hat{y}_{1,i} \equiv y_{1,i} - l_{1,i}\hat{\beta}_l$. Then the unknown function $\phi(\cdot)$ is estimated by

$$\hat{\phi}\left(\cdot\right) \equiv P_1\left(\cdot\right)' (\mathbf{P}_1'\mathbf{P}_1)^{-1}(\mathbf{P}_1'\hat{\mathbf{Y}}_1) \tag{SA.3}$$

where $\mathbf{P}_1 \equiv (P_1(x_{1,1}), \ldots, P_1(x_{1,n}))'$ and $\hat{\mathbf{Y}}_1 \equiv (\hat{y}_{1,1}, \ldots, \hat{y}_{1,n})'$.

Step 2. Estimating $g(\cdot; \beta_k)$ for any $\beta_k \in \Theta_k$. With $\hat{\beta}_l$ and $\hat{\phi}(\cdot)$ obtained in the first step, one can estimate $y^*_{2,i}$ by $\hat{y}^*_{2,i} \equiv y_{2,i} - \hat{\beta}_l l_{2,i}$ and estimate $\omega_{1,i}(\beta_k)$ by $\hat{\omega}_{1,i}(\beta_k) \equiv \hat{\phi}(x_{1,i}) - \beta_k k_{1,i}$. Let $P_2(\omega) \equiv (p_{2,1}(\omega), \ldots, p_{2,m_2}(\omega))'$ be an m_2-dimensional approximating functions. Then $g(\cdot; \beta_k)$ is estimated by

$$\hat{g}(\cdot;\beta_k) \equiv P_2(\cdot)'\hat{\beta}_g(\beta_k), \text{ where } \hat{\beta}_g(\beta_k) \equiv (\hat{\mathbf{P}}_2(\beta_k)'\hat{\mathbf{P}}_2(\beta_k))^{-1}\hat{\mathbf{P}}_2(\beta_k)'\hat{\mathbf{Y}}_2^*(\beta_k) \tag{SA.4}$$

where $\hat{\mathbf{P}}_2(\beta_k) \equiv (P_2(\hat{\omega}_{1,1}(\beta_k)), \ldots, P_2(\hat{\omega}_{1,n}(\beta_k)))'$ and $\hat{\mathbf{Y}}_2^*(\beta_k) \equiv (\hat{y}_{2,1}^* - \beta_k k_{2,1}, \ldots, \hat{y}_{2,n}^* - \beta_k k_{2,n})'$.

Step 3. Estimating $\beta_{k,0}$. The finite dimensional parameter $\beta_{k,0}$ is estimated by $\hat{\beta}_k$ through the following semiparametric nonlinear regression

$$\hat{\beta}_k \equiv \arg\min_{\beta_k \in \Theta_k} n^{-1} \sum_{i=1}^{n} \hat{\tau}_i(\beta_k)^2, \text{ where } \hat{\tau}_i(\beta_k) \equiv \hat{y}_{2,i}^* - k_{2,i}\beta_k - \hat{g}(\hat{\omega}_{1,i}(\beta_k);\beta_k). \tag{SA.5}$$

We shall derive the root-n normality of $\hat{\beta}_k$ and provide asymptotically valid inference for $\beta_{k,0}$.

SB Asymptotic Properties of $\hat{\beta}_k$

In this section, we derive the asymptotic properties of $\hat{\beta}_k$. The consistency and the asymptotic distribution of $\hat{\beta}_k$ are presented in Subsection SB.1. In Subsection SB.2, we provide a consistent estimator of the asymptotic variance of $\hat{\beta}_k$ which can be used to construct confidence interval for $\beta_{k,0}$. Proofs of the consistency and the asymptotic normality of $\hat{\beta}_k$, and the consistency of the standard deviation estimator are included in Subsection SB.3.

SB.1 Consistency and asymptotic normality

To show the consistency of $\hat{\beta}_k$, we use the standard arguments for showing the consistency of the extremum estimator which requires two primitive conditions: (i) the identification uniqueness condition of the unknown parameter $\beta_{k,0}$; and (ii) the convergence of the estimation criterion function $n^{-1}\sum_{i=1}^{n} \hat{\tau}_i(\beta_k)^2$ to the population criterion function uniformly over $\beta_k \in \Theta_k$. We impose the identification uniqueness condition of $\beta_{k,0}$ in condition (SB.6) below, which can be verified under low-level sufficient conditions. The uniform convergence of the estimation criterion function is proved in Lemma SB1 in Subsection SB.3.

Lemma SB1. *Let $\tau_i(\beta_k) \equiv y_{2,i} - l_{2,i}\beta_{l,0} - \beta_k k_{2,i} - g(\omega_{1,i}(\beta_k);\beta_k)$ for any $\beta_k \in \Theta_k$. Suppose that for any $\varepsilon > 0$, there exists a constant $\delta_\varepsilon > 0$ such that*

$$\inf_{\{\beta_k \in \Theta_k : |\beta_k - \beta_{k,0}| \geq \varepsilon\}} \mathbb{E}\left[\tau_i(\beta_k)^2 - \tau_i(\beta_{k,0})^2\right] > \delta_\varepsilon. \tag{SB.6}$$

Then under Assumptions SC1 and SC2 in Section SC, we have $\hat{\beta}_k = \beta_{k,0} + o_p(1)$.

The asymptotic normality of $\hat{\beta}_k$ can be derived from its first-order condition:

$$n^{-1}\sum_{i=1}^{n}\hat{\tau}_i(\hat{\beta}_k)\left(k_{2,i}+\frac{\partial\hat{g}(\hat{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k)}{\partial\beta_k}\right)=0 \tag{SB.7}$$

where for any $\beta_k\in\Theta_k$

$$\frac{\partial\hat{g}(\hat{\omega}_{1,i}(\beta_k);\beta_k)}{\partial\beta_k}=\hat{\beta}_g(\beta_k)'\frac{\partial P_2(\hat{\omega}_{1,i}(\beta_k))}{\partial\beta_k}+P_2(\hat{\omega}_{1,i}(\beta_k))'\frac{\partial\hat{\beta}_g(\beta_k)}{\partial\beta_k}. \tag{SB.8}$$

By the definition of $\hat{g}(\hat{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k)$ in (SA.4), we can write

$$n^{-1}\sum_{i=1}^{n}P_2(\hat{\omega}_{1,i}(\hat{\beta}_k))\hat{g}(\hat{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k)=n^{-1}\sum_{i=1}^{n}P_2(\hat{\omega}_{1,i}(\hat{\beta}_k))(\hat{y}^*_{2,i}-k_{2,i}\hat{\beta}_k)$$

which implies that

$$n^{-1}\sum_{i=1}^{n}\hat{\tau}_i(\hat{\beta}_k)P_2(\hat{\omega}_{1,i}(\hat{\beta}_k))=0.$$

Therefore, the first-order condition (SB.7) can be reduced to

$$n^{-1}\sum_{i=1}^{n}\hat{\tau}_i(\hat{\beta}_k)\left(k_{2,i}-k_{1,i}\hat{\beta}_g(\hat{\beta}_k)'\frac{\partial P_2(\hat{\omega}_{1,i}(\hat{\beta}_k))}{\partial\omega}\right)=0 \tag{SB.9}$$

which slightly simplifies the derivation of the asymptotic normality of $\hat{\beta}_k$.

Theorem SB1. *Let $g_1(\omega)\equiv\partial g(\omega)/\partial\omega$. Suppose that*

$$\Upsilon\equiv\mathbb{E}\left[(v_{2,i}-v_{1,i}g_1(\omega_{1,i}))^2\right]>0 \tag{SB.10}$$

where $v_{j,i}\equiv k_{j,i}-\mathbb{E}[k_{j,i}|\omega_{1,i}]$ for $j=1,2$. Then under (SB.6) in Lemma SB1, and Assumptions SC1, SC2 and SC3 in Section SC

$$\begin{aligned}n^{1/2}(\hat{\beta}_k-\beta_{k,0}) &= \Upsilon^{-1}n^{-1/2}\sum_{i=1}^{n}u_{2,i}\left(v_{2,i}-v_{1,i}g_1(\omega_{1,i})\right)\\ &\quad -\Upsilon^{-1}n^{-1/2}\sum_{i=1}^{n}\eta_{1,i}g_1(\omega_{1,i})\left(v^*_{2,i}-v_{1,i}g_1(\omega_{1,i})\right)\\ &\quad -\Upsilon^{-1}\Gamma n^{1/2}(\hat{\beta}_l-\beta_{l,0})+o_p(1),\end{aligned} \tag{SB.11}$$

where $\Gamma\equiv\mathbb{E}\left[(l_{2,i}-l_{1,i}g_1(\omega_{1,i}))\left(v^*_{2,i}-v_{1,i}g_1(\omega_{1,i})\right)\right]$ and $v^*_{2,i}\equiv\mathbb{E}[k_{2,i}|x_{1,i}]-\mathbb{E}[k_{2,i}|\omega_{1,i}]$. Moreover

$$n^{1/2}(\hat{\beta}_k-\beta_{k,0})\to_d N(0,\Upsilon^{-1}\Omega\Upsilon^{-1}) \tag{SB.12}$$

where $\Omega \equiv \mathbb{E}\left[\left(u_{2,i}\left(v_{2,i} - v_{1,i}g_1(\omega_{1,i})\right) - \eta_{1,i}g_1(\omega_{1,i})\left(v_{2,i}^* - v_{1,i}g_1(\omega_{1,i})\right) - \Gamma\varepsilon_{1,i}\right)^2\right]$.

REMARK. The local identification condition of $\beta_{k,0}$ is imposed in (SB.10) which is important to ensure the root-n consistency of $\hat{\beta}_k$. □

REMARK. The random variable $\varepsilon_{1,i}$ in the definition of Ω is from the linear representation of the estimator error

$$\hat{\beta}_l - \beta_{l,0} = n^{-1}\sum_{i=1}^{n}\varepsilon_{1,i} + o_p(n^{-1/2})$$

which is maintained in Assumption SC1(iii) in Section SC. Different estimation procedures of $\hat{\beta}_l$ may give different forms for $\varepsilon_{1,i}$. Therefore, the specific form of $\varepsilon_{1,i}$ has to be derived case by case. □

REMARK. Since $\mathbb{E}[v_{j,i}|\omega_{1,i}] = 0$ for $j = 1, 2$,

$$\mathbb{E}\left[l_{2,i}\left(v_{2,i} - v_{1,i}g_1(\omega_{1,i})\right)\right] = \mathbb{E}\left[\left(l_{2,i} - \mathbb{E}[l_{2,i}|\omega_{1,i}]\right)\left(v_{2,i} - v_{1,i}g_1(\omega_{1,i})\right)\right].$$

Therefore we can write

$$\Gamma = \mathbb{E}\left[\left(l_{2,i} - \mathbb{E}[l_{2,i}|\omega_{1,i}] - h(x_{1,i})g_1(\omega_{1,i})\right)\left(v_{2,i} - v_{1,i}g_1(\omega_{1,i})\right)\right]. \tag{SB.13}$$

which is the form used in the main text of the paper. Moreover when the perpetual inventory method (PIM) i.e., $k_{2,i} = (1-\delta)\,k_{1,i} + i_{1,i}$ holds, $v_{1,i}$, $v_{2,i}$ and $\omega_{1,i}$ are functions of $x_{1,i}$. Therefore

$$\mathbb{E}\left[h(x_{1,i})g_1(\omega_{1,i})\left(v_{2,i} - v_{1,i}g_1(\omega_{1,i})\right)\right] = \mathbb{E}\left[l_{1,i}g_1(\omega_{1,i})\left(v_{2,i} - v_{1,i}g_1(\omega_{1,i})\right)\right]$$

by the law of iterated expectation. Hence we deduce that

$$\Gamma = \mathbb{E}\left[\left(l_{2,i} - l_{1,i}g_1(\omega_{1,i})\right)\left(v_{2,i} - v_{1,i}g_1(\omega_{1,i})\right)\right] \tag{SB.14}$$

under PIM. □

REMARK. From the asymptotic expansion in (SB.11), we see that the asymptotic variance of $\hat{\beta}_k$ is determined by three components. The first component, $n^{-1/2}\sum_{i=1}^{n}u_{2,i}\left(v_{2,i} - v_{1,i}g_1(\omega_{1,i})\right)$ comes from the third-step estimation with known $\omega_{1,i}$. The second and the third components are from the first-step estimation. Specifically, the second one, $n^{-1/2}\sum_{i=1}^{n}\eta_{1,i}g_1(\omega_{1,i})\left(v_{2,i} - v_{1,i}g_1(\omega_{1,i})\right)$ is from estimating $\phi(\cdot)$ in the first step, while the third component $\Gamma n^{1/2}(\hat{\beta}_l - \beta_{l,0})$ is due to the estimation error in $\hat{\beta}_l$. □

REMARK. The estimator $\hat{\beta}_k$ depends on the numbers of approximating functions, i.e., m_1 and m_2 used in estimating $\phi(\cdot)$ and $g(\cdot,)$. In practice, one may select m_1 and m_2 in a data-dependent way, such as through the cross-validation. The cross-validated series nonparametric regression estimator is shown to be asymptotically optimal in the literature (see, e.g., Li (1987) and Andrews (1991a)). Therefore, we expect that the estimator $\hat{\beta}_k$ based on the cross-validated m_1 and m_2 enjoys good asymptotic properties such as the root-n asymptotic normality in (SB1). A formal justification of this conjecture will be an interesting research topic but is beyond the scope of this paper. □

SB.2 Consistent variance estimation

The asymptotic variance of $\hat{\beta}_k$ can be estimated using its explicit form and the estimators of $v_{1,i}$, $v_{2,i}$, $\varepsilon_{1,i}$, $\eta_{1,i}$, $u_{2,i}$, $v^*_{2,i}$, $h(x_{1,i})$ and $g_1(\omega_{1,i})$. The unknown functions in $v_{1,i}$, $v_{2,i}$, $\varepsilon_{1,i}$, $\eta_{1,i}$, $u_{2,i}$, $v^*_{2,i}$, $h(x_{1,i})$ and $g_1(\omega_{1,i})$ can be estimated by the kernel or the series method. Since $\hat{\beta}_k$ is constructed using the series method and its asymptotic properties have been established in the previous subsection, we next provide the asymptotic variance estimator of $\hat{\beta}_k$ using the series method.

First, it is clear that $g_1(\omega_{1,i})$ can be estimated by $\hat{g}_1(\hat{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k)$ where

$$\hat{g}_1(\hat{\omega}_{1,i}(\beta_k);\beta_k) \equiv \hat{\beta}_g(\beta_k)' \frac{\partial P_2(\hat{\omega}_{1,i}(\beta_k))}{\partial \omega} \text{ for any } \beta_k \in \Theta_k. \tag{SB.15}$$

Second, the residual $\varsigma_i \equiv v_{2,i} - v_{1,i} g_1(\omega_{1,i})$ can be estimated by

$$\hat{\varsigma}_i \equiv \Delta\hat{k}_{2,i} - P_2(\hat{\omega}_{1,i}(\hat{\beta}_k))'(\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k))^{-1}\sum_{i=1}^{n} P_2(\hat{\omega}_{1,i}(\hat{\beta}_k))\Delta\hat{k}_{2,i}$$

where $\Delta\hat{k}_{2,i} \equiv k_{2,i} - k_{1,i}\hat{g}_1(\hat{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k)$.

Given the estimated residual $\hat{\varsigma}_i$, the Hessian term Υ in the asymptotic variance of $\hat{\beta}_k$ can be estimated by

$$\hat{\Upsilon}_n \equiv n^{-1}\sum_{i=1}^{n} \hat{\varsigma}_i^2. \tag{SB.16}$$

Moreover the Jacobian term Γ can be estimated by

$$\hat{\Gamma}_n \equiv n^{-1}\sum_{i=1}^{n}(l_{2,i} - \hat{h}_i \hat{g}_1(\hat{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k))\hat{\varsigma}_i \tag{SB.17}$$

where $\hat{h}_i = P_1(x_{1,i})'(\mathbf{P}_1'\mathbf{P}_1)^{-1}\sum_{i=1}^{n} P_1(x_{1,i}) l_{1,i}$. Define

$$\hat{u}_{2,i} \equiv \hat{y}_{2,i} - l_{2,i}\hat{\beta}_l - k_{2,i}\hat{\beta}_k - \hat{g}(\hat{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k) \quad \text{and} \quad \hat{\eta}_{1,i} \equiv y_{1,i} - l_{1,i}\hat{\beta}_l - \hat{\phi}(x_{1,i}).$$

Then Ω is estimated by

$$\hat{\Omega}_n \equiv n^{-1}\sum_{i=1}^{n}\left((\hat{u}_{2,i}-\hat{\eta}_{1,i}\hat{g}_1(\hat{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k))\hat{\varsigma}_i-\hat{\Gamma}_n\hat{\varepsilon}_{1,i}\right)^2 \tag{SB.18}$$

where $\hat{\varepsilon}_{1,i}$ denotes the estimator of $\varepsilon_{1,i}$ for $i = 1, \ldots, n$.

Theorem SB2. *Under Assumptions SC1, SC2, SC3 and SC4 in Section SC, we have*

$$\hat{\Upsilon}_n = \Upsilon + o_p(1) \quad \textit{and} \quad \hat{\Omega}_n = \Omega + o_p(1) \tag{SB.19}$$

and moreover

$$\frac{n^{1/2}(\hat{\beta}_k-\beta_{k,0})}{(\hat{\Upsilon}_n^{-1}\hat{\Omega}_n\hat{\Upsilon}_n^{-1})^{1/2}} \to_d N(0,1) \tag{SB.20}$$

where $\hat{\Omega}_n$ *is defined in (SB.18).*

SB.3 Proof of the asymptotic properties

In this subsection, we prove the main results presented in the previous subsection. Throughout this subsection, we use $C > 1$ to denote a generic finite constant which does not depend on n, m_1 or m_2 but whose value may change in different places.

Proof of Lemma SB1. By (SC.72) in the proof of Lemma SC8 and Assumption SC2(i)

$$\sup_{\beta_k\in\Theta_k} \mathbb{E}\left[\tau_i(\beta_k)^2\right] \leq C \tag{SB.21}$$

which together with Lemma SC8 implies that

$$\sup_{\beta_k\in\Theta_k} n^{-1}\sum_{i=1}^{n}\tau_i(\beta_k)^2 = O_p(1). \tag{SB.22}$$

By the Markov inequality, Assumptions SC1(i, iii) and SC2(i), we obtain

$$n^{-1}\sum_{i=1}^{n}(\hat{y}^*_{2,i}-y^*_{2,i})^2 = (\hat{\beta}_l-\beta_l)^2 n^{-1}\sum_{i=1}^{n} l_{2,i}^2 = O_p(n^{-1}). \tag{SB.23}$$

By the definition of $\hat{\tau}_i(\beta_k)$ and $\tau_i(\beta_k)$, we can write

$$\begin{aligned}
& n^{-1}\sum_{i=1}^{n}\hat{\tau}_i(\beta_k)^2 - \mathbb{E}\left[\tau_i(\beta_k)^2\right] \\
= \; & n^{-1}\sum_{i=1}^{n}\left(\tau_i(\beta_k)^2 - \mathbb{E}\left[\tau_i(\beta_k)^2\right]\right) + 2n^{-1}\sum_{i=1}^{n}\tau_i(\beta_k)(\hat{y}^*_{2,i} - y^*_{2,i}) \\
& -2n^{-1}\sum_{i=1}^{n}\tau_i(\beta_k)(\hat{g}(\hat{\omega}_{1,i}(\beta_k);\beta_k) - g(\omega_{1,i}(\beta_k);\beta_k)) \\
& -2n^{-1}\sum_{i=1}^{n}(\hat{y}^*_{2,i} - y^*_{2,i})(\hat{g}(\hat{\omega}_{1,i}(\beta_k);\beta_k) - g(\omega_{1,i}(\beta_k);\beta_k)) \\
& +n^{-1}\sum_{i=1}^{n}(\hat{y}^*_{2,i} - y^*_{2,i})^2 + n^{-1}\sum_{i=1}^{n}(\hat{g}(\hat{\omega}_{1,i}(\beta_k);\beta_k) - g(\omega_{1,i}(\beta_k);\beta_k))^2,
\end{aligned}$$

which together with Assumption SC2(vi), Lemma SC7, Lemma SC8, (SB.22), (SB.23) and the Cauchy-Schwarz inequality implies that

$$\sup_{\beta_k\in\Theta_k}\left|n^{-1}\sum_{i=1}^{n}\hat{\tau}_i(\beta_k)^2 - \mathbb{E}\left[\tau_i(\beta_k)^2\right]\right| = o_p(1). \tag{SB.24}$$

The consistency of $\hat{\beta}_k$ follows from its definition in (SA.5), (SB.24), the identification uniqueness condition of $\beta_{k,0}$ assumed in (SB.6) and the standard arguments of showing the consistency of the extremum estimator. *Q.E.D.*

Lemma SB2. *Let $g_{1,i} \equiv g_1(\omega_{1,i})$ and $\hat{J}_i(\beta_k) \equiv \hat{\tau}_i(\beta_k)\left(k_{2,i} - k_{1,i}\hat{g}_1(\hat{\omega}_{1,i}(\beta_k);\beta_k)\right)$ for any $\beta_k \in \Theta_k$, where $\hat{g}_1(\hat{\omega}_{1,i}(\beta_k);\beta_k)$ is defined in (SB.15). Then under Assumptions SC1, SC2 and SC3, we have*

$$n^{-1}\sum_{i=1}^{n}\hat{J}_i(\beta_{k,0}) = n^{-1}\sum_{i=1}^{n}(u_{2,i} - \eta_{1,i}g_{1,i})(v_{2,i} - v_{1,i}g_{1,i}) - \Gamma(\hat{\beta}_l - \beta_{l,0}) + o_p(n^{-1/2}). \tag{SB.25}$$

Proof of Lemma SB2. By the definition of $\hat{\tau}_i(\beta_{k,0})$ and Lemma SC10,

$$\begin{aligned}
& n^{-1}\sum_{i=1}^{n}\hat{\tau}_i(\beta_{k,0})\left(k_{2,i} - k_{1,i}\hat{g}_1(\hat{\omega}_{1,i}(\beta_{k,0});\beta_{k,0})\right) \\
&= n^{-1}\sum_{i=1}^{n}(\hat{y}^*_{2,i}(\beta_{k,0}) - g(\omega_{1,i}))(k_{2,i} - k_{1,i}g_{1,i}) \\
&\quad - n^{-1}\sum_{i=1}^{n}(\hat{g}(\hat{\omega}_{1,i}(\beta_{k,0});\beta_{k,0}) - g(\omega_{1,i}))(k_{2,i} - k_{1,i}g_{1,i}) + o_p(n^{-1/2})
\end{aligned} \tag{SB.26}$$

where $\hat{y}^*_{2,i}(\beta_{k,0}) \equiv y_{2,i} - l_{2,i}\hat{\beta}_l - k_{2,i}\beta_{k,0}$, and by Lemma SC12

$$\begin{aligned} & n^{-1}\sum_{i=1}^{n}(\hat{g}(\hat{\omega}_{1,i}(\beta_{k,0});\beta_{k,0}) - g(\omega_{1,i}))(k_{2,i} - k_{1,i}g_{1,i}) \\ = \; & n^{-1}\sum_{i=1}^{n} u_{2,i}\varphi(\omega_{1,i}) - \mathbb{E}[l_{2,i}\varphi(\omega_{1,i})](\hat{\beta}_l - \beta_{l,0}) \\ & + n^{-1}\sum_{i=1}^{n} g_{1,i}(\hat{\phi}(x_{1,i}) - \phi(x_{1,i}))(v_{2,i} - v_{1,i}g_{1,i}) + o_p(n^{-1/2}), \end{aligned} \tag{SB.27}$$

where $\varphi(\omega_{1,i}) \equiv \mathbb{E}[k_{2,i}|\omega_{1,i}] - \mathbb{E}[k_{1,i}|\omega_{1,i}]g_{1,i}$. By the definition of $\hat{y}^*_{2,i}(\beta_{k,0})$, we get

$$\begin{aligned} & n^{-1}\sum_{i=1}^{n}(\hat{y}^*_{2,i}(\beta_{k,0}) - g(\omega_{1,i}))(k_{2,i} - k_{1,i}g_{1,i}) \\ = \; & n^{-1}\sum_{i=1}^{n} u_{2,i}(k_{2,i} - k_{1,i}g_{1,i}) - (\hat{\beta}_l - \beta_{l,0})n^{-1}\sum_{i=1}^{n} l_{2,i}(k_{2,i} - k_{1,i}g_{1,i}) \\ = \; & n^{-1}\sum_{i=1}^{n} u_{2,i}(k_{2,i} - k_{1,i}g_{1,i}) - (\hat{\beta}_l - \beta_{l,0})\mathbb{E}[l_{2,i}(k_{2,i} - k_{1,i}g_{1,i})] + o_p(n^{-1/2}) \end{aligned} \tag{SB.28}$$

where the second equality is by Assumption SC1(iii) and

$$n^{-1}\sum_{i=1}^{n} l_{2,i}(k_{2,i} - k_{1,i}g_{1,i}) = \mathbb{E}[l_{2,i}(k_{2,i} - k_{1,i}g_{1,i})] + O_p(n^{-1/2})$$

which holds by the Markov inequality, Assumptions SC1(i) and SC2(i, ii). Therefore by (SB.26), (SB.27) and (SB.28), we obtain

$$\begin{aligned} & n^{-1}\sum_{i=1}^{n}\hat{\tau}_i(\beta_{k,0})\,(k_{2,i} - k_{1,i}\hat{g}_1(\hat{\omega}_{1,i}(\beta_{k,0});\beta_{k,0})) \\ = \; & n^{-1}\sum_{i=1}^{n} u_{2,i}(v_{2,i} - v_{1,i}g_{1,i}) - (\hat{\beta}_l - \beta_{l,0})\mathbb{E}[l_{2,i}(v_{2,i} - v_{1,i}g_{1,i})] \\ & - n^{-1}\sum_{i=1}^{n} g_{1,i}(\hat{\phi}(x_{1,i}) - \phi(x_{1,i}))(v_{2,i} - v_{1,i}g_{1,i}) + o_p(n^{-1/2}). \end{aligned} \tag{SB.29}$$

The claim of the lemma follows from (SB.29) and Lemma SC13. *Q.E.D.*

Lemma SB3. *Under Assumptions SC1, SC2 and SC3, we have*

$$n^{-1}\sum_{i=1}^{n}(\hat{J}_i(\hat{\beta}_k) - \hat{J}_i(\beta_{k,0})) = -(\hat{\beta}_k - \beta_{k,0})\left[\mathbb{E}[(v_{2,i} - v_{1,i}g_{1,i})^2] + o_p(1)\right] + o_p(n^{-1/2}).$$

Proof of Lemma SB3. First note that by the definition of $\hat{J}_i(\beta_k)$ and $\hat{\tau}_i(\beta_k)$, we can write

$$\begin{aligned}
& n^{-1}\sum_{i=1}^{n}(\hat{J}_i(\hat{\beta}_k)-\hat{J}_i(\beta_{k,0})) \\
= \; & -(\hat{\beta}_k-\beta_{k,0})n^{-1}\sum_{i=1}^{n}k_{2,i}(k_{2,i}-k_{1,i}\hat{g}_1(\hat{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k)) \\
& -n^{-1}\sum_{i=1}^{n}(\hat{g}(\hat{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k)-\hat{g}(\hat{\omega}_{1,i}(\beta_{k,0});\beta_{k,0}))(k_{2,i}-k_{1,i}\hat{g}_1(\hat{\omega}_{1,i}(\beta_{k,0});\beta_{k,0})) \\
& -n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(\hat{g}_1(\hat{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k)-\hat{g}_1(\hat{\omega}_{1,i}(\beta_{k,0});\beta_{k,0})) \\
& +(\hat{\beta}_l-\beta_{l,0})n^{-1}\sum_{i=1}^{n}l_{2,i}k_{1,i}(\hat{g}_1(\hat{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k)-\hat{g}_1(\hat{\omega}_{1,i}(\beta_{k,0});\beta_{k,0})) \qquad \text{(SB.30)}
\end{aligned}$$

which together with Assumption SC1(iii), Lemma SC17, Lemma SC21 and Lemma SC23 implies that

$$\begin{aligned}
n^{-1}\sum_{i=1}^{n}(\hat{J}_i(\hat{\beta}_k)-\hat{J}_i(\beta_{k,0})) &= -(\hat{\beta}_k-\beta_{k,0})\mathbb{E}[k_{2,i}(k_{2,i}-k_{1,i}g_{1,i})] \\
&\quad +(\hat{\beta}_k-\beta_{k,0})\left[\mathbb{E}[k_{1,i}g_{1,i}(v_{2,i}-v_{1,i}g_{1,i})]+\mathbb{E}[k_{2,i}\varphi(\omega_{1,i})]\right] \\
&\quad +(\hat{\beta}_k-\beta_{k,0})o_p(1)+o_p(n^{-1/2}) \\
&= -(\hat{\beta}_k-\beta_{k,0})\left[\mathbb{E}[(v_{2,i}-v_{1,i}g_{1,i})^2]+o_p(1)\right]+o_p(n^{-1/2})
\end{aligned}$$

which finishes the proof. *Q.E.D.*

Proof of Theorem SB1. By Assumptions SC1(ii, iii) and SC2(i, ii), and Hölder's inequality

$$\Gamma=\mathbb{E}\left[(l_{2,i}-h_i g_1(\omega_{1,i}))\,(v_{2,i}-v_{1,i}g_1(\omega_{1,i}))\right]\leq C \qquad \text{(SB.31)}$$

and

$$\begin{aligned}
\Omega &= \mathbb{E}[((u_{2,i}-\eta_{1,i}g_1(\omega_{1,i}))(v_{2,i}-v_{1,i}g_1(\omega_{1,i}))-\Gamma\varepsilon_{1,i})^2] \\
&\leq C\mathbb{E}[u_{2,i}^4+\eta_{1,i}^4+v_{1,i}^4+v_{2,i}^4+\varepsilon_{1,i}^2]\leq C. \qquad \text{(SB.32)}
\end{aligned}$$

By Assumption SC1(i), (SB.32) and the Lindeberg–Lévy central limit theorem,

$$n^{-1/2}\sum_{i=1}^{n}((u_{2,i}-\eta_{1,i}g_1(\omega_{1,i}))(v_{2,i}-v_{1,i}g_1(\omega_{1,i}))-\Gamma\varepsilon_{1,i})\to_d N(0,\Omega). \qquad \text{(SB.33)}$$

By (SB.9), Assumption SC1(iii), Lemma SB2 and Lemma SB3, we can write

$$\begin{aligned} 0 &= n^{-1}\sum_{i=1}^{n}\hat{J}_i(\beta_{k,0}) + n^{-1}\sum_{i=1}^{n}(\hat{J}_i(\hat{\beta}_k) - \hat{J}_i(\beta_{k,0})) \\ &= n^{-1}\sum_{i=1}^{n}(u_{2,i} - \eta_{1,i}g_1(\omega_{1,i}))(v_{2,i} - v_{1,i}g_1(\omega_{1,i})) - \Gamma n^{1/2}(\hat{\beta}_l - \beta_{l,0}) \\ &\quad -(\hat{\beta}_k - \beta_{k,0})\left[\mathbb{E}[(v_{2,i} - v_{1,i}g_1(\omega_{1,i}))^2] + o_p(1)\right] + o_p(n^{-1/2}) \end{aligned} \tag{SB.34}$$

which together with (SB.10) and (SB.33) implies that

$$n^{1/2}(\hat{\beta}_k - \beta_{k,0}) = \Upsilon^{-1}n^{-1/2}\sum_{i=1}^{n}(u_{2,i} - \eta_{1,i}g_1(\omega_{1,i}))(v_{2,i} - v_{1,i}g_1(\omega_{1,i})) - \Upsilon^{-1}\Gamma n^{1/2}(\hat{\beta}_l - \beta_{l,0}) + o_p(1). \tag{SB.35}$$

This proves (SB.11). The claim in (SB.12) follows from Assumption SC1(iii), (SB.33) and (SB.35). *Q.E.D.*

Proof of Theorem SB2. The results in (SB.19) are proved in Lemma SC25(i, iii), which together with Theorem SB1, Assumption SC4(iii) and the Slutsky Theorem proves the claim in (SB.20). *Q.E.D.*

SC Auxiliary Results

In this section, we provide the auxiliary results which are used to show Lemma SB1, Theorem SB1 and Theorem SB2. The following notations are used throughout this section. We use $\|\cdot\|_2$ to denote the L_2-norm under the joint distribution of $(y_{t,i}, i_{t,i}, k_{t,i}, l_{t,i})_{t=1,2}$, $\|\cdot\|$ to denote the Euclidean norm and $\|\cdot\|_S$ to denote the matrix operator norm. For any real symmetric square matrix A, we use $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$ to denote the smallest and largest eigenvalues of A respectively. Throughout this appendix, we use $C > 1$ to denote a generic finite constant which does not depend on n, m_1 or m_2 but whose value may change in different places.

SC.1 The asymptotic properties of the first-step estimators

Let $Q_{m_1} \equiv \mathbb{E}[P_1(x_{1,i})P_1(x_{1,i})']$. The following assumptions are needed for studying the first-step estimator $\hat{\phi}(\cdot)$.

Assumption SC1. *(i) The data* $\{(y_{t,i}, i_{t,i}, k_{t,i}, l_{t,i})_{t=1,2}\}_{i=1}^n$ *are i.i.d.; (ii)* $\mathbb{E}[\eta_{1,i}|\, x_{1,i}] = 0$ *and* $\mathbb{E}[l_{1,i}^2 + \eta_{1,i}^4\big|\, x_{1,i}] \leq C$*; (iii) there exist i.i.d. random variables* $\varepsilon_{1,i}$ *with* $\mathbb{E}[\varepsilon_{1,i}^4] \leq C$ *such that*

$$\hat{\beta}_l - \beta_{l,0} = n^{-1}\sum_{i=1}^{n}\varepsilon_{1,i} + o_p(n^{-1/2});$$

(iv) there exist $r_\phi > 0$ *and* $\beta_{\phi,m} \in \mathbb{R}^m$ *such that* $\sup_{x\in\mathcal{X}} |\phi_m(x) - \phi(x)| = O(m^{-r_\phi})$ *where* $\phi_m(x) \equiv P_1(x)' \beta_{\phi,m}$ *and* $\mathcal{X}$ *denotes the support of* $x_{1,i}$ *which is compact; (v)* $C^{-1} \leq \lambda_{\min}(Q_{m_1}) \leq \lambda_{\max}(Q_{m_1}) \leq C$ *uniformly over* m_1*; (vi)* $m_1^2 n^{-1} + n^{1/2} m_1^{-r_\phi} = O(1)$ *and* $\log(m_1)\xi_{0,m_1}^2 n^{-1} = o(1)$ *where* ξ_{0,m_1} *is a nondecreasing sequence such that* $\sup_{x\in\mathcal{X}} \|P_1(x)\| \leq \xi_{0,m_1}$*.*

Assumption SC1(iii) assumes that there exists a root-n consistent estimator $\hat{\beta}_l$ of $\beta_{l,0}$. Different estimation procedures of $\hat{\beta}_l$ may give different forms for $\varepsilon_{1,i}$. For example, $\hat{\beta}_l$ may be obtained together with the nonparametric estimator of $\phi(\cdot)$ in the partially linear regression proposed in Olley and Pakes (1996), or from the GMM estimation proposed in Ackerberg, Caves, and Frazer (2015). Therefore, the specific form of $\varepsilon_{1,i}$ has to be derived case by case. The rest conditions in Assumption SC1 are fairly standard in series estimation; see, for example, Andrews (1991b), Newey (1997) and Chen (2007).[1] In particular, condition (iv) specifies the precision for approximating the unknown function $\phi(\cdot)$ via approximating functions, for which comprehensive results are available from numerical approximation theory.

The properties of the first-step estimator $\hat{\phi}(\cdot)$ are presented in the following lemma.

Lemma SC4. *Under Assumption SC1, we have*

$$n^{-1}\sum_{i=1}^{n} |\hat{\phi}(x_{1,i}) - \phi(x_{1,i})|^2 = O_p(m_1 n^{-1}) \tag{SC.36}$$

and moreover

$$\sup_{x_1\in\mathcal{X}} |\hat{\phi}(x_1) - \phi(x_1)| = O_p(\xi_{0,m_1} m_1^{1/2} n^{-1/2}). \tag{SC.37}$$

Proof of Lemma SC4. Under Assumption SC1(i, v, vi), we can invoke Lemma 6.2 in Belloni, Chernozhukov, Chetverikov, and Kato (2015) to obtain

$$\left\| n^{-1}\mathbf{P}_1'\mathbf{P}_1 - Q_{m_1} \right\|_S = O_p((\log m_1)^{1/2}\xi_{0,m_1} n^{-1/2}) = o_p(1) \tag{SC.38}$$

which together with Assumption SC1(v) implies that

$$C^{-1} \leq \lambda_{\min}(n^{-1}\mathbf{P}_1'\mathbf{P}_1) \leq \lambda_{\max}(n^{-1}\mathbf{P}_1'\mathbf{P}_1) \leq C \tag{SC.39}$$

uniformly over m_1 with probability approaching 1 (wpa1). Since $\hat{y}_{1,i} = y_{1,i} - l_{1,i}\hat{\beta}_l = \phi(x_{1,i}) +$

[1] For some approximating functions such as power series, Assumptions SC1(v, vi) hold under certain nonsingular transformation on the vector approximating functions $P_1(\cdot)$, i.e., $BP_1(\cdot)$, where B is some non-singular constant matrix. Since the nonparametric series estimator is invariant to any nonsingular transformation of $P_1(\cdot)$, we do not distinguish between $BP_1(\cdot)$ and $P_1(\cdot)$ throughout this appendix.

$\eta_{1,i} - l_{1,i}(\hat{\beta}_l - \beta_{l,0})$, we can write

$$\begin{aligned}\hat{\beta}_\phi - \beta_{\phi,m_1} &= (\mathbf{P}_1'\mathbf{P}_1)^{-1}\sum_{i=1}^n P_1(x_{1,i})\eta_{1,i} \\ &\quad +(\mathbf{P}_1'\mathbf{P}_1)^{-1}\sum_{i=1}^n P_1(x_{1,i})(\phi(x_{1,i}) - \phi_{m_1}(x_{1,i})) \\ &\quad -(\hat{\beta}_l - \beta_{l,0})(\mathbf{P}_1'\mathbf{P}_1)^{-1}\sum_{i=1}^n P_1(x_{1,i})l_{1,i}. \qquad \text{(SC.40)}\end{aligned}$$

By Assumption SC1(i, ii, v) and the Markov inequality

$$n^{-1}\sum_{i=1}^n P_1(x_{1,i})\eta_{1,i} = O_p(m_1^{1/2}n^{-1/2}) \qquad \text{(SC.41)}$$

which together with Assumption SC1(vi), (SC.38) and (SC.39) implies that

$$\left[(n^{-1}\mathbf{P}_1'\mathbf{P}_1)^{-1} - Q_{m_1}^{-1}\right] n^{-1}\sum_{i=1}^n P_1(x_{1,i})\eta_{1,i} = O_p((\log m_1)^{1/2}\xi_{0,m_1}m_1^{1/2}n^{-1}) = o_p(n^{-1/2}). \qquad \text{(SC.42)}$$

By Assumption SC1(iv, vi) and (SC.39)

$$(\mathbf{P}_1'\mathbf{P}_1)^{-1}\sum_{i=1}^n P_1(x_{1,i})(\phi(x_{1,i}) - \phi_{m_1}(x_{1,i})) = O_p(m^{-r_\phi}) = O_p(n^{-1/2}). \qquad \text{(SC.43)}$$

Under Assumption SC1(i, ii, v, vi), we can use similar arguments in showing (SC.41) to get

$$n^{-1}\sum_{i=1}^n P_1(x_{1,i})l_{1,i} - \mathbb{E}[P_1(x_{1,i})l_{1,i}] = O_p(m_1^{1/2}n^{-1/2}) = o_p(1). \qquad \text{(SC.44)}$$

By Assumption SC1(i, ii, v),

$$\|\mathbb{E}[l_{1,i}P_1(x_{1,i})]\|^2 \leq \lambda_{\max}(Q_{m_1})\mathbb{E}[l_{1,i}P_1(x_{1,i})']Q_{m_1}^{-1}\mathbb{E}[P_1(x_{1,i})l_{1,i}] \leq C\mathbb{E}[l_{1,i}^2] \leq C \qquad \text{(SC.45)}$$

which combined with (SC.44) implies that

$$n^{-1}\sum_{i=1}^n P_1(x_{1,i})l_{1,i} = O_p(1). \qquad \text{(SC.46)}$$

By Assumption SC1(iii, v, vi), (SC.38), (SC.44), (SC.45) and (SC.46),

$$(\hat{\beta}_l - \beta_{l,0})(\mathbf{P}_1'\mathbf{P}_1)^{-1}\sum_{i=1}^n P_1(x_{1,i})l_{1,i} = Q_{m_1}^{-1}\mathbb{E}[P_1(x_{1,i})l_{1,i}](\hat{\beta}_l - \beta_{l,0}) + O_p(n^{-1/2})$$

which combined with Assumption SC1(vi), (SC.40), (SC.42) and (SC.43) shows that

$$\hat{\beta}_\phi - \beta_{\phi,m_1} = Q_{m_1}^{-1}\left(\sum_{i=1}^n P_1(x_{1,i})\eta_{1,i} - \mathbb{E}[P_1(x_{1,i})l_{1,i}](\hat{\beta}_l - \beta_{l,0})\right) + O_p(n^{-1/2}) = O_p(m_1^{1/2}n^{-1/2}) \tag{SC.47}$$

where the second equality follows from Assumptions SC1(iii, v), (SC.41) and (SC.45). By the Cauchy-Schwarz inequality

$$\begin{aligned}
n^{-1}\sum_{i=1}^n |\hat{\phi}(x_{1,i}) - \phi(x_{1,i})|^2 &\leq 2n^{-1}\sum_{i=1}^n |\hat{\phi}(x_{1,i}) - \phi_{m_1}(x_{1,i})|^2 + 2n^{-1}\sum_{i=1}^n |\phi_{m_1}(x_{1,i}) - \phi(x_{1,i})|^2 \\
&\leq 2\lambda_{\max}(n^{-1}\mathbf{P}_1'\mathbf{P}_1)\left\|\hat{\beta}_\phi - \beta_{\phi,m_1}\right\|^2 + 2\sup_{x\in\mathcal{X}_1}|\phi_{m_1}(x) - \phi(x)| = O_p(m_1^{1/2}n^{-1/2}) \qquad \text{(SC.48)}
\end{aligned}$$

where the equality is by Assumptions SC1(iv, vi), (SC.39) and (SC.47), which proves (SC.36). By the triangle inequality, the Cauchy-Schwarz inequality, Assumption SC1(iv, vi) and (SC.47)

$$\begin{aligned}
\sup_{x_1\in\mathcal{X}}|\hat{\phi}(x_1) - \phi(x_1)| &\leq \sup_{x_1\in\mathcal{X}}|\hat{\phi}(x_1) - \phi_{m_1}(x_1)| + \sup_{x_1\in\mathcal{X}}|\phi_{m_1}(x_1) - \phi(x_1)| \\
&\leq \xi_{0,m_1}\left\|\hat{\beta}_\phi - \beta_{\phi,m_1}\right\| + O(m_1^{-r_\phi}) = O_p(\xi_{0,m_1}m_1^{1/2}n^{-1/2}) \qquad \text{(SC.49)}
\end{aligned}$$

which proves the claim in (SC.36). *Q.E.D.*

SC.2 Auxiliary results for the consistency of $\hat{\beta}_k$

Recall that $\omega_{1,i}(\beta_k) \equiv \phi(x_{1,i}) - \beta_k k_{1,i}$ and $g(\omega;\beta_k) \equiv \mathbb{E}[y_{2,i}^* - \beta_k k_{2,i}|\omega_{1,i}(\beta_k) = \omega]$. For any $\beta_k \in \Theta_k$, let $\Omega(\beta_k) \equiv [a_{\beta_k}, b_{\beta_k}]$ denote the support of $\omega_{1,i}(\beta_k)$ with $c_\omega < a_{\beta_k} < b_{\beta_k} < C_\omega$, where c_ω and C_ω are finite constants. Define $\Omega_\varepsilon(\beta_k) \equiv [a_{\beta_k} - \varepsilon, b_{\beta_k} + \varepsilon]$ for any constant $\varepsilon > 0$. For an integer $d \geq 0$, let $|g(\beta_k)|_d = \max_{0\leq j\leq d}\sup_{\omega\in\Omega(\beta_k)}\left|\partial^j g(\omega;\beta_k)/\partial\omega^j\right|$.

Assumption SC2. *(i)* $\mathbb{E}[(y_{2,i}^*)^4 + l_{2,i}^4 + k_{2,i}^4|x_{1,i}] \leq C$*; (ii)* $g(\omega;\beta_k)$ *is continuously differentiable with uniformly bounded derivatives; (iii) for some* $d \geq 1$ *there exist* $\beta_{g,m_2}(\beta_k) \in \mathbb{R}^{m_2}$ *and* $r_g > 0$ *such that* $\sup_{\beta_k\in\Theta_k}|g(\beta_k) - g_{m_2}(\beta_k)|_d = O(m_2^{-r_g})$ *where* $g_{m_2}(\omega;\beta_k) \equiv P_2(\omega)'\beta_{g,m_2}(\beta_k)$*; (iv) for any* $\beta_k \in \Theta_k$ *there exists a nonsingular matrix* $B(\beta_k)$ *such that for* $\tilde{P}_2(\omega_1(\beta_k);\beta_k) \equiv B(\beta_k)P_2(\omega_1(\beta_k))$*,*

$$C^{-1} \leq \lambda_{\min}(Q_{m_2}(\beta_k)) \leq \lambda_{\max}(Q_{m_2}(\beta_k)) \leq C$$

uniformly over $\beta_k \in \Theta_k$*, where* $Q_{m_2}(\beta_k) \equiv \mathbb{E}[\tilde{P}_2(\omega_1(\beta_k);\beta_k)\,\tilde{P}_2(\omega_1(\beta_k);\beta_k)']$*; (v) for* $j = 0, 1, 2, 3$*, there exist sequences* ξ_{j,m_2} *such that* $\sup_{\beta_k\in\Theta_k}\sup_{\omega\in\Omega_\varepsilon(\beta_k)}\left\|\partial^j \tilde{P}_2(\omega;\beta_k)/\partial\omega^{j_1}\partial\beta_k^{j-j_1}\right\| \leq \xi_{j,m_2}$ *where* $j_1 \leq j$ *and* $\varepsilon = m_2^{-2}$*; (vi)* $\xi_{j,m_2} \leq Cm_2^{j+1}$ *and* $\xi_{0,m_1}(m_1^{1/2}m_2^3 + (\log(n))^{1/2})n^{-1/2} + n^{1/2}m_2^{-r_g} = o(1)$*.*

Assumption SC2(i) imposes upper bound on the conditional moments of $y_{2,i}^*$, $l_{2,i}$ and $k_{2,i}$ given $x_{1,i}$. Assumptions SC2(ii, iii) require that the conditional moment function $g(\omega;\beta_k)$ is smooth and can be well approximated by linear combinations of $P_2(\omega)$. Assumption SC2(iv) imposes normalization on the approximating functions $P_2(\omega)$, and uniform lower and upper bounds on the eigenvalues of $Q_{m_2}(\beta_k)$. Assumption SC2(v, vi) restrict the magnitudes of the normalized approximating functions and their derivatives, and the convergence rate of the series approximation error.

Since the series estimator $\hat{g}(\hat{\omega}_{1,i}(\beta_k);\beta_k) = P_2(\hat{\omega}_{1,i}(\beta_k))'\hat{\beta}_g(\beta_k)$ is invariant to any non-singular transformation on $P_2(\omega)$, throughout the rest of the Appendix we let

$$\tilde{\mathbf{P}}_2(\beta_k) \equiv (\tilde{P}_{2,1}(\beta_k), \ldots, \tilde{P}_{2,n}(\beta_k))' \quad \text{and} \quad \hat{\mathbf{P}}_2(\beta_k) \equiv (\hat{P}_{2,1}(\beta_k), \ldots, \hat{P}_{2,n}(\beta_k))'$$

where $\tilde{P}_{2,i}(\beta_k) \equiv B(\beta_k)P_2(\omega_{1,i}(\beta_k))$, $\hat{P}_{2,i}(\beta_k) \equiv B(\beta_k)P_2(\hat{\omega}_{1,i}(\beta_k))$ and $\hat{\omega}_{1,i}(\beta_k) \equiv \hat{\phi}(x_{1,i}) - k_{1,i}\beta_k$.[2] Define

$$\partial^j \tilde{P}_2(\omega;\beta_k) \equiv \frac{\partial^j \tilde{P}_2(\omega;\beta_k)}{\partial\omega^j} \quad \text{and} \quad \partial^j \tilde{P}_{2,i}(\beta_k) \equiv \partial^j \tilde{P}_2(\omega_{1,i}(\beta_k);\beta_k)$$

for $j = 1, 2, 3$ and $i = 1, \ldots, n$.

Lemma SC5. *Under Assumptions SC1 and SC2, we have*

$$\sup_{\beta_k\in\Theta_k}\left\|n^{-1}\hat{\mathbf{P}}_2(\beta_k)'\hat{\mathbf{P}}_2(\beta_k) - n^{-1}\tilde{\mathbf{P}}_2(\beta_k)'\tilde{\mathbf{P}}_2(\beta_k)\right\|_S = O_p(\xi_{1,m_2}m_1^{1/2}n^{-1/2}).$$

Proof of Lemma SC5. Since $\hat{\omega}_{1,i}(\beta_k) = \hat{\phi}(x_{1,i}) - \beta_k k_{1,i}$, by Lemma SC4

$$\sup_{\beta_k\in\Theta_k}\max_{i\leq n}|\hat{\omega}_{1,i}(\beta_k) - \omega_{1,i}(\beta_k)| = \max_{i\leq n}|\hat{\phi}(x_{1,i}) - \phi(x_{1,i})| = O_p(\xi_{0,m_1}m_1^{1/2}n^{-1/2}) = o_p(1) \quad \text{(SC.50)}$$

which together with Assumption SC2(vi) implies that

$$\hat{\omega}_{1,i}(\beta_k) \in \Omega_\varepsilon(\beta_k) \text{ wpa1} \quad \text{(SC.51)}$$

for any $i \leq n$ and uniformly over $\beta_k \in \Theta_k$. By the mean value expansion, we have for any $\upsilon_2 \in \mathbb{R}^{m_2}$

$$\left|\upsilon_2'(\tilde{P}_{2,i}(\beta_k) - \hat{P}_{2,i}(\beta_k))\right| = \left|\upsilon_2'\partial^1\tilde{P}_2(\tilde{\omega}_{1,i}(\beta_k);\beta_k)(\hat{\omega}_{1,i}(\beta_k) - \omega_{1,i}(\beta_k))\right| \quad \text{(SC.52)}$$

[2] Note that we define $\hat{P}_{2,i}(\beta_k) \equiv P_2(\hat{\omega}_{1,i}(\beta_k))$ in Section SA. We change its definition here since the asymptotic properties of the sereis estimator $\hat{g}(\hat{\omega}_{1,i}(\beta_k);\beta_k) = P_2(\hat{\omega}_{1,i}(\beta_k))'\hat{\beta}_g(\beta_k)$ shall be investigated under the new definition $\hat{P}_{2,i}(\beta_k) \equiv B(\beta_k)P_2(\hat{\omega}_{1,i}(\beta_k))$.

where $\tilde{\omega}_{1,i}(\beta_k)$ lies between $\omega_{1,i}(\beta_k)$ and $\hat{\omega}_{1,i}(\beta_k)$. Since $\omega_{1,i}(\beta_k)$ and $\hat{\omega}_{1,i}(\beta_k)$ are in $\Omega_\varepsilon(\beta_k)$ uniformly over $\beta_k \in \Theta_k$ and for any $i = 1, \ldots, n$ wpa1, the same property holds for $\tilde{\omega}_{1,i}(\beta_k)$. By the Cauchy-Schwarz inequality, Assumption SC2(v) and (SC.52)

$$\left| \upsilon_2'(\tilde{P}_{2,i}(\beta_k) - \hat{P}_{2,i}(\beta_k)) \right| \leq \|\upsilon_2\| \, \xi_{1,m_2} |\hat{\phi}(x_{1,i}) - \phi(x_{1,i})| \text{ wpa1.}$$

Therefore,

$$\begin{aligned} &\upsilon_2'(\hat{\mathbf{P}}_2(\beta_k) - \tilde{\mathbf{P}}_2(\beta_k))'(\hat{\mathbf{P}}_2(\beta_k) - \tilde{\mathbf{P}}_2(\beta_k))\upsilon_2 \\ &= \sum_{i=1}^{n} (\upsilon_2'(\tilde{P}_{2,i}(\beta_k) - \hat{P}_{2,i}(\beta_k)))^2 \leq \|\upsilon_2\|^2 \, \xi_{1,m_2}^2 \sum_{i=1}^{n} |\hat{\phi}(x_{1,i}) - \phi(x_{1,i})|^2 \end{aligned}$$

wpa1, which together with Lemma SC4 implies that

$$\sup_{\beta_k \in \Theta_k} ||\hat{\mathbf{P}}_2(\beta_k) - \tilde{\mathbf{P}}_2(\beta_k)||_S = O_p(\xi_{1,m_2} m_1^{1/2}). \tag{SC.53}$$

By Lemma SC27 and Assumption SC2(iv, vi), we have uniformly over $\beta_k \in \Theta_k$

$$C^{-1} \leq \lambda_{\min}(n^{-1}\tilde{\mathbf{P}}_2(\beta_k)'\tilde{\mathbf{P}}_2(\beta_k)) \leq \lambda_{\max}(n^{-1}\tilde{\mathbf{P}}_2(\beta_k)'\tilde{\mathbf{P}}_2(\beta_k)) \leq C \text{ wpa1.} \tag{SC.54}$$

By the triangle inequality, Assumption SC2(vi), (SC.53) and (SC.54), we get

$$\begin{aligned} &\sup_{\beta_k \in \Theta_k} \left\| n^{-1}\hat{\mathbf{P}}_2(\beta_k)'\hat{\mathbf{P}}_2(\beta_k) - n^{-1}\tilde{\mathbf{P}}_2(\beta_k)'\tilde{\mathbf{P}}_2(\beta_k) \right\|_S \\ \leq\; &\sup_{\beta_k \in \Theta_k} n^{-1} \left\| (\hat{\mathbf{P}}_2(\beta_k) - \tilde{\mathbf{P}}_2(\beta_k))'(\hat{\mathbf{P}}_2(\beta_k) - \tilde{\mathbf{P}}_2(\beta_k)) \right\|_S \\ &+ \sup_{\beta_k \in \Theta_k} n^{-1} \left\| (\hat{\mathbf{P}}_2(\beta_k) - \tilde{\mathbf{P}}_2(\beta_k))'\tilde{\mathbf{P}}_2(\beta_k) \right\|_S \\ &+ \sup_{\beta_k \in \Theta_k} n^{-1} \left\| \tilde{\mathbf{P}}_2(\beta_k)'(\hat{\mathbf{P}}_2(\beta_k) - \tilde{\mathbf{P}}_2(\beta_k)) \right\|_S = O_p(\xi_{1,m_2} m_1^{1/2} n^{-1/2}) \end{aligned}$$

which proves the claim of the lemma. *Q.E.D.*

Lemma SC6. *Under Assumptions SC1 and SC2, we have*

$$\sup_{\beta_k \in \Theta_k} n^{-1} \sum_{i=1}^{n} \left| \tilde{P}_{2,i}(\beta_k)'\hat{\beta}_g(\beta_k) - g(\omega_{1,i}(\beta_k); \beta_k) \right|^2 = O_p((m_2^2 + \xi_{1,m_2}^2 m_1) n^{-1}) = o_p(1)$$

where $\hat{\beta}_g(\beta_k) \equiv (\hat{\mathbf{P}}_2(\beta_k)'\hat{\mathbf{P}}_2(\beta_k))^{-1}\hat{\mathbf{P}}_2(\beta_k)'\hat{\mathbf{Y}}_2^*(\beta_k)$.

Proof of Lemma SC6. By the Cauchy-Schwarz inequality and Assumption SC2(iii)

$$
\begin{aligned}
& n^{-1}\sum_{i=1}^{n}\left|\tilde{P}_{2,i}(\beta_k)'\hat{\beta}_g(\beta_k)-g(\omega_{1,i}(\beta_k);\beta_k)\right|^2 \\
\leq\ & 2n^{-1}\sum_{i=1}^{n}\left|\tilde{P}_{2,i}(\beta_k)'\hat{\beta}_g(\beta_k)-g_{m_2}(\omega_{1,i}(\beta_k);\beta_k)\right|^2 \\
& +2n^{-1}\sum_{i=1}^{n}|g_{m_2}(\omega_{1,i}(\beta_k);\beta_k)-g(\omega_{1,i}(\beta_k);\beta_k)|^2 \\
\leq\ & 2\lambda_{\max}(n^{-1}\tilde{\mathbf{P}}_2(\beta_k)'\tilde{\mathbf{P}}_2(\beta_k))||\hat{\beta}_g(\beta_k)-\tilde{\beta}_{g,m_2}(\beta_k)||^2+Cm_2^{-2r_g} \qquad \text{(SC.55)}
\end{aligned}
$$

for any $\beta_k \in \Theta_k$, where $\tilde{\beta}_{g,m_2}(\beta_k) \equiv (B(\beta_k)')^{-1}\beta_{g,m_2}(\beta_k)$ and $\beta_{g,m_2}(\beta_k)$ is defined in Assumption SC2(iii). We next show that

$$
\sup_{\beta_k\in\Theta_k}\left\|\hat{\beta}_g(\beta_k)-\tilde{\beta}_{g,m_2}(\beta_k)\right\|^2 = O_p((m_2^2+\xi_{1,m_2}^2 m_1)n^{-1}) = o_p(1) \qquad \text{(SC.56)}
$$

which together with (SC.54) and (SC.55) proves the claim of the lemma.

Let $u_{2,i}(\beta_k) \equiv y_{2,i}^* - k_{2,i}\beta_k - g(\omega_{1,i}(\beta_k),\beta_k)$. Then we can write

$$
\begin{aligned}
\hat{\beta}_g(\beta_k)-\tilde{\beta}_{g,m_2}(\beta_k) &= (\hat{\mathbf{P}}_2(\beta_k)'\hat{\mathbf{P}}_2(\beta_k))^{-1}\hat{\mathbf{P}}_2(\beta_k)'(\hat{\mathbf{Y}}_2^*(\beta_k)-\hat{\mathbf{P}}_2(\beta_k)'\tilde{\beta}_{g,m_2}(\beta_k)) \\
&= (\hat{\mathbf{P}}_2(\beta_k)'\hat{\mathbf{P}}_2(\beta_k))^{-1}\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_k)(g(\omega_{1,i}(\beta_k),\beta_k)-g_{m_2}(\hat{\omega}_{1,i}(\beta_k),\beta_k)) \\
&\quad -(\hat{\beta}_l-\beta_{l,0})(\hat{\mathbf{P}}_2(\beta_k)'\hat{\mathbf{P}}_2(\beta_k))^{-1}\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_k)l_{2,i} \\
&\quad +(\hat{\mathbf{P}}_2(\beta_k)'\hat{\mathbf{P}}_2(\beta_k))^{-1}\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_k)u_{2,i}(\beta_k) \qquad \text{(SC.57)}
\end{aligned}
$$

where $g_{m_2}(\hat{\omega}_{1,i}(\beta_k),\beta_k) = \hat{P}_{2,i}(\beta_k)'\tilde{\beta}_{g,m_2}(\beta_k)$. By Assumption SC2(vi), Lemma SC5 and (SC.54), we have uniformly over $\beta_k \in \Theta_k$

$$
C^{-1} \leq \lambda_{\min}(n^{-1}\hat{\mathbf{P}}_2(\beta_k)'\hat{\mathbf{P}}_2(\beta_k)) \leq \lambda_{\max}(n^{-1}\hat{\mathbf{P}}_2(\beta_k)'\hat{\mathbf{P}}_2(\beta_k)) \leq C \text{ wpa1} \qquad \text{(SC.58)}
$$

which implies that $\hat{\mathbf{P}}_2(\beta_k)(\hat{\mathbf{P}}_2(\beta_k)'\hat{\mathbf{P}}_2(\beta_k))^{-1}\hat{\mathbf{P}}_2(\beta_k)'$ is an idempotent matrix uniformly over $\beta_k \in$

Θ_k wpa1. Therefore,

$$\begin{aligned}
&\left\| (\hat{\mathbf{P}}_2(\beta_k)'\hat{\mathbf{P}}_2(\beta_k))^{-1} \sum_{i=1}^{n} \hat{P}_{2,i}(\beta_k)(g(\omega_{1,i}(\beta_k),\beta_k) - g_{m_2}(\hat{\omega}_{1,i}(\beta_k),\beta_k)) \right\|^2 \\
\leq\ & O_p(1)n^{-1}\sum_{i=1}^{n}(g(\omega_{1,i}(\beta_k),\beta_k) - g_{m_2}(\hat{\omega}_{1,i}(\beta_k),\beta_k))^2. \qquad \text{(SC.59)}
\end{aligned}$$

uniformly over $\beta_k \in \Theta_k$. Since $\omega_{1,i}(\beta_k) = \phi(x_{1,i}) - k_{1,i}\beta_k$, we can use Assumptions SC1(i) and SC2(i) to deduce

$$\sup_{\beta_k \in \Theta_k} |g(\omega_{1,i}(\beta_k);\beta_k)| \leq C. \qquad \text{(SC.60)}$$

Therefore,

$$\begin{aligned}
\sup_{\beta_k \in \Theta_k} \left\| \tilde{\beta}_{g,m_2}(\beta_k) \right\|^2 &\leq \sup_{\beta_k \in \Theta_k} (\lambda_{\min}(Q_{m_2}(\beta_k)))^{-1} \left\| \tilde{P}_{2,i}(\beta_k)'\tilde{\beta}_{g,m_2}(\beta_k) \right\|_2^2 \\
&\leq C \sup_{\beta_k \in \Theta_k} \|g(\omega_{1,i}(\beta_k);\beta_k) - g_{m_2}(\omega_{1,i}(\beta_k);\beta_k)\|_2^2 \\
&\quad + C \sup_{\beta_k \in \Theta_k} \|g(\omega_{1,i}(\beta_k);\beta_k)\|_2^2 \leq C. \qquad \text{(SC.61)}
\end{aligned}$$

By the second order expansion, Assumption SC2(iii, v, vi), Lemma SC4, (SC.60) and (SC.61), we have uniformly over $\beta_k \in \Theta_k$,

$$\begin{aligned}
& n^{-1}\sum_{i=1}^{n}(g_{m_2}(\omega_{1,i}(\beta_k),\beta_k) - g_{m_2}(\hat{\omega}_{1,i}(\beta_k),\beta_k))^2 \\
\leq\ & 2n^{-1}\sum_{i=1}^{n}(\partial^1 \tilde{P}_{2,i}(\beta_k)'\tilde{\beta}_{g,m_2}(\beta_k)(\hat{\phi}(x_{1,i}) - \phi(x_{1,i}))^2 \\
& + 2n^{-1}\sum_{i=1}^{n}(\partial^2 \tilde{P}_2\left(\tilde{\omega}_{1,i}(\beta_k);\beta_k\right)'\tilde{\beta}_{g,m_2}(\beta_k)(\hat{\phi}(x_{1,i}) - \phi(x_{1,i})^2)^2 \\
=\ & O_p(m_1 n^{-1}) + O_p(\xi_{2,m_2}^2 \xi_{0,m_1}^2 m_1^2 n^{-2}) = O_p(m_1 n^{-1})
\end{aligned}$$

where $\tilde{\omega}_{1,i}(\beta_k)$ is between $\omega_{1,i}(\beta_k)$ and $\hat{\omega}_{1,i}(\beta_k)$ and it lies in $\Omega_\varepsilon(\beta_k)$ uniformly over $\beta_k \in \Theta_k$ wpa1 by (SC.51), which together with Assumption SC2(iii, vi) implies that

$$\begin{aligned}
& n^{-1}\sum_{i=1}^{n}(g(\omega_{1,i}(\beta_k),\beta_k)-g_{m_2}(\hat{\omega}_{1,i}(\beta_k),\beta_k))^2 \\
\leq\ & Cn^{-1}\sum_{i=1}^{n}(g(\omega_{1,i}(\beta_k),\beta_k)-g_{m_2}(\omega_{1,i}(\beta_k),\beta_k))^2 \\
& +Cn^{-1}\sum_{i=1}^{n}(g_{m_2}(\omega_{1,i}(\beta_k),\beta_k)-g_{m_2}(\hat{\omega}_{1,i}(\beta_k),\beta_k))^2 \\
=\ & O_p(m_1 n^{-1}+m_2^{-2r_g})=O_p(m_1 n^{-1}). \qquad \text{(SC.62)}
\end{aligned}$$

From (SC.59) and (SC.62), we get uniformly over $\beta_k\in\Theta_k$

$$(\hat{\mathbf{P}}_2(\beta_k)'\hat{\mathbf{P}}_2(\beta_k))^{-1}\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_k)(g(\omega_{1,i}(\beta_k),\beta_k)-g_{m_2}(\hat{\omega}_{1,i}(\beta_k),\beta_k))=O_p(m_1^{1/2}n^{-1/2}). \qquad \text{(SC.63)}$$

By Assumptions SC1(i) and SC2(i), and the Markov inequality,

$$n^{-1}\sum_{i=1}^{n}l_{2,i}^2=O_p(1) \qquad \text{(SC.64)}$$

which together with Assumption SC1(iii) and (SC.58) implies that

$$(\hat{\beta}_l-\beta_{l,0})(\hat{\mathbf{P}}_2(\beta_k)'\hat{\mathbf{P}}_2(\beta_k))^{-1}\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_k)l_{2,i}=O_p(n^{-1/2}) \qquad \text{(SC.65)}$$

uniformly over $\beta_k\in\Theta_k$. By the mean value expansion, the Cauchy-Schwarz inequality and the triangle inequality, we have for any $\upsilon_2\in\mathbb{R}^{m_2}$

$$\begin{aligned}
& \left|n^{-1}\sum_{i=1}^{n}\upsilon_2'(\hat{P}_{2,i}(\beta_k)-\tilde{P}_{2,i}(\beta_k))u_{2,i}(\beta_k)\right| \\
=\ & \left|n^{-1}\sum_{i=1}^{n}\upsilon_2'\partial^1\tilde{P}_2\left(\tilde{\omega}_{1,i}(\beta_k);\beta_k\right)\left(\hat{\omega}_{1,i}(\beta_k)-\omega_{1,i}(\beta_k)\right)u_{2,i}(\beta_k)\right| \\
\leq\ & \|\upsilon_2\|\,\xi_{1,m_2}n^{-1}\sum_{i=1}^{n}\left|(\hat{\phi}(x_{1,i})-\phi(x_{1,i}))u_{2,i}(\beta_k)\right|. \qquad \text{(SC.66)}
\end{aligned}$$

By the definition of $u_{2,i}(\beta_k)$, we can use Assumptions SC1(i) and SC2(i), (SC.60) and the Markov inequality to deduce

$$\sup_{\beta_k\in\Theta_k}n^{-1}\sum_{i=1}^{n}(u_{2,i}(\beta_k))^2=O_p(1). \qquad \text{(SC.67)}$$

Thus by the Cauchy-Schwarz inequality, Lemma SC4 and (SC.67),

$$\sup_{\beta_k\in\Theta_k} n^{-1}\sum_{i=1}^{n}\left|(\hat{\phi}(x_{1,i})-\phi(x_{1,i}))u_{2,i}(\beta_k)\right| = O_p(m_1^{1/2}n^{-1/2})$$

which together with (SC.58) and (SC.66) implies that

$$(\hat{\mathbf{P}}_2(\beta_k)'\hat{\mathbf{P}}_2(\beta_k))^{-1}\sum_{i=1}^{n}(\hat{P}_{2,i}(\beta_k)-\tilde{P}_{2,i}(\beta_k))u_{2,i}(\beta_k) = O_p(\xi_{1,m_2}m_1^{1/2}n^{-1/2}) \tag{SC.68}$$

uniformly over $\beta_k \in \Theta_k$. Applying Lemma SC28 and (SC.58) yields

$$(\hat{\mathbf{P}}_2(\beta_k)'\hat{\mathbf{P}}_2(\beta_k))^{-1}\sum_{i=1}^{n}\tilde{P}_{2,i}(\beta_k)u_{2,i}(\beta_k) = O_p(m_2 n^{-1/2}) \tag{SC.69}$$

uniformly over $\beta_k \in \Theta_k$. The claim in (SC.56) then follows from Assumption SC2(vi), (SC.57), (SC.63), (SC.65), (SC.68) and (SC.69). *Q.E.D.*

Lemma SC7. *Under Assumptions SC1 and SC2, we have*

$$\sup_{\beta_k\in\Theta_k} n^{-1}\sum_{i=1}^{n}\left|\hat{g}(\hat{\omega}_{1,i}(\beta_k);\beta_k)-g(\omega_{1,i}(\beta_k);\beta_k)\right|^2 = O_p((m_2^2+\xi_{1,m_2}^2 m_1)n^{-1}) = o_p(1).$$

Proof of Lemma SC7. By the triangle inequality, (SC.56) and (SC.61)

$$\sup_{\beta_k\in\Theta_k}\left\|\hat{\beta}_g(\beta_k)\right\| \leq \sup_{\beta_k\in\Theta_k}\left\|\tilde{\beta}_{g,m_2}(\beta_k)\right\| + \sup_{\beta_k\in\Theta_k}\left\|\hat{\beta}_g(\beta_k)-\tilde{\beta}_{g,m_2}(\beta_k)\right\| = O_p(1). \tag{SC.70}$$

By the mean value expansion, the Cauchy-Schwarz inequality, Assumption SC2(v, vi), Lemma SC4 and (SC.70),

$$\begin{aligned}
&\sup_{\beta_k\in\Theta_k} n^{-1}\sum_{i=1}^{n}\left|(\hat{P}_{2,i}(\beta_k)-\tilde{P}_{2,i}(\beta_k))'\hat{\beta}_g(\beta_k)\right|^2 \\
=\;& \sup_{\beta_k\in\Theta_k} n^{-1}\sum_{i=1}^{n}\left|\partial^1\tilde{P}_2\left(\tilde{\omega}_{1,i}(\beta_k);\beta_k\right)'\hat{\beta}_g(\beta_k)(\hat{\omega}_{1,i}(\beta_k)-\omega_{1,i}(\beta_k)\right|^2 \\
\leq\;& \xi_{1,m_2}^2 n^{-1}\sum_{i=1}^{n}(\hat{\phi}(x_{1,i})-\phi(x_{1,i}))^2\sup_{\beta_k\in\Theta_k}\left\|\hat{\beta}_g(\beta_k)\right\| = O_p(\xi_{1,m_2}^2 m_1 n^{-1}) = o_p(1) \quad \text{(SC.71)}
\end{aligned}$$

where $\tilde{\omega}_{1,i}(\beta_k)$ is between $\hat{\omega}_{1,i}(\beta_k)$ and $\omega_{1,i}(\beta_k)$ and hence by (SC.51) it lies in $\Omega_\varepsilon(\beta_k)$ wpa1 for any $i \leq n$ and uniformly over $\beta_k \in \Theta_k$. The claim of the lemma directly follows from Lemma SC6 and (SC.71). *Q.E.D.*

Lemma SC8. *Under Assumptions SC1 and SC2, we have*

$$\sup_{\beta_k \in \Theta_k} n^{-1} \sum_{i=1}^{n} (\tau_i(\beta_k)^2 - \mathbb{E}\left[\tau_i(\beta_k)^2\right]) = O_p(n^{-1/2}).$$

PROOF OF LEMMA SC8. For any $\beta_k \in \Theta_k$, by the Cauchy-Schwarz inequality and (SC.60),

$$\tau_i(\beta_k)^2 \leq C\left[(y_{2,i}^*)^2 + k_{2,i}^2\beta_k^2 + g(\omega_{1,i}(\beta_k);\beta_k)^2\right] \leq C(1 + (y_{2,i}^*)^2 + k_{2,i}^2). \tag{SC.72}$$

For any $\beta_{k,1}$, $\beta_{k,2} \in \Theta_k$, by the triangle inequality and Assumption SC2(ii),

$$|\tau_i(\beta_{k,1}) - \tau_i(\beta_{k,2})| \leq (C + k_{2,i})\left|\beta_{k,1} - \beta_{k,2}\right|. \tag{SC.73}$$

By Assumption SC2(ii), (SC.72) and (SC.73), we get

$$\mathbb{E}\left[\tau_i(\beta_k)^2 \left|\tau_i(\beta_{k,1}) - \tau_i(\beta_{k,2})\right|^2\right] \leq C(\beta_{k,1} - \beta_{k,2})^2 \tag{SC.74}$$

for any $\beta_k \in \Theta_k$, which implies that

$$\mathbb{E}\left[\left|\tau_i(\beta_{k,1})^2 - \tau_i(\beta_{k,2})^2\right|^2\right] \leq C(\beta_{k,1} - \beta_{k,2})^2.$$

Therefore we have for any $\beta_{k,1}, \beta_{k,2} \in \Theta_k$,

$$\left\|\tau_i(\beta_{k,1})^2 - \tau_i(\beta_{k,2})^2\right\|_2 \leq C\left|\beta_{k,1} - \beta_{k,2}\right|. \tag{SC.75}$$

By Assumptions SC1(i) and SC2(i), and (SC.60),

$$\begin{aligned}
&\mathbb{E}\left[\left|n^{-1/2}\sum_{i=1}^{n}\left(\tau_i(\beta_k)^2 - \mathbb{E}\left[\tau_i(\beta_k)^2\right]\right)\right|^2\right] \\
&= \mathbb{E}\left[\tau_i(\beta_k)^4\right] - \left(\mathbb{E}\left[\tau_i(\beta_k)^2\right]\right)^2 \leq C\left(\mathbb{E}\left[(y_{2,i}^*)^4 + k_{2,i}^4 + (g(\omega_{1,i}(\beta_k);\beta_k))^4\right]\right) \leq C
\end{aligned}$$

for any $\beta_k \in \Theta_k$, which implies that

$$n^{-1/2}\sum_{i=1}^{n}\left(\tau_i(\beta_k)^2 - \mathbb{E}\left[\tau_i(\beta_k)^2\right]\right) = O_p(1) \tag{SC.76}$$

for any $\beta_k \in \Theta_k$. Moreover, by Assumption SC1(i) and (SC.75)

$$\mathbb{E}\left[\left|n^{-1/2}\sum_{i=1}^{n}\left(\tau_i(\beta_{k,1})^2-\tau_i(\beta_{k,2})^2-\mathbb{E}\left[\tau_i(\beta_{k,1})^2-\tau_i(\beta_{k,1})^2\right]\right)\right|^2\right]$$
$$\leq \mathbb{E}\left[\left|\tau_i(\beta_{k,1})^2-\tau_i(\beta_{k,2})^2\right|^2\right]\leq C\left|\beta_{k,1}-\beta_{k,2}\right|^2. \tag{SC.77}$$

Collecting the results in (SC.76) and (SC.77), we can invoke Theorem 2.2.4 in van der Vaart and Wellner (1996) to deduce that

$$\left\|\sup_{\beta_k\in\Theta_k}\left|n^{-1/2}\sum_{i=1}^{n}\left(\tau_i(\beta_k)^2-\mathbb{E}\left[\tau_i(\beta_k)^2\right]\right)\right|\right\|_2\leq C$$

which together with the Markov inequality finishes the proof. *Q.E.D.*

SC.3 Auxiliary results for the asymptotic normality of $\hat{\beta}_k$

Let $a_j(\omega)\equiv\mathbb{E}[k_{j,i}|\omega_{1,i}=\omega]$ and $v_{j,i}\equiv k_{j,i}-a_j(\omega_{1,i})$ for $j=1,2$. Define

$$h_1(x_{1,i})\equiv\mathbb{E}[l_{1,i}|x_{1,i}] \quad \text{and} \quad \varphi(\omega)\equiv a_2(\omega)-a_1(\omega)g_1(\omega)$$

where $g_1(\omega)\equiv\partial g(\omega)/\partial\omega$. For any $\beta_k\in\Theta_k$ and $i=1,\ldots,n$, let

$$\hat{g}_i(\beta_k)\equiv\hat{g}(\hat{\omega}_{1,i}(\beta_k);\beta_k) \quad \text{and} \quad \hat{g}_{1,i}(\beta_k)\equiv\hat{g}_1(\hat{\omega}_{1,i}(\beta_k);\beta_k).$$

The following assumptions are needed for showing the asymptotic normality of $\hat{\beta}_k$.

Assumption SC3. *(i) $\varphi(\omega)$ is continuously differentiable with uniformly bounded derivatives over $\omega\in\Omega(\beta_{k,0})$; (ii) there exist $\beta_{\varphi,m_2}\in\mathbb{R}^{m_2}$ and $r_\varphi>0$ such that*

$$\sup_{\omega\in\Omega(\beta_{k,0})}|\varphi(\omega)-\varphi_{m_2}(\omega)|=O(m_2^{-r_\varphi})$$

where $\varphi_{m_2}(\omega)\equiv P_2(\omega)'\beta_{\varphi,m_2}$; (iii) for any function $\psi(\cdot)$ with $\|\psi(x_{1,i})\|_2<\infty$, there exists $\beta_{\psi,m_1}\in\mathbb{R}^{m_2}$ such that $\|\psi-\psi_{m_1}\|_2\to 0$ as $m_1\to\infty$ where $\psi_{m_1}(x_1)\equiv P(x_1)'\beta_{\psi,m_1}$; (iv) $n^{1/2}m_2^{-r_\varphi}+m_1m_2^4n^{-1/2}=o(1)$.

Assumptions SC3(i, ii) require that the function $\varphi(\omega)$ is smooth and can be well approximated by the approximating functions $P_2(\omega)$. Assumption SC3(iii) requires that any function of $x_{1,i}$ with finite L_2-norm can be approximated by the approximating functions $P_1(x_{1,i})$. Assumption SC3(iv) restricts the numbers of the approximating functions, and the smoothness of $\varphi(\omega)$.

Lemma SC9. *Under Assumptions SC1, SC2 and SC2(iv), we have*

$$\left\|\hat{\beta}_g(\beta_{k,0}) - \tilde{\beta}_{g,m_2}(\beta_{k,0})\right\| = O_p((m_1^{1/2} + m_2^{1/2})n^{-1/2})$$

where $\tilde{\beta}_{g,m_2}(\beta_{k,0}) \equiv (B(\beta_{k,0})')^{-1}\beta_{g,m_2}(\beta_{k,0})$ *and* $\beta_{g,m_2}(\beta_{k,0})$ *is defined in Assumption SC2(iii).*

Proof of Lemma SC9. By the definition of $\hat{\beta}_g(\beta_k)$, we can utilize the decomposition in (SC.57), and the results in (SC.63) and (SC.65) to get

$$\hat{\beta}_g(\beta_{k,0}) - \tilde{\beta}_{g,m_2}(\beta_{k,0}) = (\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_{k,0})u_{2,i} + O_p(m_1^{1/2}n^{-1/2}). \tag{SC.78}$$

By the second order expansion, we have for any $\upsilon_2 \in \mathbb{R}^{m_2}$

$$\begin{aligned} n^{-1}\sum_{i=1}^{n}\upsilon_2'(\hat{P}_{2,i}(\beta_{k,0}) - \tilde{P}_{2,i}(\beta_{k,0}))u_{2,i} &= n^{-1}\sum_{i=1}^{n}\upsilon_2'\partial^1\tilde{P}_{2,i}(\beta_{k,0})(\hat{\phi}_i - \phi_i)u_{2,i} \\ &\quad + n^{-1}\sum_{i=1}^{n}\upsilon_2'\partial^2\tilde{P}_2(\tilde{\omega}_{1,i};\beta_{k,0})(\hat{\phi}_i - \phi_i)^2u_{2,i} \end{aligned} \tag{SC.79}$$

By Assumption SC2(i) and (SC.60),

$$\mathbb{E}\left[u_{2,i}^2|x_{1,i}\right] \le C. \tag{SC.80}$$

By Assumptions SC1(i, v) and SC2(vi), (SC.80) and the Markov inequality

$$\left\|n^{-1}\sum_{i=1}^{n}|u_{2,i}|\,P_1(x_{1,i})P_1(x_{1,i})' - \mathbb{E}\left[|u_{2,i}|\,P_1(x_{1,i})P_1(x_{1,i})'\right]\right\|_S = o_p(1). \tag{SC.81}$$

Since $\lambda_{\max}(\mathbb{E}[|u_{2,i}|\,P_1(x_{1,i})P_1(x_{1,i})']) \le C$ by Assumption SC1(v) and (SC.80), from (SC.81) we deduce that

$$\lambda_{\max}\left(n^{-1}\sum_{i=1}^{n}|u_{2,i}|\,P_1(x_{1,i})P_1(x_{1,i})'\right) \le C \text{ wpa1}. \tag{SC.82}$$

By (SC.47) and (SC.82), we get

$$n^{-1}\sum_{i=1}^{n}\left|u_{2,i}(\hat{\phi}_i - \phi_{m_1,i})^2\right| = O_p(m_1n^{-1}). \tag{SC.83}$$

By Assumptions SC1(i, iv) and SC2(i), and the Markov inequality

$$n^{-1}\sum_{i=1}^{n}\left|u_{2,i}(\phi_{m_1,i} - \phi_i)^2\right| = O_p(m_1^{-2r_\phi})$$

which together with (SC.83) and Assumption SC2(vi) implies that

$$n^{-1}\sum_{i=1}^{n}|u_{2,i}|\,(\hat{\phi}_i-\phi_i)^2=O_p(m_1n^{-1}). \tag{SC.84}$$

By the Cauchy-Schwarz inequality and the triangle inequality, Assumption SC2(v) and (SC.84)

$$\left|n^{-1}\sum_{i=1}^{n}\upsilon_2'\partial^2\tilde{P}_2\left(\tilde{\omega}_{1,i};\beta_{k,0}\right)(\hat{\phi}_i-\phi_i)^2u_{2,i}\right|\leq\|\upsilon_2\|\,O_p(\xi_{2,m_2}m_1n^{-1}). \tag{SC.85}$$

By Assumptions SC1(i, v) and SC2(v), and (SC.80),

$$\mathbb{E}\left[\left\|n^{-1}\sum_{i=1}^{n}u_{2,i}\partial^1\tilde{P}_{2,i}\left(\beta_{k,0}\right)P_1(x_{1,i})'\right\|^2\right]\leq C\xi_{1,m_2}^2m_1n^{-1}$$

which together with the Cauchy-Schwarz inequality, the Markov inequality and (SC.47) implies that

$$\begin{aligned}&\left|n^{-1}\sum_{i=1}^{n}u_{2,i}\upsilon_2'\partial^1\tilde{P}_{2,i}\left(\beta_{k,0}\right)P_1(x_{1,i})'(\hat{\beta}_\phi-\beta_{\phi,m_1})\right|\\ \leq\;&\|\upsilon_2\|\left\|\hat{\beta}_\phi-\beta_{\phi,m_1}\right\|\left\|n^{-1}\sum_{i=1}^{n}u_{2,i}\partial^1\tilde{P}_{2,i}\left(\beta_{k,0}\right)P_1(x_{1,i})'\right\|=\|\upsilon_2\|\,O_p(\xi_{1,m_2}m_1n^{-1})\end{aligned} \tag{SC.86}$$

By Assumptions SC1(i) and SC2(iii, v, vi), and (SC.80),

$$\mathbb{E}\left[\left\|n^{-1}\sum_{i=1}^{n}\partial^1\tilde{P}_{2,i}\left(\beta_{k,0}\right)(\phi_{m_2,i}-\phi_i)u_{2,i}\right\|^2\right]\leq C\xi_{1,m_2}^2n^{-2}$$

which together with the Cauchy-Schwarz inequality and the Markov inequality implies that

$$\left|n^{-1}\sum_{i=1}^{n}\upsilon_2'\partial^1\tilde{P}_{2,i}\left(\beta_{k,0}\right)(\phi_{m_2,i}-\phi_i)u_{2,i}\right|\leq\|\upsilon_2\|\,O_p(\xi_{1,m_2}n^{-1}). \tag{SC.87}$$

Collecting the results in (SC.86) and (SC.87) obtains

$$\left|n^{-1}\sum_{i=1}^{n}\upsilon_2'\partial^1\tilde{P}_{2,i}\left(\beta_{k,0}\right)(\hat{\phi}_i-\phi_i)u_{2,i}\right|\leq\|\upsilon_2\|\,O_p(\xi_{1,m_2}m_1n^{-1}). \tag{SC.88}$$

Therefore, from Assumptions SC2(vi) and SC3(iv), (SC.58), (SC.79), (SC.85) and (SC.88) we can deduce

$$(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}\sum_{i=1}^{n}(\hat{P}_{2,i}(\beta_{k,0})-\tilde{P}_{2,i}(\beta_{k,0}))u_{2,i}=O_p(m_1^{1/2}n^{-1/2}). \tag{SC.89}$$

By Assumptions SC1(i) and SC2(v), and (SC.80),

$$n^{-1}\sum_{i=1}^{n}\tilde{P}_{2,i}(\beta_{k,0})u_{2,i}=O_p(m_2^{1/2}n^{-1/2})$$

which together with (SC.58) implies that

$$(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}\sum_{i=1}^{n}\tilde{P}_{2,i}(\beta_{k,0})u_{2,i}=O_p(m_2^{1/2}n^{-1/2}). \tag{SC.90}$$

The claim of the lemma follows from (SC.78), (SC.89) and (SC.90). *Q.E.D.*

Lemma SC10. *Under Assumptions SC1, SC2 and SC3, we have:*

$$n^{-1}\sum_{i=1}^{n}\hat{\tau}_i(\beta_{k,0})k_{1,i}(\hat{g}_{1,i}(\beta_{k,0})-g_1(\omega_{1,i}))=o_p(n^{-1/2}).$$

PROOF OF LEMMA SC10. By the definition of $\hat{\tau}_i(\beta_{k,0})$, we can write

$$\begin{aligned}
& n^{-1}\sum_{i=1}^{n}\hat{\tau}_i(\beta_{k,0})k_{1,i}(\hat{g}_{1,i}(\beta_{k,0})-g_1(\omega_{1,i})) \\
= \; & n^{-1}\sum_{i=1}^{n}k_{1,i}(g(\omega_{1,i})-\hat{g}_i(\beta_{k,0}))(\hat{g}_{1,i}(\beta_{k,0})-g_1(\omega_{1,i})) \\
& +n^{-1}\sum_{i=1}^{n}k_{1,i}(\hat{y}^*_{2,i}(\beta_{k,0})-g(\omega_{1,i}))(\hat{g}_{1,i}(\beta_{k,0})-g_1(\omega_{1,i})).
\end{aligned} \tag{SC.91}$$

We shall show that both terms in the right hand side of the above equation are $o_p(n^{-1/2})$. By the

Cauchy-Schwarz inequality, (SC.58), (SC.62) and Lemma SC9

$$
\begin{aligned}
& n^{-1}\sum_{i=1}^{n}(\hat{g}_i(\beta_{k,0})-g(\omega_{1,i}))^2 \\
\leq\ & Cn^{-1}\sum_{i=1}^{n}(\hat{P}_{2,i}(\beta_{k,0})'(\hat{\beta}_g(\beta_{k,0})-\tilde{\beta}_{g,m_2}(\beta_{k,0})))^2 \\
& +Cn^{-1}\sum_{i=1}^{n}(\hat{P}_{2,i}(\beta_{k,0})'\tilde{\beta}_{g,m_2}(\beta_{k,0})-g(\omega_{1,i})))^2 \\
\leq\ & 2\left\|\hat{\beta}_g(\beta_{k,0})-\tilde{\beta}_{g,m_2}(\beta_{k,0})\right\|^2\lambda_{\max}(n^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))+O_p(m_1n^{-1}) \\
=\ & O_p((m_1+m_2)n^{-1})
\end{aligned}
\tag{SC.92}
$$

Similarly, we can show that

$$
\begin{aligned}
& n^{-1}\sum_{i=1}^{n}(\hat{g}_{1,i}(\beta_{k,0})-g_1(\omega_{1,i}))^2 \\
\leq\ & Cn^{-1}\sum_{i=1}^{n}(\partial^1\hat{P}_{2,i}(\beta_{k,0})'(\hat{\beta}_g(\beta_{k,0})-\tilde{\beta}_{g,m_2}(\beta_{k,0})))^2 \\
& +Cn^{-1}\sum_{i=1}^{n}((\partial^1\hat{P}_{2,i}(\beta_{k,0})-\partial^1\tilde{P}_{2,i}(\beta_{k,0}))'\tilde{\beta}_{g,m_2}(\beta_{k,0}))^2 \\
& +Cn^{-1}\sum_{i=1}^{n}(\partial^1\tilde{P}_{2,i}(\beta_{k,0})'\tilde{\beta}_{g,m_2}(\beta_{k,0})-g_1(\omega_{1,i}))^2 \\
\leq\ & C\xi_{1,m_2}^2\left\|\hat{\beta}_g(\beta_{k,0})-\tilde{\beta}_{g,m_2}(\beta_{k,0})\right\|^2+O_p(\xi_{2,m_2}^2m_1n^{-1}) \\
=\ & O_p(\xi_{1,m_2}^2(m_1+m_2)n^{-1}+\xi_{2,m_2}^2m_1n^{-1}).
\end{aligned}
\tag{SC.93}
$$

Therefore, by the Cauchy-Schwarz inequality, Assumption SC3(iv), (SC.92) and (SC.93),

$$
n^{-1}\sum_{i=1}^{n}k_{1,i}(\hat{g}_i(\beta_{k,0})-g(\omega_{1,i}))(\hat{g}_{1,i}(\beta_{k,0})-g_1(\omega_{1,i}))=o_p(n^{-1/2}).
\tag{SC.94}
$$

Since $\hat{y}^*_{2,i}(\beta_{k,0}) - g(\omega_{1,i}) = u_{2,i} - l_{1,i}(\hat{\beta}_l - \beta_{l,0})$, we can write

$$\begin{aligned} & n^{-1}\sum_{i=1}^{n}(\hat{y}^*_{2,i}(\beta_{k,0}) - g(\omega_{1,i}))k_{1,i}(\hat{g}_{1,i}(\beta_{k,0}) - g_1(\omega_{1,i})) \\ &= n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}(\hat{g}_{1,i}(\beta_{k,0}) - g_1(\omega_{1,i})) \\ &\quad - (\hat{\beta}_l - \beta_{l,0})n^{-1}\sum_{i=1}^{n} l_{1,i}k_{1,i}(\hat{g}_{1,i}(\beta_{k,0}) - g_1(\omega_{1,i})). \end{aligned} \tag{SC.95}$$

Since $k_{1,i}$ has bounded support, by Assumptions SC1(i, ii, iii), SC2(vi) and SC3(iv), (SC.93) and the Markov inequality,

$$(\hat{\beta}_l - \beta_{l,0})n^{-1}\sum_{i=1}^{n} l_{1,i}k_{1,i}(\hat{g}_{1,i}(\beta_{k,0}) - g_1(\omega_{1,i})) = o_p(n^{-1/2}). \tag{SC.96}$$

Let

$$\partial^1\hat{P}_{2,i}(\beta_k) \equiv \partial^1\tilde{P}_2(\hat{\omega}_{1,i}(\beta_k);\beta_k) \text{ for any } \beta_k \in \Theta_k.$$

Then we can write

$$\begin{aligned} & n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}(\hat{g}_{1,i}(\beta_{k,0}) - g_1(\omega_{1,i})) \\ = \; & n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}(\partial^1\hat{P}_{2,i}(\beta_{k,0}) - \partial^1\tilde{P}_{2,i}(\beta_{k,0}))'\hat{\beta}_g(\beta_{k,0}) \\ & + n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}\partial^1\tilde{P}_{2,i}(\beta_{k,0})'(\hat{\beta}_g(\beta_{k,0}) - \tilde{\beta}_{g,m_2}(\beta_{k,0})) \\ & + n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}\left(\partial^1\tilde{P}_{2,i}(\beta_{k,0})'\tilde{\beta}_{g,m_2}(\beta_{k,0}) - g_1(\omega_{1,i})\right). \end{aligned} \tag{SC.97}$$

By Assumptions SC1(i) and SC2(iii), (SC.80) and the Markov inequality, we have

$$n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}\left(\partial^1\tilde{P}_{2,i}(\beta_{k,0})'\tilde{\beta}_{g,m_2}(\beta_{k,0}) - g_1(\omega_{1,i})\right) = o_p(n^{-1/2}). \tag{SC.98}$$

Similarly,

$$n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}\partial^1\tilde{P}_{2,i}(\beta_{k,0}) = O_p(\xi_{1,m_2}n^{-1/2})$$

which together with Assumptions SC2(vi) and SC3(iv), and Lemma SC9 implies that

$$n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}\partial^{1}\tilde{P}_{2,i}(\beta_{k,0})'(\hat{\beta}_{g}(\beta_{k,0})-\tilde{\beta}_{g,m_2}(\beta_{k,0}))=o_p(n^{-1/2}). \tag{SC.99}$$

By Assumption SC1(i), (SC.80) and the Markov inequality

$$n^{-1}\sum_{i=1}^{n}u_{2,i}^{2}k_{1,i}^{2}=O_p(1). \tag{SC.100}$$

Let $\hat{\phi}_i\equiv\hat{\phi}(x_{1,i})$ and $\phi_i\equiv\phi(x_{1,i})$. By the second order expansion,

$$\begin{aligned}&n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(\partial^{1}\hat{P}_{2,i}(\beta_{k,0})-\partial^{1}\tilde{P}_{2,i}(\beta_{k,0}))'\hat{\beta}_{g}(\beta_{k,0})\\ =\;&n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(\hat{\phi}_i-\phi_i)\partial^{2}\tilde{P}_{2,i}(\beta_{k,0})'\hat{\beta}_{g}(\beta_{k,0})\\ &+n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(\hat{\phi}_i-\phi_i)^{2}\partial^{3}\tilde{P}_{2}(\tilde{\omega}_{1,i}(\beta_{k,0});\beta_{k,0})'\hat{\beta}_{g}(\beta_{k,0})\end{aligned} \tag{SC.101}$$

where $\tilde{\omega}_{1,i}(\beta_{k,0})$ is between $\hat{\omega}_{1,i}(\beta_{k,0})$ and $\omega_{1,i}(\beta_{k,0})$. By the Cauchy-Schwarz inequality, Assumption SC2(v), Lemma SC4, (SC.70) and (SC.100)

$$n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(\hat{\phi}_i-\phi_i)^{2}\partial^{3}\tilde{P}_{2}(\tilde{\omega}_{1,i}(\beta_{k,0});\beta_{k,0})'\hat{\beta}_{g}(\beta_{k,0})=O_p(\xi_{3,m_2}m_1n^{-1})=o_p(n^{-1/2}) \tag{SC.102}$$

where the second equality is by Assumptions SC2(vi) and SC3(iv). By Assumptions SC1(i, v) and SC2(v), and (SC.80)

$$n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}P_1(x_{1,i})\partial^{2}\tilde{P}_{2,i}(\beta_{k,0})'=O_p(\xi_{2,m_2}m_1^{1/2}n^{-1/2})$$

which together with Lemma SC4 and (SC.70) implies that

$$n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(\hat{\phi}_i-\phi_i)\partial^{2}\tilde{P}_{2,i}(\beta_{k,0})'\hat{\beta}_{g}(\beta_{k,0})=O_p(\xi_{2,m_2}m_1n^{-1})=o_p(n^{-1/2}) \tag{SC.103}$$

where the second equality is by Assumptions SC2(vi) and SC3(iv). Similarly, we can show that

$$n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(\phi_{m_1}(x_{1,i})-\phi(x_{1,i}))\partial^{2}\tilde{P}_{2,i}(\beta_{k,0})'\hat{\beta}_{g}(\beta_{k,0})=o_p(n^{-1/2})$$

which together with (SC.103) implies that

$$n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(\hat{\phi}_i-\phi_i)\partial^2\tilde{P}_{2,i}(\beta_{k,0})'\hat{\beta}_g(\beta_{k,0})=o_p(n^{-1/2}). \tag{SC.104}$$

Collecting the results in (SC.101), (SC.102) and (SC.104) we get

$$n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(\partial^1\hat{P}_{2,i}(\beta_{k,0})-\partial^1\tilde{P}_{2,i}(\beta_{k,0}))'\hat{\beta}_g(\beta_{k,0})=o_p(n^{-1/2}). \tag{SC.105}$$

By (SC.95), (SC.96), (SC.97), (SC.98), (SC.99) and (SC.105),

$$n^{-1}\sum_{i=1}^{n}(\hat{y}^*_{2,i}(\beta_{k,0})-g(\omega_{1,i}))k_{1,i}(\hat{g}_{1,i}(\beta_{k,0})-g_1(\omega_{1,i}))=o_p(n^{-1/2}). \tag{SC.106}$$

The claim of the lemma follows from (SC.91), (SC.94) and (SC.106). *Q.E.D.*

Lemma SC11. *Under Assumptions SC1, SC2 and SC3, we have*

$$\begin{aligned}&n^{-1}\sum_{i=1}^{n}(\hat{P}_{2,i}(\beta_{k,0})'\tilde{\beta}_{g,m_2}(\beta_{k,0})-g(\omega_{1,i}))(k_{2,i}-k_{1,i}g_1(\omega_{1,i}))\\ =\;&n^{-1}\sum_{i=1}^{n}g_1(\omega_{1,i})(\hat{\phi}(x_{1,i})-\phi(x_{1,i}))(k_{2,i}-k_{1,i}g_1(\omega_{1,i}))+o_p(n^{-1/2}).\end{aligned}$$

Proof of Lemma SC11. First we write

$$\begin{aligned}&n^{-1}\sum_{i=1}^{n}(\hat{P}_{2,i}(\beta_{k,0})'\tilde{\beta}_{g,m_2}(\beta_{k,0})-g(\omega_{1,i}))(k_{2,i}-k_{1,i}g_1(\omega_{1,i}))\\ =\;&n^{-1}\sum_{i=1}^{n}(\hat{P}_{2,i}(\beta_{k,0})-\tilde{P}_{2,i}(\beta_{k,0}))'\tilde{\beta}_{g,m_2}(\beta_{k,0})(k_{2,i}-k_{1,i}g_1(\omega_{1,i}))\\ &+n^{-1}\sum_{i=1}^{n}(\tilde{P}_{2,i}(\beta_{k,0})'\tilde{\beta}_{g,m_2}(\beta_{k,0})-g(\omega_{1,i}))(k_{2,i}-k_{1,i}g_1(\omega_{1,i})).\end{aligned} \tag{SC.107}$$

By Assumptions SC1(i) and SC2(i, ii), and the Markov inequality

$$n^{-1}\sum_{i=1}^{n}(k_{2,i}-k_{1,i}g_1(\omega_{1,i}))^2=O_p(1). \tag{SC.108}$$

Therefore by Assumption SC2(iii, vi) and (SC.108), we have

$$n^{-1}\sum_{i=1}^{n}(\tilde{P}_{2,i}(\beta_{k,0})'\tilde{\beta}_{g,m_2}(\beta_{k,0})-g(\omega_{1,i}))(k_{2,i}-k_{1,i}g_1(\omega_{1,i}))=o_p(n^{-1/2}). \quad \text{(SC.109)}$$

Recall that $\hat{\phi}_i \equiv \hat{\phi}(x_{1,i})$ and $\phi_i \equiv \phi(x_{1,i})$. By the second order expansion,

$$\begin{aligned}
& n^{-1}\sum_{i=1}^{n}(\hat{P}_{2,i}(\beta_{k,0})-\tilde{P}_{2,i}(\beta_{k,0}))'\tilde{\beta}_{g,m_2}(\beta_{k,0})(k_{2,i}-k_{1,i}g_1(\omega_{1,i})) \\
= \;& n^{-1}\sum_{i=1}^{n}\partial^1\tilde{P}_{2,i}(\beta_{k,0})'\tilde{\beta}_{g,m_2}(\beta_{k,0})(\hat{\phi}_i-\phi_i)(k_{2,i}-k_{1,i}g_1(\omega_{1,i})) \\
& +n^{-1}\sum_{i=1}^{n}\partial^2\tilde{P}_2(\tilde{\omega}_{1,i}(\beta_{k,0});\beta_{k,0})'\tilde{\beta}_{g,m_2}(\beta_{k,0})(\hat{\phi}_i-\phi_i)^2(k_{2,i}-k_{1,i}g_1(\omega_{1,i})). \quad \text{(SC.110)}
\end{aligned}$$

By the Cauchy-Schwarz inequality and the triangle inequality, Assumption SC2(v), (SC.51) and (SC.61)

$$\begin{aligned}
& \left|n^{-1}\sum_{i=1}^{n}\partial^2\tilde{P}_2(\tilde{\omega}_{1,i}(\beta_{k,0});\beta_{k,0})'\tilde{\beta}_{g,m_2}(\beta_{k,0})(\hat{\phi}_i-\phi_i)^2(k_{2,i}-k_{1,i}g_1(\omega_{1,i}))\right| \\
\leq \;& O_p(\xi_{2,m_2})n^{-1}\sum_{i=1}^{n}\left|k_{2,i}-k_{1,i}g_1(\omega_{1,i})\right|(\hat{\phi}_i-\phi_i)^2. \quad \text{(SC.111)}
\end{aligned}$$

Since by Assumption SC2(i, ii) $\mathbb{E}[\left|k_{2,i}-k_{1,i}g_1(\omega_{1,i})\right|^2|x_{1,i}] \leq C$, we can use the similar arguments for showing (SC.84) to get

$$n^{-1}\sum_{i=1}^{n}\left|k_{2,i}-k_{1,i}g_1(\omega_{1,i})\right|(\hat{\phi}_i-\phi_i)^2=O_p(m_1 n^{-1})$$

which combined with Assumption SC3(iv) and (SC.111) implies that

$$n^{-1}\sum_{i=1}^{n}\partial^2\tilde{P}_2(\tilde{\omega}_{1,i}(\beta_{k,0});\beta_{k,0})'\tilde{\beta}_{g,m_2}(\beta_{k,0})(\hat{\phi}_i-\phi_i)^2(k_{2,i}-k_{1,i}g_1(\omega_{1,i}))=o_p(n^{-1/2}). \quad \text{(SC.112)}$$

By the Cauchy-Schwarz inequality, Assumption SC2(iii), Lemma SC4 and (SC.108)

$$\begin{aligned}
& n^{-1}\sum_{i=1}^{n}\partial^1\tilde{P}_{2,i}(\beta_{k,0})'\tilde{\beta}_{g,m_2}(\beta_{k,0})(\hat{\phi}_i-\phi_i)(k_{2,i}-k_{1,i}g_1(\omega_{1,i})) \\
= \;& n^{-1}\sum_{i=1}^{n}g_1(\omega_{1,i})(\hat{\phi}_i-\phi_i)(k_{2,i}-k_{1,i}g_1(\omega_{1,i}))+o_p(n^{-1/2})
\end{aligned}$$

which together with (SC.110) and (SC.112) shows that

$$
\begin{aligned}
& n^{-1}\sum_{i=1}^{n}(\hat{P}_{2,i}(\beta_{k,0})-\tilde{P}_{2,i}(\beta_{k,0}))'\tilde{\beta}_{g,m_2}(\beta_{k,0})(k_{2,i}-k_{1,i}g_1(\omega_{1,i})) \\
= \; & n^{-1}\sum_{i=1}^{n}g_1(\omega_{1,i})(\hat{\phi}_i-\phi_i)(k_{2,i}-k_{1,i}g_1(\omega_{1,i}))+o_p(n^{-1/2}). \qquad \text{(SC.113)}
\end{aligned}
$$

The claim of the lemma follows from (SC.107), (SC.109) and (SC.112). *Q.E.D.*

Lemma SC12. *Under Assumptions SC1, SC2 and SC3, we have*

$$
\begin{aligned}
& n^{-1}\sum_{i=1}^{n}(\hat{g}_i(\beta_{k,0})-g(\omega_{1,i}))(k_{2,i}-k_{1,i}g_1(\omega_{1,i})) \\
& = n^{-1}\sum_{i=1}^{n}u_{2,i}\varphi(\omega_{1,i})-\mathbb{E}[l_{2,i}\varphi(\omega_{1,i})](\hat{\beta}_l-\beta_{l,0}) \\
& \quad + n^{-1}\sum_{i=1}^{n}g_1(\omega_{1,i})(\hat{\phi}(x_{1,i})-\phi(x_{1,i}))(v_{2,i}-v_{1,i}g_1(\omega_{1,i}))+o_p(n^{-1/2})
\end{aligned}
$$

where $\varphi(\omega_{1,i})\equiv\mathbb{E}[k_{2,i}-k_{1,i}g_1(\omega_{1,i})|\omega_{1,i}]$ *and* $v_{j,i}\equiv k_{j,i}-\mathbb{E}[k_{j,i}|\omega_{1,i}]$ *for* $j=1,2$.

Proof of Lemma SC12. By the definition of $\hat{g}_i(\beta_{k,0})$, we can write

$$
\begin{aligned}
& n^{-1}\sum_{i=1}^{n}(\hat{g}_i(\beta_{k,0})-g(\omega_{1,i}))(k_{2,i}-k_{1,i}g_1(\omega_{1,i})) \\
= \; & (\hat{\beta}_g(\beta_{k,0})-\tilde{\beta}_{g,m_2}(\beta_{k,0}))'n^{-1}\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_{k,0})(k_{2,i}-k_{1,i}g_1(\omega_{1,i})) \\
& +n^{-1}\sum_{i=1}^{n}(\hat{P}_{2,i}(\beta_{k,0})'\tilde{\beta}_{g,m_2}(\beta_{k,0})-g(\omega_{1,i}))(k_{2,i}-k_{1,i}g_1(\omega_{1,i})). \qquad \text{(SC.114)}
\end{aligned}
$$

In view of Lemma SC11 and (SC.114), the claim of the lemma follows if

$$
\begin{aligned}
& (\hat{\beta}_g(\beta_{k,0})-\tilde{\beta}_{g,m_2}(\beta_{k,0}))'n^{-1}\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_{k,0})(k_{2,i}-k_{1,i}g_1(\omega_{1,i})) \\
= \; & n^{-1}\sum_{i=1}^{n}u_{2,i}\varphi(\omega_{1,i})-\mathbb{E}[l_{2,i}\varphi(\omega_{1,i})](\hat{\beta}_l-\beta_{l,0}) \\
& -n^{-1}\sum_{i=1}^{n}g_1(\omega_{1,i})(\hat{\phi}(x_{1,i})-\phi(x_{1,i}))\varphi(\omega_{1,i})+o_p(n^{-1/2}). \qquad \text{(SC.115)}
\end{aligned}
$$

We next prove (SC.115).

Let $\hat{\beta}_{\varphi}(\beta_{k,0}) \equiv (\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_{k,0})(k_{2,i}-k_{1,i}g_1(\omega_{1,i}))$. Then we can write

$$\begin{aligned}&(\hat{\beta}_g(\beta_{k,0})-\tilde{\beta}_{g,m_2}(\beta_{k,0}))'n^{-1}\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_{k,0})(k_{2,i}-k_{1,i}g_1(\omega_{1,i}))\\ =\;&(\hat{\beta}_g(\beta_{k,0})-\tilde{\beta}_{g,m_2}(\beta_{k,0}))'(n^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))\hat{\beta}_{\varphi}(\beta_{k,0}).\end{aligned} \tag{SC.116}$$

Under Assumptions SC1, SC2 and SC3, we can use the same arguments for proving Lemma SC9 to show that

$$\hat{\beta}_{\varphi}(\beta_{k,0})-\tilde{\beta}_{\varphi,m_2}(\beta_{k,0})=O_p((m_1^{1/2}+m_2^{1/2})n^{-1/2})=o_p(1) \tag{SC.117}$$

where $\tilde{\beta}_{\varphi,m_2}(\beta_{k,0}) \equiv (B(\beta_{k,0})')^{-1}\beta_{\varphi,m_2}$ and β_{φ,m_2} is defined in Assumption SC3(ii). By Assumptions SC2(v, vi) and SC3, Lemma SC4, Lemma SC9, (SC.58) and (SC.117)

$$\begin{aligned}&(\hat{\beta}_g(\beta_{k,0})-\tilde{\beta}_{g,m_2}(\beta_{k,0}))'(n^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))\hat{\beta}_{\varphi}(\beta_{k,0})\\ =\;&(\hat{\beta}_g(\beta_{k,0})-\tilde{\beta}_{g,m_2}(\beta_{k,0}))'(n^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))\tilde{\beta}_{\varphi,m_2}(\beta_{k,0})+o_p(n^{-1/2})\\ =\;&(\hat{\beta}_g(\beta_{k,0})-\tilde{\beta}_{g,m_2}(\beta_{k,0}))'n^{-1}\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_{k,0})\varphi(\hat{\omega}_{1,i}(\beta_{k,0}))+o_p(n^{-1/2})\\ =\;&(\hat{\beta}_g(\beta_{k,0})-\tilde{\beta}_{g,m_2}(\beta_{k,0}))'n^{-1}\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_{k,0})\varphi(\omega_{1,i})+o_p(n^{-1/2}).\end{aligned} \tag{SC.118}$$

Using the decomposition in (SC.57), we can write

$$\begin{aligned}&(\hat{\beta}_g(\beta_{k,0})-\tilde{\beta}_{g,m_2}(\beta_{k,0}))'n^{-1}\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_{k,0})\varphi(\omega_{1,i})\\ =\;&\frac{\boldsymbol{\varphi}_n'\hat{\mathbf{P}}_2(\beta_{k,0})(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}}{n}\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_k)(g(\omega_{1,i})-g_{m_2}(\hat{\omega}_{1,i}(\beta_{k,0}),\beta_{k,0}))\\ &-(\hat{\beta}_l-\beta_{l,0})\frac{\boldsymbol{\varphi}_n'\hat{\mathbf{P}}_2(\beta_{k,0})(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}}{n}\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_k)l_{2,i}\\ &+\frac{\boldsymbol{\varphi}_n'\hat{\mathbf{P}}_2(\beta_{k,0})(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}}{n}\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_{k,0})u_{2,i},\end{aligned} \tag{SC.119}$$

where $\boldsymbol{\varphi}_n \equiv (\varphi(\omega_{1,1}),\ldots,\varphi(\omega_{1,n}))'$. The rest of the proof is divided into 3 steps. The claim in (SC.115) follows from (SC.116), (SC.118), (SC.119), (SC.120), (SC.124) and (SC.130) below.

Step 1. In this step, we show that

$$\begin{aligned}
&\frac{\boldsymbol{\varphi}_n'\hat{\mathbf{P}}_2(\beta_{k,0})(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}}{n}\sum_{i=1}^n \hat{P}_{2,i}(\beta_k)(g(\omega_{1,i}) - g_{m_2}(\hat{\omega}_{1,i}(\beta_{k,0}),\beta_{k,0})) \\
= &-n^{-1}\sum_{i=1}^n g_1(\omega_{1,i})(\hat{\phi}(x_{1,i}) - \phi(x_{1,i}))\varphi(\omega_{1,i}) + o_p(n^{-1/2}). \qquad \text{(SC.120)}
\end{aligned}$$

Recall that $\hat{\phi}_i \equiv \hat{\phi}(x_{1,i})$ and $\phi_i \equiv \phi(x_{1,i})$. By the second order expansion, Assumptions SC2(iii, v, vi) and SC3(iv), Lemma SC4, (SC.51), (SC.58) and (SC.61),

$$\begin{aligned}
&\sum_{i=1}^n \hat{P}_{2,i}(\beta_{k,0})(g_{m_2}(\hat{\omega}_{1,i}(\beta_{k,0}),\beta_{k,0}) - g(\omega_{1,i})) \\
= &\sum_{i=1}^n \hat{P}_{2,i}(\beta_{k,0})(\hat{P}_{2,i}(\beta_{k,0}) - \tilde{P}_{2,i}(\beta_{k,0}))'\tilde{\beta}_{g,m_2}(\beta_{k,0}) + o_p(n^{-1/2}) \\
= &\sum_{i=1}^n \hat{P}_{2,i}(\beta_{k,0})(\hat{\phi}_i - \phi_i)\partial^1\tilde{P}_{2,i}(\beta_{k,0})'\tilde{\beta}_{g,m_2}(\beta_{k,0}) \\
&+\sum_{i=1}^n \hat{P}_{2,i}(\beta_{k,0})(\hat{\phi}_i - \phi_i)^2\partial^2\tilde{P}_2(\tilde{\omega}_{1,i}(\beta_{k,0});\beta_{k,0})'\tilde{\beta}_{g,m_2}(\beta_{k,0}) + o_p(n^{-1/2}) \\
= &\sum_{i=1}^n \hat{P}_{2,i}(\beta_{k,0})(\hat{\phi}_i - \phi_i)g_1(\omega_{1,i}) + o_p(n^{-1/2}). \qquad \text{(SC.121)}
\end{aligned}$$

By Assumptions SC1(i) and SC3(i), and (SC.58),

$$n^{-1}\sum_{i=1}^n \hat{P}_{2,i}(\beta_{k,0})\varphi(\omega_{1,i}) = O_p(1)$$

which together with (SC.58) and (SC.121) implies that

$$\begin{aligned}
&\frac{\boldsymbol{\varphi}_n'\hat{\mathbf{P}}_2(\beta_{k,0})(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}}{n}\sum_{i=1}^n \hat{P}_{2,i}(\beta_k)(g(\omega_{1,i}) - g_{m_2}(\hat{\omega}_{1,i}(\beta_{k,0}),\beta_{k,0})) \\
= &\frac{\boldsymbol{\varphi}_n'\hat{\mathbf{P}}_2(\beta_{k,0})(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}}{n}\sum_{i=1}^n \hat{P}_{2,i}(\beta_{k,0})g_1(\omega_{1,i})(\hat{\phi}_i - \phi_i) + o_p(n^{-1/2}) \qquad \text{(SC.122)}
\end{aligned}$$

By Assumptions SC2(ii) and SC3(ii, iv), Lemma SC4 and (SC.58)

$$\begin{aligned}
&n^{-1}\sum_{i=1}^n (\hat{\phi}_i - \phi_i)g_1(\omega_{1,i})\hat{P}_{2,i}(\beta_{k,0})'(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'(\boldsymbol{\varphi}_n - \boldsymbol{\varphi}_{m_2,n}) \\
= &O_p(m_1^{1/2}n^{-1/2})O_p(m_2^{-r_\varphi}) = o_p(n^{-1/2})
\end{aligned}$$

which further implies that

$$
\begin{aligned}
& n^{-1}\sum_{i=1}^{n}(\hat{\phi}_i-\phi_i)g_1(\omega_{1,i})\hat{P}_{2,i}(\beta_{k,0})'(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'\boldsymbol{\varphi}_n \\
= \; & n^{-1}\sum_{i=1}^{n}(\hat{\phi}_i-\phi_i)g_1(\omega_{1,i})\hat{P}_{2,i}(\beta_{k,0})'\tilde{\beta}_{\varphi,m_2}(\beta_{k,0})+o_p(n^{-1/2}) \\
= \; & n^{-1}\sum_{i=1}^{n}(\hat{\phi}_i-\phi_i)g_1(\omega_{1,i})\varphi(\hat{\omega}_{1,i}(\beta_{k,0}))+o_p(n^{-1/2}) \\
= \; & n^{-1}\sum_{i=1}^{n}(\hat{\phi}_i-\phi_i)g_1(\omega_{1,i})\varphi(\omega_{1,i})+o_p(n^{-1/2}) \qquad \text{(SC.123)}
\end{aligned}
$$

The claim in (SC.120) follows from (SC.122) and (SC.123).

Step 2. In this step, we show that

$$
\frac{\boldsymbol{\varphi}_n'\hat{\mathbf{P}}_2(\beta_{k,0})(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}}{n}\hat{\mathbf{P}}_2(\beta_{k,0})'\mathbf{L}_2(\hat{\beta}_l-\beta_{l,0})=\mathbb{E}[l_{2,i}\varphi(\omega_{1,i})](\hat{\beta}_l-\beta_{l,0})+o_p(n^{-1/2}). \qquad \text{(SC.124)}
$$

By Assumptions SC1(i) and SC2(i), and (SC.58),

$$
n^{-1}\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_{k,0})l_{2,i}=O_p(1) \qquad \text{(SC.125)}
$$

Using the similar arguments for showing (SC.62), we get

$$
n^{-1}\sum_{i=1}^{n}\left(\hat{P}_{2,i}(\beta_{k,0})'\tilde{\beta}_{\varphi,m_2}(\beta_{k,0})-\varphi(\omega_{1,i})\right)^2=O_p(m_1n^{-1})
$$

which together with (SC.58) and (SC.125) implies that

$$
n^{-1}(\boldsymbol{\varphi}_n-\hat{\boldsymbol{\varphi}}_{m_2,n})'\hat{\mathbf{P}}_2(\beta_{k,0})(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'\mathbf{L}_2=O_p(m_1^{1/2}n^{-1/2}) \qquad \text{(SC.126)}
$$

where $\hat{\boldsymbol{\varphi}}_{m_2,n}=(\varphi_{m_2}(\hat{\omega}_{1,1}(\beta_{k,0})),\ldots,\varphi_{m_2}(\hat{\omega}_{1,n}(\beta_{k,0})))'$. Therefore,

$$
n^{-1}\boldsymbol{\varphi}_n'\hat{\mathbf{P}}_2(\beta_{k,0})(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'\mathbf{L}_2=n^{-1}\sum_{i=1}^{n}l_{2,i}\hat{P}_{2,i}(\beta_{k,0})'\tilde{\beta}_{\varphi,m_2}(\beta_{k,0})+O_p(m_1^{1/2}n^{-1/2}). \qquad \text{(SC.127)}
$$

By the first order expansion,

$$\begin{aligned}
& n^{-1}\sum_{i=1}^{n} l_{2,i}\hat{P}_{2,i}(\beta_{k,0})'\tilde{\beta}_{\varphi,m_2}(\beta_{k,0}) \\
= & n^{-1}\sum_{i=1}^{n} l_{2,i}\varphi(\omega_{1,i}) + n^{-1}\sum_{i=1}^{n} l_{2,i}(\varphi_{m_2}(\omega_{1,i}) - \varphi(\omega_{1,i})) \\
& + n^{-1}\sum_{i=1}^{n} l_{2,i}(\hat{\phi}_i - \phi_i)\partial^1\tilde{P}_2(\tilde{\omega}_{1,i}(\beta_{k,0});\beta_{k,0})'\tilde{\beta}_{\varphi,m_2}(\beta_{k,0}) \\
= & n^{-1}\sum_{i=1}^{n} l_{2,i}\varphi(\omega_{1,i}) + O_p(\xi_{1,m_2}m_1^{1/2}n^{-1/2}) \qquad \text{(SC.128)}
\end{aligned}$$

where the second equality is by Assumtions SC1(i), SC2(i, v, vi) and SC3(ii, iv), Lemma SC4, (SC.51) and (SC.61). Collecting the results in (SC.127) and (SC.128), we deduce that

$$\begin{aligned}
& n^{-1}\boldsymbol{\varphi}_n'\hat{\mathbf{P}}_2(\beta_{k,0})(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'\mathbf{L}_2 \\
= & n^{-1}\sum_{i=1}^{n} l_{2,i}\varphi(\omega_{1,i}) + O_p(\xi_{1,m_2}m_1^{1/2}n^{-1/2}) = \mathbb{E}[l_{2,i}\varphi(\omega_{1,i})] + o_p(1) \qquad \text{(SC.129)}
\end{aligned}$$

where the second equality is by the Markov inequality, Assumptions SC1(i), SC2(i) and SC3(i, iv). The claim in (SC.124) follows from Assumption SC1(iii) and (SC.129).

Step 3. In this step, we show that

$$\frac{\boldsymbol{\varphi}_n'\hat{\mathbf{P}}_2(\beta_{k,0})(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}}{n}\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_{k,0})u_{2,i} = n^{-1}\sum_{i=1}^{n} u_{2,i}\varphi(\omega_{1,i}) + o_p(n^{-1/2}). \qquad \text{(SC.130)}$$

By the second order expansion, we have for any we have $\upsilon_2 \in \mathbb{R}^{m_2}$

$$\begin{aligned}
\sum_{i=1}^{n}\upsilon_2'\hat{P}_{2,i}(\beta_{k,0})u_{2,i} = & \sum_{i=1}^{n}\upsilon_2'\tilde{P}_{2,i}(\beta_{k,0})u_{2,i} \\
& + \sum_{i=1}^{n}\upsilon_2'\partial^1\tilde{P}_{2,i}(\beta_{k,0})(\hat{\phi}_i - \phi_i)u_{2,i} \\
& + \sum_{i=1}^{n}\upsilon_2'\partial^2\tilde{P}_2(\tilde{\omega}_{1,i}(\beta_{k,0});\beta_{k,0})(\hat{\phi}_i - \phi_i)^2u_{2,i} \qquad \text{(SC.131)}
\end{aligned}$$

By the Markov inequality, Assumptions SC1(i, iv, v), SC2(v, vi) and SC3(iv), Lemma SC4 and (SC.80), we can show that

$$n^{-1}\sum_{i=1}^{n}\upsilon_2'\partial^1\tilde{P}_{2,i}(\beta_{k,0})(\hat{\phi}_i - \phi_i)u_{2,i} = \|\upsilon_2\|\, o_p(n^{-1/2}). \qquad \text{(SC.132)}$$

By the Cauchy-Schwarz inequality, Assumptions SC2(v) and SC3(iv), and (SC.51)

$$n^{-1}\sum_{i=1}^{n}\upsilon_2'\partial^2\tilde{P}_2(\tilde{\omega}_{1,i}(\beta_{k,0});\beta_{k,0})(\hat{\phi}_i-\phi_i)^2u_{2,i}=\|\upsilon_2\|\,o_p(n^{-1/2}). \tag{SC.133}$$

By Assumptions SC1(i) and SC3(i), and (SC.58),

$$n^{-1}\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_{k,0})\varphi(\omega_{1,i})=O_p(1) \tag{SC.134}$$

Combining the results in (SC.58), (SC.131), (SC.132), (SC.133) and (SC.134), we get

$$\begin{aligned}&\frac{\boldsymbol{\varphi}_n'\hat{\mathbf{P}}_2(\beta_{k,0})(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}}{n}\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_{k,0})u_{2,i}\\ =\;&\frac{\boldsymbol{\varphi}_n'\hat{\mathbf{P}}_2(\beta_{k,0})(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}}{n}\sum_{i=1}^{n}\tilde{P}_{2,i}(\beta_{k,0})u_{2,i}+o_p(n^{-1/2}).\end{aligned} \tag{SC.135}$$

Since $n^{-1}\sum_{i=1}^{n}\tilde{P}_{2,i}(\beta_{k,0})u_{2,i}=O_p(m_2^{1/2}n^{-1/2})$ by the Markov inequality, Assumptions SC1(i) and SC2(iv), and (SC.80), we can use similar arguments for showing (SC.126) to get

$$\frac{(\boldsymbol{\varphi}_n-\hat{\boldsymbol{\varphi}}_{m_2,n})'\hat{\mathbf{P}}_2(\beta_{k,0})(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}}{n}\sum_{i=1}^{n}\tilde{P}_{2,i}(\beta_{k,0})u_{2,i}=O_p(m_1^{1/2}m_2^{1/2}n^{-1})=o_p(n^{-1/2})$$

where the second equality is by Assumption SC3(iv). Therefore

$$\begin{aligned}&\frac{\boldsymbol{\varphi}_n'\hat{\mathbf{P}}_2(\beta_{k,0})(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}}{n}\sum_{i=1}^{n}\tilde{P}_{2,i}(\beta_{k,0})u_{2,i}\\ =\;&n^{-1}\sum_{i=1}^{n}\tilde{P}_{2,i}(\beta_{k,0})'\tilde{\beta}_{\varphi,m_2}(\beta_{k,0})u_{2,i}+o_p(n^{-1/2})\\ =\;&n^{-1}\sum_{i=1}^{n}u_{2,i}\varphi(\omega_{1,i})+o_p(n^{-1/2})\end{aligned} \tag{SC.136}$$

where the second equality is by the Markov inequality, Assumptions SC1(i) and SC3(ii, iv), and (SC.80). The claim in (SC.130) follows from (SC.135) and (SC.136). *Q.E.D.*

Lemma SC13. *Under Assumptions SC1, SC2 and SC3, we have*

$$\begin{aligned}
& n^{-1}\sum_{i=1}^{n} g_1(\omega_{1,i})(\hat{\phi}(x_{1,i})-\phi(x_{1,i}))(v_{2,i}-v_{1,i}g_1(\omega_{1,i})) \\
= \quad & n^{-1}\sum_{i=1}^{n} \eta_{1,i}g_1(\omega_{1,i})(v_{2,i}^{*}-v_{1,i}g_1(\omega_{1,i})) \\
& -\mathbb{E}[h_1(x_{1,i})g_{1,i}(v_{2,i}-v_{1,i}g_1(\omega_{1,i}))](\hat{\beta}_l-\beta_{l,0})+o_p(n^{-1/2})
\end{aligned}$$

where $h_1(x_{1,i}) \equiv \mathbb{E}[l_{1,i}|x_{1,i}]$ *and* $v_{2,i}^{*} \equiv \mathbb{E}[k_{2,i}|x_{1,i}]-\mathbb{E}[k_{2,i}|\omega_{1,i}]$ *for* $j=1,2$.

Proof of Lemma SC13. Since $\hat{\phi}(x_{1,i})-\phi(x_{1,i})=(\hat{\beta}_{\phi}-\beta_{\phi,m_1})'P_1(x_{1,i})+\phi_{m_1}(x_{1,i})-\phi(x_{1,i})$, we can write

$$\begin{aligned}
& n^{-1}\sum_{i=1}^{n} g_{1,i}(\hat{\phi}(x_{1,i})-\phi(x_{1,i}))(v_{2,i}-v_{1,i}g_{1,i}) \\
& =(\hat{\beta}_{\phi}-\beta_{\phi,m_1})'n^{-1}\sum_{i=1}^{n} P_1(x_{1,i})g_{1,i}(v_{2,i}-v_{1,i}g_{1,i}) \\
& \quad +n^{-1}\sum_{i=1}^{n} g_{1,i}(\phi_{m_1}(x_{1,i})-\phi(x_{1,i}))(v_{2,i}-v_{1,i}g_{1,i})
\end{aligned} \tag{SC.137}$$

where $g_{1,i} \equiv g_1(\omega_{1,i})$. By Assumptions SC1(i, iv, vi) and SC2(ii), and the Markov inequality

$$n^{-1}\sum_{i=1}^{n} g_{1,i}(\phi_{m_1}(x_{1,i})-\phi(x_{1,i})))(v_{2,i}-v_{1,i}g_{1,i})=o_p(n^{-1/2}). \tag{SC.138}$$

By Assumptions SC1(i, v, vi) and SC2(i, ii), and the Markov inequality

$$n^{-1}\sum_{i=1}^{n} P_1(x_{1,i})g_{1,i}(v_{2,i}-v_{1,i}g_{1,i})-\mathbb{E}\left[P_1(x_{1,i})g_{1,i}(v_{2,i}-v_{1,i}g_{1,i})\right]=O_p(m_1^{1/2}n^{-1/2})$$

which together with the LIE, Assumption SC3(iv) and (SC.47) implies that

$$\begin{aligned}
& (\hat{\beta}_{\phi}-\beta_{\phi,m_1})'n^{-1}\sum_{i=1}^{n} P_1(x_{1,i})g_{1,i}(v_{2,i}-v_{1,i}g_{1,i}) \\
= \quad & n^{-1}\sum_{i=1}^{n} \eta_{1,i}P_1(x_{1,i})'Q_{m_1}^{-1}\mathbb{E}\left[P_1(x_{1,i})g_{1,i}(v_{2,i}^{*}-v_{1,i}g_{1,i})\right] \\
& -(\hat{\beta}_l-\beta_{l,0})\mathbb{E}[h_1(x_{1,i})P_1(x_{1,i})']Q_{m_1}^{-1}\mathbb{E}\left[P_1(x_{1,i})g_{1,i}(v_{2,i}-v_{1,i}g_{1,i})\right]+o_p(n^{-1/2})
\end{aligned} \tag{SC.139}$$

By Assumptions SC1(i, ii, v), SC2(i, ii) and SC3(iii)

$$\begin{aligned}&\mathbb{E}\left[\left|n^{-1}\sum_{i=1}^{n}\eta_{1,i}\left[P_1(x_{1,i})'Q_{m_1}^{-1}\mathbb{E}[P_1(x_{1,i})g_{1,i}(v_{2,i}^{*}-v_{1,i}g_{1,i})]-g_{1,i}(v_{2,i}^{*}-v_{1,i}g_{1,i})\right]\right|^2\right]\\&\leq Cn^{-1}\mathbb{E}\left[\left|P_1(x_{1,i})'Q_{1,m_1}^{-1}\mathbb{E}[P_1(x_{1,i})g_{1,i}(v_{2,i}^{*}-v_{1,i}g_{1,i})]-g_{1,i}(v_{2,i}^{*}-v_{1,i}g_{1,i})\right|^2\right]=o(n^{-1})\end{aligned}$$

which together with the Markov inequality implies that

$$n^{-1}\sum_{i=1}^{n}\eta_{1,i}P_1(x_{1,i})'Q_{m_1}^{-1}\mathbb{E}[P_{1,i}g_{1,i}(v_{2,i}^{*}-v_{1,i}g_{1,i})]=n^{-1}\sum_{i=1}^{n}\eta_{1,i}g_{1,i}(v_{2,i}^{*}-v_{1,i}g_{1,i})+o_p(n^{-1/2}). \tag{SC.140}$$

By Hölder's inequality, Assumptions SC1(ii, v), SC2(ii) and SC3(iii)

$$\begin{aligned}&\left|\mathbb{E}\left[l_{1,i}P_1(x_{1,i})'\right]Q_{m_1}^{-1}\mathbb{E}\left[P_1(x_{1,i})g_{1,i}(v_{2,i}^{*}-v_{1,i}g_{1,i})\right]-\mathbb{E}[l_{1,i}g_{1,i}(v_{2,i}^{*}-v_{1,i}g_{1,i})]\right|^2\\=&\left|\mathbb{E}\left[l_{1,i}\left(P_1(x_{1,i})'Q_{m_1}^{-1}\mathbb{E}\left[P_1(x_{1,i})g_{1,i}(v_{2,i}^{*}-v_{1,i}g_{1,i})\right]-g_{1,i}(v_{2,i}^{*}-v_{1,i}g_{1,i})\right)\right]\right|^2\\\leq&\ \mathbb{E}\left[l_{1,i}^2\right]\mathbb{E}\left[\left(P_1(x_{1,i})'Q_{m_1}^{-1}\mathbb{E}\left[P_1(x_{1,i})g_{1,i}(v_{2,i}^{*}-v_{1,i}g_{1,i})\right]-g_{1,i}(v_{2,i}^{*}-v_{1,i}g_{1,i})\right)^2\right]=o(1)\end{aligned}$$

which combined with Assumption SC1(iii) implies that

$$\begin{aligned}&(\hat{\beta}_l-\beta_{l,0})\mathbb{E}[l_{1,i}P_1(x_{1,i})']Q_{m_1}^{-1}\mathbb{E}\left[P_1(x_{1,i})g_{1,i}(v_{2,i}-v_{1,i}g_{1,i})\right]\\=&\ (\hat{\beta}_l-\beta_{l,0})\mathbb{E}[l_{1,i}P_1(x_{1,i})']Q_{m_1}^{-1}\mathbb{E}\left[P_1(x_{1,i})g_{1,i}(v_{2,i}^{*}-v_{1,i}g_{1,i})\right]\\=&\ (\hat{\beta}_l-\beta_{l,0})\mathbb{E}[l_{1,i}g_{1,i}(v_{2,i}^{*}-v_{1,i}g_{1,i})]+o_p(n^{-1/2})\\=&\ (\hat{\beta}_l-\beta_{l,0})\mathbb{E}[h_1(x_{1,i})g_{1,i}(v_{2,i}-v_{1,i}g_{1,i})]+o_p(n^{-1/2})\end{aligned} \tag{SC.141}$$

The claim of the lemma follows from (SC.137), (SC.138), (SC.139), (SC.140) and (SC.141).*Q.E.D.*

Lemma SC14. *Under Assumptions SC1, SC2 and SC3, we have*

$$\hat{\beta}_g(\hat{\beta}_k)-\tilde{\beta}_{g,m_2}(\hat{\beta}_k)=(\hat{\beta}_k-\beta_{k,0})O_p(m_2^3m_1^{1/2}n^{-1/2})+O_p((m_1^{1/2}+m_2)n^{-1/2}).$$

Proof of Lemma SC14. Using the decomposition in (SC.57), and applying the results in (SC.63), (SC.65) and (SC.69), we have

$$\hat{\beta}_g(\hat{\beta}_k)-\tilde{\beta}_{g,m_2}(\hat{\beta}_k)=(\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k))^{-1}\sum_{i=1}^{n}(\hat{P}_{2,i}(\hat{\beta}_k)-\tilde{P}_{2,i}(\hat{\beta}_k))u_{2,i}(\hat{\beta}_k)+O_p((m_1^{1/2}+m_2)n^{-1/2}). \tag{SC.142}$$

By the second-order expansion, we have for any $\upsilon_2 \in \mathbb{R}^{m_2}$

$$
\begin{aligned}
& n^{-1}\sum_{i=1}^{n}\upsilon_2'(\hat{P}_{2,i}(\hat{\beta}_k) - \tilde{P}_{2,i}(\hat{\beta}_k))u_{2,i}(\hat{\beta}_k) \\
= \; & n^{-1}\sum_{i=1}^{n}\upsilon_2'\partial^1\tilde{P}_{2,i}(\hat{\beta}_k)(\hat{\phi}_i - \phi_i)u_{2,i}(\hat{\beta}_k) \\
& + n^{-1}\sum_{i=1}^{n}\upsilon_2'\partial^2\hat{P}_2(\tilde{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k)(\hat{\phi}_i - \phi_i)^2 u_{2,i}(\hat{\beta}_k) \qquad \text{(SC.143)}
\end{aligned}
$$

where $\tilde{\omega}_{1,i}(\hat{\beta}_k)$ lies between $\hat{\omega}_{1,i}(\hat{\beta}_k)$ and $\omega_{1,i}(\hat{\beta}_k)$. By (SC.60) and the compactness of Θ_k,

$$
\sup_{\beta_k \in \Theta_k} |u_{2,i}(\beta_k)| \leq C + |y_{2,i}^*| + |k_{2,i}|. \qquad \text{(SC.144)}
$$

Using similar arguments in showing (SC.84), we have

$$
n^{-1}\sum_{i=1}^{n}(\hat{\phi}_i - \phi_i)^2(C + |y_{2,i}^*| + |k_{2,i}|) = O_p(m_1 n^{-1})
$$

which together with the Cauchy-Schwarz inequality, the triangle inequality, Assumptions SC2(vi) and SC3(iv), and (SC.144) implies that

$$
\begin{aligned}
& \left| n^{-1}\sum_{i=1}^{n}\upsilon_2'\partial^2\hat{P}_{2,i}(\tilde{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k)(\hat{\phi}_i - \phi_i)^2 u_{2,i}(\hat{\beta}_k) \right| \\
\leq \; & \|\upsilon_2\| \, \xi_{2,m_2} n^{-1}\sum_{i=1}^{n}(\hat{\phi}_i - \phi_i)^2(C + |y_{2,i}^*|) = \|\upsilon_2\| \, o_p(m_1^{1/2} n^{-1/2}). \qquad \text{(SC.145)}
\end{aligned}
$$

Since $u_{2,i}(\hat{\beta}_k) = u_{2,i} - k_{2,i}(\hat{\beta}_k - \beta_{k,0}) - (g(\omega_{1,i}(\hat{\beta}_k), \hat{\beta}_k) - g(\omega_{1,i}))$, we can write

$$
\begin{aligned}
& n^{-1}\sum_{i=1}^{n}\upsilon_2'\partial^1\tilde{P}_{2,i}(\hat{\beta}_k)(\hat{\phi}_i - \phi_i)u_{2,i}(\hat{\beta}_k) \\
= \; & n^{-1}\sum_{i=1}^{n}\upsilon_2'\partial^1\tilde{P}_{2,i}(\hat{\beta}_k)(\hat{\phi}_i - \phi_i)u_{2,i} \\
& -(\hat{\beta}_k - \beta_{k,0})n^{-1}\sum_{i=1}^{n}\upsilon_2'\partial^1\tilde{P}_{2,i}(\hat{\beta}_k)k_{2,i}(\hat{\phi}_i - \phi_i) \\
& -n^{-1}\sum_{i=1}^{n}\upsilon_2'\partial^1\tilde{P}_{2,i}(\hat{\beta}_k)(\hat{\phi}_i - \phi_i)(g(\omega_{1,i}(\hat{\beta}_k), \hat{\beta}_k) - g(\omega_{1,i})). \qquad \text{(SC.146)}
\end{aligned}
$$

By the Cauchy-Schwarz inequality, the triangle inequality, Assumption SC2(v) and Lemma SC4

$$\left| n^{-1}\sum_{i=1}^{n} \upsilon_2' \partial^1 \tilde{P}_{2,i}(\hat{\beta}_k) k_{2,i}(\hat{\phi}_i - \phi_i) \right| \leq \|\upsilon_2\| \, O_p(\xi_{1,m_2} m_1^{1/2} n^{-1/2}). \tag{SC.147}$$

Similarly we can show that

$$\begin{aligned} & \left| n^{-1}\sum_{i=1}^{n} \upsilon_2' \partial^1 \tilde{P}_{2,i}(\hat{\beta}_k)(\hat{\phi}_i - \phi_i)(g(\omega_{1,i}(\hat{\beta}_k), \hat{\beta}_k) - g(\omega_{1,i})) \right| \\ \leq \quad & \|\upsilon_2\| \, |\hat{\beta}_k - \beta_{k,0}| O_p(\xi_{1,m_2} m_1^{1/2} n^{-1/2}). \end{aligned} \tag{SC.148}$$

By the Cauchy-Schwarz inequality, the triangle inequality, Assumption SC3(iv), Lemma SC30 and (SC.47),

$$\left| n^{-1}\sum_{i=1}^{n} \upsilon_2' \partial^1 \tilde{P}_{2,i}(\hat{\beta}_k)(\hat{\phi}_i - \phi_{m_1,i}) u_{2,i} \right| \leq \|\upsilon_2\| \, O_p((m_1^{1/2} + m_2) n^{-1/2}). \tag{SC.149}$$

Using similar arguments in the proof of Lemma SC30, we can show that

$$n^{-1}\sum_{i=1}^{n} \partial^1 \tilde{P}_{2,i}(\hat{\beta}_k)(\phi_{m_1,i} - \phi_i) u_{2,i} = O_p(m_2^{5/2} n^{-1})$$

which together with the Cauchy-Schwarz inequality and Assumption SC3(iv) implies that

$$\left| n^{-1}\sum_{i=1}^{n} \upsilon_2' \partial^1 \tilde{P}_{2,i}(\hat{\beta}_k)(\phi_{m_1,i} - \phi_i) u_{2,i} \right| \leq \|\upsilon_2\| \, O_p((m_1^{1/2} + m_2) n^{-1/2}). \tag{SC.150}$$

Collecting the results in (SC.143), (SC.145), (SC.146), (SC.147), (SC.148), (SC.149) and (SC.150), we have

$$n^{-1}\sum_{i=1}^{n} (\hat{P}_{2,i}(\hat{\beta}_k) - \tilde{P}_{2,i}(\hat{\beta}_k)) u_{2,i}(\hat{\beta}_k) = (\hat{\beta}_k - \beta_{k,0}) O_p(\xi_{1,m_2} m_1^{1/2} n^{-1/2}) + O_p((m_1^{1/2} + m_2) n^{-1/2}). \tag{SC.151}$$

The claim of the lemma follows from (SC.142) and (SC.151). *Q.E.D.*

Lemma SC15. *Under Assumptions SC1, SC2 and SC3, we have*

$$n^{-1}\sum_{i=1}^{n} \left| \hat{g}_i(\hat{\beta}_k) - g(\omega_{1,i}(\hat{\beta}_k); \hat{\beta}_k) \right|^2 = (\hat{\beta}_k - \beta_{k,0})^2 O_p(\xi_{1,m_2}^2 m_1 n^{-1}) + O_p\left((m_1 + m_2^2) n^{-1}\right).$$

Proof of Lemma SC15. For any $\beta_k \in \Theta_k$ we deduce by the Cauchy-Schwarz inequality, Assumption SC2(iii) and (SC.54) that

$$\begin{aligned} & n^{-1}\sum_{i=1}^{n}\left|\tilde{P}_{2,i}(\hat{\beta}_k)'\hat{\beta}_g(\hat{\beta}_k)-g(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k)\right|^2 \\ & \leq 2n^{-1}\sum_{i=1}^{n}\left|\tilde{P}_{2,i}(\hat{\beta}_k)'\hat{\beta}_g(\hat{\beta}_k)-g_{m_2}(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k)\right|^2 \\ & \qquad +2n^{-1}\sum_{i=1}^{n}\left|g_{m_2}(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k)-g(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k)\right|^2 \\ & \leq C||\hat{\beta}_g(\hat{\beta}_k)-\tilde{\beta}_{g,m_2}(\hat{\beta}_k)||^2+Cm_2^{-2r_g} \end{aligned} \tag{SC.152}$$

wpa1, which together with Assumption SC2(vi) and Lemma SC14 implies that

$$\begin{aligned} & n^{-1}\sum_{i=1}^{n}\left|\tilde{P}_{2,i}(\hat{\beta}_k)'\hat{\beta}_g(\hat{\beta}_k)-g(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k)\right|^2 \\ = \ & (\hat{\beta}_k-\beta_{k,0})^2O_p(\xi_{1,m_2}^2m_1n^{-1})+O_p\left((m_1+m_2^2)n^{-1}\right). \end{aligned} \tag{SC.153}$$

By the mean value expansion, the Cauchy-Schwarz inequality, Assumptions SC2(v) and SC3(iv), Lemma SC4, Lemma SC14 and (SC.56)

$$\begin{aligned} & n^{-1}\sum_{i=1}^{n}\left|(\hat{P}_{2,i}(\hat{\beta}_k)-\tilde{P}_{2,i}(\hat{\beta}_k))'(\hat{\beta}_g(\hat{\beta}_k)-\tilde{\beta}_{g,m_2}(\hat{\beta}_k))\right|^2 \\ = \ & n^{-1}\sum_{i=1}^{n}(\hat{\phi}_i-\phi_i)^2\left(\partial^1\tilde{P}_2(\tilde{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k)'(\hat{\beta}_g(\hat{\beta}_k)-\tilde{\beta}_{g,m_2}(\hat{\beta}_k))\right)^2 \\ \leq \ & \xi_{1,m_2}^2||\hat{\beta}_g(\hat{\beta}_k)-\tilde{\beta}_{g,m_2}(\hat{\beta}_k)||^2n^{-1}\sum_{i=1}^{n}(\hat{\phi}_i-\phi_i)^2 \\ = \ & (\hat{\beta}_k-\beta_{k,0})^2O_p(\xi_{1,m_2}^2m_1n^{-1})+O_p\left((m_1+m_2^2)n^{-1}\right). \end{aligned} \tag{SC.154}$$

By Assumptions SC2(ii, iii, vi), and Lemma SC4, we have

$$\begin{aligned} & n^{-1}\sum_{i=1}^{n}\left|(\hat{P}_{2,i}(\hat{\beta}_k)-\tilde{P}_{2,i}(\hat{\beta}_k))'\tilde{\beta}_{g,m_2}(\hat{\beta}_k)\right|^2 \\ \leq \ & 2n^{-1}\sum_{i=1}^{n}\left|(g_1(\hat{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k)-g_1(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k))\right|^2+O_p(n^{-1}) \\ \leq \ & Cn^{-1}\sum_{i=1}^{n}(\hat{\phi}_i-\phi_i)^2+O_p(n^{-1})=O_p(m_1n^{-1}) \end{aligned} \tag{SC.155}$$

which together with (SC.154) implies that

$$\begin{aligned}
& n^{-1}\sum_{i=1}^{n}\left|(\hat{P}_{2,i}(\hat{\beta}_k)-\tilde{P}_{2,i}(\hat{\beta}_k))'\hat{\beta}_g(\hat{\beta}_k)\right|^2 \\
= \;& (\hat{\beta}_k-\beta_{k,0})^2 O_p(\xi_{1,m_2}^2 m_1 n^{-1}) + O_p\left((m_1+m_2^2)n^{-1}\right). \quad \text{(SC.156)}
\end{aligned}$$

The claim of the lemma follows from (SC.153) and (SC.156). *Q.E.D.*

Lemma SC16. *Under Assumptions SC1, SC2 and SC3, we have*

$$\begin{aligned}
& n^{-1}\sum_{i=1}^{n}\left|\hat{g}_{1,i}(\hat{\beta}_k)-g_1(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k)\right|^2 \\
= \;& (\hat{\beta}_k-\beta_{k,0})^2 O_p(\xi_{1,m_2}^4 m_1 n^{-1}) + O_p\left(\xi_{1,m_2}^2(m_1+m_2^2)n^{-1}\right).
\end{aligned}$$

Proof of Lemma SC16. Since $\hat{g}_{1,i}(\hat{\beta}_k)=\partial^1\hat{P}_{2,i}(\hat{\beta}_k)'\hat{\beta}_g(\hat{\beta}_k)$, we can use similar arguments in showing (SC.153) to get

$$\begin{aligned}
& n^{-1}\sum_{i=1}^{n}\left|\partial^1\tilde{P}_{2,i}(\hat{\beta}_k)'\hat{\beta}_g(\hat{\beta}_k)-g_1(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k)\right|^2 \\
= \;& (\hat{\beta}_k-\beta_{k,0})^2 O_p(\xi_{1,m_2}^4 m_1 n^{-1}) + O_p\left(\xi_{1,m_2}^2(m_1+m_2^2)n^{-1}\right). \quad \text{(SC.157)}
\end{aligned}$$

Using similar arguments in showing (SC.154) and (SC.155), we can show that

$$\begin{aligned}
& n^{-1}\sum_{i=1}^{n}\left|(\partial^1\hat{P}_{2,i}(\hat{\beta}_k)-\partial^1\tilde{P}_{2,i}(\hat{\beta}_k))'(\hat{\beta}_g(\hat{\beta}_k)-\tilde{\beta}_{g,m_2}(\hat{\beta}_k))\right|^2 \\
\leq \;& (\hat{\beta}_k-\beta_{k,0})^2 O_p(\xi_{2,m_2}^2\xi_{1,m_2}^2 m_1^2 n^{-2}) + O_p\left(\xi_{2,m_2}^2 m_1(m_1+m_2^2)n^{-2}\right) \quad \text{(SC.158)}
\end{aligned}$$

and

$$n^{-1}\sum_{i=1}^{n}\left|(\partial^1\hat{P}_{2,i}(\hat{\beta}_k)-\partial^1\tilde{P}_{2,i}(\hat{\beta}_k))'\tilde{\beta}_{g,m_2}(\hat{\beta}_k)\right|^2 = O_p(m_1 n^{-1}), \quad \text{(SC.159)}$$

which implies that

$$\begin{aligned}
& n^{-1}\sum_{i=1}^{n}\left|(\partial^1\hat{P}_{2,i}(\hat{\beta}_k)-\partial^1\tilde{P}_{2,i}(\hat{\beta}_k))'\hat{\beta}_g(\hat{\beta}_k)\right|^2 \\
\leq \;& (\hat{\beta}_k-\beta_{k,0})^2 O_p(\xi_{2,m_2}^2\xi_{1,m_2}^2 m_1^2 n^{-2}) + O_p\left(\xi_{2,m_2}^2(m_1+m_2^2)m_1 n^{-2}\right). \quad \text{(SC.160)}
\end{aligned}$$

The claim of the lemma follows from Assumption SC3(iv), (SC.157) and (SC.160). *Q.E.D.*

Lemma SC17. *Under Assumptions SC1, SC2 and SC3, we have*

$$n^{-1}\sum_{i=1}^{n} k_{2,i}(k_{2,i} - k_{1,i}\hat{g}_{1,i}(\hat{\beta}_k)) = \mathbb{E}[k_{2,i}(k_{2,i} - k_{1,i}g_1(\omega_{1,i}))] + o_p(1) \tag{SC.161}$$

and

$$n^{-1}\sum_{i=1}^{n} l_{2,i}k_{1,i}(\hat{g}_{1,i}(\hat{\beta}_k) - \hat{g}_{1,i}(\beta_{k,0}) = o_p(1). \tag{SC.162}$$

Proof of Lemma SC17. By the Cauchy-Schwarz inequality, Assumptions SC2(ii, vi) and SC3(iv), Lemma SC16 and the consistency of $\hat{\beta}_k$, we have

$$\begin{aligned} n^{-1}\sum_{i=1}^{n} k_{2,i}(k_{2,i} - k_{1,i}\hat{g}_{1,i}(\hat{\beta}_k)) &= n^{-1}\sum_{i=1}^{n} k_{2,i}(k_{2,i} - k_{1,i}g_{1,i}(\hat{\beta}_k)) + o_p(1) \\ &= n^{-1}\sum_{i=1}^{n} k_{2,i}(k_{2,i} - k_{1,i}g_1(\omega_{1,i})) + o_p(1) \\ &= \mathbb{E}[k_{2,i}(k_{2,i} - k_{1,i}g_1(\omega_{1,i}))] + o_p(1) \end{aligned}$$

where the third equality is by the Markov inequality. This proves the claim in (SC.161). Similarly, by Assumptions SC2(ii, vi) and SC3(iv), Lemma SC16 and the consistency of $\hat{\beta}_k$, we have

$$\begin{aligned} & n^{-1}\sum_{i=1}^{n} \left|\hat{g}_{1,i}(\hat{\beta}_k) - \hat{g}_{1,i}(\beta_{k,0})\right|^2 \\ & \leq 2n^{-1}\sum_{i=1}^{n} \left|g_1(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k) - g_1(\omega_{1,i}(\beta_{k,0});\beta_{k,0})\right|^2 + o_p(1) \\ & \leq C(\hat{\beta}_k - \beta_{k,0})^2 + o_p(1) = o_p(1). \end{aligned} \tag{SC.163}$$

By the Markov inequality and Assumption SC2(i), $n^{-1}\sum_{i=1}^{n} l_{2,i}^2 k_{1,i}^2 = O_p(1)$ which together with (SC.163) proves the claim in (SC.162). *Q.E.D.*

Lemma SC18. *Let $a_{2,i} = a_2(\omega_{1,i})$. Then under Assumptions SC1, SC2 and SC3, we have*

$$\begin{aligned} & n^{-1}\sum_{i=1}^{n}(\hat{g}_i(\hat{\beta}_k) - \hat{g}_i(\beta_{k,0}))(k_{2,i} - k_{1,i}\hat{g}_{1,i}(\beta_{k,0})) \\ = & -(\hat{\beta}_k - \beta_{k,0})\left(\mathbb{E}[(a_{2,i} + v_{1,i}g_{1,i})(k_{2,i} - k_{1,i}g_{1,i})] + o_p(1)\right) + O_p((m_2 + m_1^{1/2})n^{-1/2}). \end{aligned}$$

Proof of Lemma SC18. First note that

$$
\begin{aligned}
& n^{-1}\sum_{i=1}^{n}(\hat{g}_i(\hat{\beta}_k)-\hat{g}_i(\beta_{k,0}))(k_{2,i}-k_{1,i}\hat{g}_{1,i}(\beta_{k,0})) \\
= \quad & -n^{-1}\sum_{i=1}^{n}k_{1,i}(\hat{g}_i(\hat{\beta}_k)-\hat{g}_i(\beta_{k,0}))(\hat{g}_{1,i}(\beta_{k,0})-g_{1,i}) \\
& +n^{-1}\sum_{i=1}^{n}(\hat{g}_i(\hat{\beta}_k)-g(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k)-\hat{g}_i(\beta_{k,0})+g(\omega_{1,i}))(k_{2,i}-k_{1,i}g_{1,i}) \\
& +n^{-1}\sum_{i=1}^{n}(g(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k)-g(\omega_{1,i}))(k_{2,i}-k_{1,i}g_{1,i}). \qquad \text{(SC.164)}
\end{aligned}
$$

By the Cauchy-Schwarz inequality, Assumption SC3(iv), Lemma SC15 and (SC.93),

$$
n^{-1}\sum_{i=1}^{n}k_{1,i}(\hat{g}_i(\hat{\beta}_k)-\hat{g}_i(\beta_{k,0}))(\hat{g}_{1,i}(\beta_{k,0})-g_{1,i})=(\hat{\beta}_k-\beta_{k,0})o_p(1)+o_p(n^{-1/2}). \qquad \text{(SC.165)}
$$

Similarly, we can use Lemma SC15 to get

$$
\begin{aligned}
& n^{-1}\sum_{i=1}^{n}(\hat{g}_i(\hat{\beta}_k)-g(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k)-\hat{g}_i(\beta_{k,0})+g(\omega_{1,i}))(k_{2,i}-k_{1,i}g_{1,i}) \\
= \quad & (\hat{\beta}_k-\beta_{k,0})o_p(1)+O_p((m_2+m_1^{1/2})n^{-1/2}). \qquad \text{(SC.166)}
\end{aligned}
$$

Moreover, by Assumptions SC2(ii) and the consistency of $\hat{\beta}_k$

$$
\begin{aligned}
& n^{-1}\sum_{i=1}^{n}(g(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k)-g(\omega_{1,i}(\beta_{k,0});\beta_{k,0}))(k_{2,i}-k_{1,i}g_{1,i}) \\
= \quad & (\hat{\beta}_k-\beta_{k,0})n^{-1}\sum_{i=1}^{n}\frac{\partial g(\omega_{1,i}(\beta_{k,0});\beta_{k,0})}{\partial \beta_k}(k_{2,i}-k_{1,i}g_{1,i})+(\hat{\beta}_k-\beta_{k,0})o_p(1). \qquad \text{(SC.167)}
\end{aligned}
$$

Since

$$
\frac{\partial g(\omega_{1,i}(\beta_{k,0});\beta_{k,0})}{\partial \beta_k}=-a_2(\omega_{1,i})-g_1(\omega_{1,i})v_{1,i},
$$

by Assumptions SC1(i) and SC2(ii), and the Markov inequality,

$$
\begin{aligned}
& n^{-1}\sum_{i=1}^{n}\frac{\partial g(\omega_{1,i}(\beta_{k,0});\beta_{k,0})}{\partial \beta_k}(k_{2,i}-k_{1,i}g_{1,i}) \\
& = -n^{-1}\sum_{i=1}^{n}(a_2(\omega_{1,i})+v_{1,i}g_1(\omega_{1,i}))(k_{2,i}-k_{1,i}g_{1,i}) \\
& = -\mathbb{E}[(a_{2\,i}+v_{1,i}g_{1,i})(k_{2,i}-k_{1,i}g_{1,i})]+O_p(n^{-1/2})
\end{aligned}
$$

which together with (SC.167) implies that

$$
\begin{aligned}
& n^{-1}\sum_{i=1}^{n}(g(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k)-g(\omega_{1,i}(\beta_{k,0});\beta_{k,0}))(k_{2,i}-k_{1,i}g_{1,i}) \\
= & -(\hat{\beta}_k-\beta_{k,0})\left(\mathbb{E}[(a_{2\,i}+v_{1,i}g_{1,i})(k_{2,i}-k_{1,i}g_{1,i})]+o_p(1)\right)+o_p(n^{-1/2}). \qquad \text{(SC.168)}
\end{aligned}
$$

The claim of the lemma follows from (SC.164), (SC.165), (SC.166) and (SC.168). *Q.E.D.*

Lemma SC19. *Under Assumptions SC1, SC2 and SC3, we have*

$$
\hat{\beta}_k-\beta_{k,0}=O_p((m_1^{1/2}+m_2)n^{-1/2}). \qquad \text{(SC.169)}
$$

PROOF OF LEMMA SC19. Using Assumption SC1(iii), Lemma SC17 and Lemma SC18, we can use the decomposition in (SB.30) to deduce that

$$
\begin{aligned}
& n^{-1}\sum_{i=1}^{n}(\hat{J}_i(\hat{\beta}_k)-\hat{J}_i(\beta_{k,0})) \\
= & -(\hat{\beta}_k-\beta_{k,0})\left(\mathbb{E}[(v_{2,i}-v_{1,i}g_{1,i})^2]+o_p(1)\right) \\
& -n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(\hat{g}_{1,i}(\hat{\beta}_k)-\hat{g}_{1,i}(\beta_{k,0}))+O_p((m_1^{1/2}+m_2)n^{-1/2}). \qquad \text{(SC.170)}
\end{aligned}
$$

In view of the first order condition of $\hat{\beta}_k$, (SB.10), (SB.33), (SC.170) and Lemma SB2, the claim of the lemma follows if one can show that

$$
n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(\hat{g}_{1,i}(\hat{\beta}_k)-\hat{g}_{1,i}(\beta_{k,0}))=(\hat{\beta}_k-\beta_{k,0})o_p(1)+O_p((m_1^{1/2}+m_2)n^{-1/2}). \qquad \text{(SC.171)}
$$

We next prove the above claim. By the mean value expansion,

$$
\begin{aligned}
n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}\hat{g}_{1,i}(\hat{\beta}_k) &= n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}\partial^1\hat{P}_{2,i}(\hat{\beta}_k)'\hat{\beta}_g(\hat{\beta}_k) \\
&= n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}\partial^1\tilde{P}_{2,i}(\hat{\beta}_k)'\hat{\beta}_g(\hat{\beta}_k) \\
&\quad +n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(\hat{\phi}_i-\phi_i)\partial^2\tilde{P}_2(\tilde{\omega}_{1,i};\hat{\beta}_k)'\hat{\beta}_g(\hat{\beta}_k) \qquad \text{(SC.172)}
\end{aligned}
$$

where $\tilde{\omega}_{1,i}$ lies between $\hat{\omega}_{1,i}(\hat{\beta}_k)$ and $\omega_{1,i}(\hat{\beta}_k)$. By Assumption SC1(i), (SC.80) and the Markov inequality,

$$
n^{-1}\sum_{i=1}^{n}u_{2,i}^2k_{1,i}^2=O_p(1). \qquad \text{(SC.173)}
$$

By the triangle inequality

$$
\begin{aligned}
& \left| n^{-1} \sum_{i=1}^{n} u_{2,i} k_{1,i} (\partial^1 \tilde{P}_{2,i}(\hat{\beta}_k)' \hat{\beta}_g(\hat{\beta}_k) - g_1(\omega_{1,i}(\hat{\beta}_k); \hat{\beta}_k)) \right| \\
\leq & \left| n^{-1} \sum_{i=1}^{n} u_{2,i} k_{1,i} (g_1(\omega_{1,i}(\hat{\beta}_k); \hat{\beta}_k) - \partial^1 \tilde{P}_{2,i}(\hat{\beta}_k)' \tilde{\beta}_{g,m_2}(\hat{\beta}_k)) \right| \\
& + \left| n^{-1} \sum_{i=1}^{n} u_{2,i} k_{1,i} \partial^1 \tilde{P}_{2,i}(\hat{\beta}_k)' (\hat{\beta}_g(\hat{\beta}_k) - \tilde{\beta}_{g,m_2}(\hat{\beta}_k)) \right| . \qquad \text{(SC.174)}
\end{aligned}
$$

By Assumptions SC2(iii, vi) and (SC.173)

$$
n^{-1} \sum_{i=1}^{n} u_{2,i} k_{1,i} (g_1(\omega_{1,i}(\hat{\beta}_k); \hat{\beta}_k) - \partial^1 \tilde{P}_{2,i}(\hat{\beta}_k)' \tilde{\beta}_{g,m_2}(\hat{\beta}_k)) = O_p((m_1^{1/2} + m_2) n^{-1/2}). \qquad \text{(SC.175)}
$$

By the Cauchy-Schwarz inequality, Assumption SC2(v), Lemma SC14 and Lemma SC29

$$
\begin{aligned}
& \left| n^{-1} \sum_{i=1}^{n} u_{2,i} k_{1,i} \partial^1 \tilde{P}_{2,i}(\hat{\beta}_k)' (\hat{\beta}_g(\hat{\beta}_k) - \tilde{\beta}_{g,m_2}(\hat{\beta}_k)) \right| \\
\leq & \left\| n^{-1} \sum_{i=1}^{n} u_{2,i} k_{1,i} \partial^1 \tilde{P}_{2,i}(\hat{\beta}_k) \right\| \left\| \hat{\beta}_g(\hat{\beta}_k) - \tilde{\beta}_{g,m_2}(\hat{\beta}_k) \right\| \\
= & |\hat{\beta}_k - \beta_{k,0}| o_p(1) + O_p((m_1^{1/2} + m_2) n^{-1/2})
\end{aligned}
$$

which together with (SC.174) and (SC.175) implies that

$$
\begin{aligned}
& n^{-1} \sum_{i=1}^{n} u_{2,i} k_{1,i} (\partial^1 \tilde{P}_{2,i}(\hat{\beta}_k)' \hat{\beta}_g(\hat{\beta}_k) - g_1(\omega_{1,i}(\hat{\beta}_k); \hat{\beta}_k)) \\
= & |\hat{\beta}_k - \beta_{k,0}| o_p(1) + O_p((m_1^{1/2} + m_2) n^{-1/2}). \qquad \text{(SC.176)}
\end{aligned}
$$

By the Cauchy-Schwarz inequality and the triangle inequality, Assumptions SC2(ii, iii, v, vi) and

SC3(iv), Lemma SC4, Lemma SC14 and (SC.173)

$$
\begin{aligned}
& \left| n^{-1} \sum_{i=1}^{n} u_{2,i} k_{1,i} (\hat{\phi}_i - \phi_i) \partial^2 \tilde{P}_2(\tilde{\omega}_{1,i}; \hat{\beta}_k)' \hat{\beta}_g(\hat{\beta}_k) \right| \\
\leq \; & n^{-1} \sum_{i=1}^{n} \left| u_{2,i} k_{1,i} (\hat{\phi}_i - \phi_i) \partial^2 \tilde{P}_2(\tilde{\omega}_{1,i}; \hat{\beta}_k)' \tilde{\beta}_{g,m_2}(\hat{\beta}_k) \right| \\
& + n^{-1} \sum_{i=1}^{n} \left| u_{2,i} k_{1,i} (\hat{\phi}_i - \phi_i) \partial^2 \tilde{P}_2(\tilde{\omega}_{1,i}; \hat{\beta}_k)' (\hat{\beta}_g(\hat{\beta}_k) - \tilde{\beta}_{g,m_2}(\hat{\beta}_k)) \right| \\
\leq \; & \left(\sup_{\omega \in \Omega_{C_\omega}} \left| \partial^2 \tilde{P}_2(\omega; \hat{\beta}_k)' \tilde{\beta}_{g,m_2}(\hat{\beta}_k) \right| + \xi_{2,m_2} \left\| \hat{\beta}_g(\hat{\beta}_k) - \tilde{\beta}_{g,m_2}(\hat{\beta}_k) \right\| \right) n^{-1} \sum_{i=1}^{n} \left| u_{2,i} k_{1,i} (\hat{\phi}_i - \phi_i) \right| \\
= \; & |\hat{\beta}_k - \beta_{k,0}| o_p(1) + O_p((m_1^{1/2} + m_2) n^{-1/2}) \qquad \text{(SC.177)}
\end{aligned}
$$

which together with (SC.172), (SC.176) and (SC.177) proves

$$
n^{-1} \sum_{i=1}^{n} u_{2,i} k_{1,i} (\hat{g}_{1,i}(\hat{\beta}_k) - g_1(\omega_{1,i}(\hat{\beta}_k); \hat{\beta}_k)) = (\hat{\beta}_k - \beta_{k,0}) o_p(1) + O_p((m_1^{1/2} + m_2) n^{-1/2}). \quad \text{(SC.178)}
$$

Similarly, we can show that

$$
n^{-1} \sum_{i=1}^{n} u_{2,i} k_{1,i} (\hat{g}_{1,i}(\beta_{k,0}) - g_1(\omega_{1,i}(\beta_{k,0}); \beta_{k,0})) = O_p((m_1^{1/2} + m_2) n^{-1/2})
$$

which together with (SC.178) implies that

$$
\begin{aligned}
& n^{-1} \sum_{i=1}^{n} u_{2,i} k_{1,i} (\hat{g}_{1,i}(\hat{\beta}_k) - \hat{g}_{1,i}(\beta_{k,0})) \\
= \; & n^{-1} \sum_{i=1}^{n} u_{2,i} k_{1,i} (g_1(\omega_{1,i}(\hat{\beta}_k); \hat{\beta}_k) - g_1(\omega_{1,i}(\beta_{k,0}); \beta_{k,0})) \\
& + (\hat{\beta}_k - \beta_{k,0}) o_p(1) + O_p((m_1^{1/2} + m_2) n^{-1/2}). \qquad \text{(SC.179)}
\end{aligned}
$$

By Assumption SC2(ii), the Markov inequality and the consistency of $\hat{\beta}_k$,

$$
n^{-1} \sum_{i=1}^{n} u_{2,i} k_{1,i} (g_1(\omega_{1,i}(\hat{\beta}_k); \hat{\beta}_k) - g_1(\omega_{1,i}(\beta_{k,0}); \beta_{k,0})) = (\hat{\beta}_k - \beta_{k,0}) o_p(1) + O_p((m_1^{1/2} + m_2) n^{-1/2})
$$

which together with (SC.179) proves (SC.171). *Q.E.D.*

Lemma SC20. *Under Assumptions SC1, SC2 and SC3, we have*

$$(B(\beta_{k,0})')^{-1}B(\hat{\beta}_k)'\hat{\beta}_g(\hat{\beta}_k) - \hat{\beta}_g(\beta_{k,0}) = (\hat{\beta}_k - \beta_{k,0})O_p(\xi_{1,m_2}).$$

Proof of Lemma SC20. We define $\hat{\mathbf{P}}_2^*(\beta_k) = (\hat{P}_{2,1}^*(\beta_k), \ldots, \hat{P}_{2,n}^*(\beta_k))'$ where $\hat{P}_{2,i}^*(\beta_k) = B(\beta_{k,0})P_2(\hat{\omega}_{1,i}(\beta_k))$. Then we can write

$$(B(\beta_{k,0})')^{-1}B(\hat{\beta}_k)'\hat{\beta}_g(\hat{\beta}_k) = (\hat{\mathbf{P}}_2^*(\hat{\beta}_k)'\hat{\mathbf{P}}_2^*(\hat{\beta}_k))^{-1}\hat{\mathbf{P}}_2^*(\hat{\beta}_k)'\hat{\mathbf{Y}}_2^*(\hat{\beta}_k)$$

and therefore,

$$\begin{aligned}
& (B(\beta_{k,0})')^{-1}B(\hat{\beta}_k)'\hat{\beta}_g(\hat{\beta}_k) - \hat{\beta}_g(\beta_{k,0}) \\
= \; & \left[(\hat{\mathbf{P}}_2^*(\hat{\beta}_k)'\hat{\mathbf{P}}_2^*(\hat{\beta}_k))^{-1} - (\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}\right]\hat{\mathbf{P}}_2^*(\hat{\beta}_k)'\hat{\mathbf{Y}}_2^*(\hat{\beta}_k) \\
& + (\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}(\hat{\mathbf{P}}_2^*(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\beta_{k,0}))'\hat{\mathbf{Y}}_2^*(\beta_k) \\
& + (\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'(\hat{\mathbf{Y}}_2^*(\hat{\beta}_k) - \hat{\mathbf{Y}}_2^*(\beta_{k,0})).
\end{aligned} \quad \text{(SC.180)}$$

By (SC.58), Assumption SC1(i) and the Markov inequality,

$$\begin{aligned}
& \left\|(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'\mathbf{K}_2\right\|^2 \\
& \leq (\lambda_{\min}(n^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0})))^{-1}n^{-1}\sum_{i=1}^{n}k_{2,i}^2 = O_p(1).
\end{aligned} \quad \text{(SC.181)}$$

Since $\hat{\mathbf{Y}}_2^*(\hat{\beta}_k) - \hat{\mathbf{Y}}_2^*(\beta_{k,0}) = -(\hat{\beta}_k - \beta_{k,0})\mathbf{K}_2$, by (SC.181) we get

$$(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'(\hat{\mathbf{Y}}_2^*(\hat{\beta}_k) - \hat{\mathbf{Y}}_2^*(\beta_{k,0})) = (\hat{\beta}_k - \beta_{k,0})O_p(1). \quad \text{(SC.182)}$$

By the mean value expansion, we have for any $\upsilon_2 \in \mathbb{R}^{m_2}$,

$$\upsilon_2'(\tilde{P}_2(\hat{\omega}_{1,i}(\hat{\beta}_k);\beta_{k,0}) - \tilde{P}_2(\hat{\omega}_{1,i}(\beta_{k,0});\beta_{k,0})) = -\upsilon_2'\partial^1\tilde{P}_2(\hat{\omega}_{1,i}(\tilde{\beta}_k);\beta_{k,0})k_{1,i}(\hat{\beta}_k - \beta_{k,0}) \quad \text{(SC.183)}$$

where $\tilde{\beta}_k$ lies between $\hat{\beta}_k$ and $\beta_{k,0}$. By Assumption SC3(iv) and Lemma SC19, $\hat{\omega}_{1,i}(\tilde{\beta}_k) \in \Omega_{\varepsilon_n}(\beta_{k,0})$ for any $i = 1, \ldots, n$ wpa1. By the Cauchy-Schwarz inequality and (SC.183)

$$\left|\upsilon_2'(\hat{P}_{2,i}^*(\hat{\beta}_k) - \hat{P}_{2,i}(\beta_{k,0}))\right| \leq \|\upsilon_2\|\,\xi_{1,m_2}\left|k_{1,i}(\hat{\beta}_k - \beta_{k,0})\right| \quad \text{(SC.184)}$$

wpa1. Therefore we have wpa1,

$$\begin{aligned}&\upsilon_2'(\hat{\mathbf{P}}_2^*(\hat{\beta}_k)-\hat{\mathbf{P}}_2(\beta_{k,0}))'(\hat{\mathbf{P}}_2^*(\hat{\beta}_k)-\hat{\mathbf{P}}_2(\beta_{k,0}))\upsilon_2\\&=\sum_{i=1}^n(\upsilon_2'(\hat{P}_{2,i}^*(\hat{\beta}_k)-\hat{P}_{2,i}(\beta_{k,0}))^2\leq\|\upsilon_2\|^2\,\xi_{1,m_2}^2(\hat{\beta}_k-\beta_{k,0})^2\sum_{i=1}^n k_{1,i}^2\end{aligned}\tag{SC.185}$$

which implies that

$$||\hat{\mathbf{P}}_2^*(\hat{\beta}_k)-\hat{\mathbf{P}}_2(\beta_{k,0})||_S=|\hat{\beta}_k-\beta_{k,0}|O_p(\xi_{1,m_2}n^{1/2}).\tag{SC.186}$$

Since $y_{2,i}^*(\beta_k)=y_{2,i}^*-\beta_k k_{2,i}$, by the Cauchy-Schwarz inequality we get

$$n^{-1}\sum_{i=1}^n(y_{2,i}^*(\beta_k))^2\leq 8\left(n^{-1}\sum_{i=1}^n(y_{2,i}^*)^2+\beta_k^2n^{-1}\sum_{i=1}^n k_{2,i}^2\right)$$

which together with the Markov inequality, Assumption SC2(i) and the compactness of Θ_k implies that

$$\sup_{\beta_k\in\Theta_k}n^{-1}\sum_{i=1}^n(y_{2,i}^*(\beta_k))^2=O_p(1).\tag{SC.187}$$

By the Cauchy-Schwarz inequality,(SC.58), (SC.186) and (SC.187),

$$\begin{aligned}&\left\|(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}(\hat{\mathbf{P}}_2^*(\hat{\beta}_k)-\hat{\mathbf{P}}_2(\beta_{k,0}))'\hat{\mathbf{Y}}_2^*(\beta_k)\right\|\\&\leq(\lambda_{\min}(n^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0})))^{-1}n^{-1}||\hat{\mathbf{P}}_2^*(\hat{\beta}_k)-\hat{\mathbf{P}}_2(\beta_{k,0})||_S\left\|\hat{\mathbf{Y}}_2^*(\hat{\beta}_k)\right\|\\&=|\hat{\beta}_k-\beta_{k,0}|O_p(\xi_{1,m_2}).\end{aligned}\tag{SC.188}$$

By the definition of $\hat{\beta}_g(\hat{\beta}_k)$, we can write

$$\begin{aligned}&[(\hat{\mathbf{P}}_2^*(\hat{\beta}_k)'\hat{\mathbf{P}}_2^*(\hat{\beta}_k))^{-1}-(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}]\hat{\mathbf{P}}_2^*(\hat{\beta}_k)'\hat{\mathbf{Y}}_2^*(\hat{\beta}_k)\\=\;&(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}(\hat{\mathbf{P}}_2(\beta_{k,0})-\hat{\mathbf{P}}_2^*(\hat{\beta}_k))'\hat{\mathbf{P}}_2(\hat{\beta}_k)\hat{\beta}_g(\hat{\beta}_k)\\&+(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'(\hat{\mathbf{P}}_2(\beta_{k,0})-\hat{\mathbf{P}}_2^*(\hat{\beta}_k))(B(\beta_{k,0})')^{-1}B(\hat{\beta}_k)'\hat{\beta}_g(\hat{\beta}_k).\end{aligned}\tag{SC.189}$$

By the Cauchy-Schwarz inequality, Assumption SC2(vi), (SC.58), (SC.70) and (SC.186),

$$\begin{aligned}&\left\|(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}(\hat{\mathbf{P}}_2(\beta_{k,0})-\hat{\mathbf{P}}_2^*(\hat{\beta}_k))'\hat{\mathbf{P}}_2(\hat{\beta}_k)\hat{\beta}_g(\hat{\beta}_k)\right\|\\\leq\;&(\lambda_{\min}(n^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0})))^{-1}n^{-1}||\hat{\mathbf{P}}_2^*(\hat{\beta}_k)-\hat{\mathbf{P}}_2(\beta_{k,0})||_S\left\|\hat{\mathbf{P}}_2(\hat{\beta}_k)\hat{\beta}_g(\hat{\beta}_k)\right\|\\=\;&|\hat{\beta}_k-\beta_{k,0}|O_p(\xi_{1,m_2}).\end{aligned}\tag{SC.190}$$

By the definition of $\hat{\mathbf{P}}_2(\beta_{k,0})$ and $\hat{\mathbf{P}}_2^*(\hat{\beta}_k)$, and the mean value expansion

$$\begin{aligned} & \left\| (\hat{\mathbf{P}}_2(\beta_{k,0}) - \hat{\mathbf{P}}_2^*(\hat{\beta}_k))(B(\beta_{k,0})')^{-1} B(\hat{\beta}_k)' \hat{\beta}_g(\hat{\beta}_k) \right\|^2 \\ = & \sum_{i=1}^n (\hat{\beta}_g(\hat{\beta}_k)' B(\hat{\beta}_k)(P_2(\hat{\omega}_{1,i}(\beta_{k,0})) - P_2(\hat{\omega}_{1,i}(\hat{\beta}_k))))^2 \\ = & \sum_{i=1}^n (\hat{\beta}_g(\hat{\beta}_k)' \partial^1 \tilde{P}_2(\hat{\omega}_{1,i}(\tilde{\beta}_k); \hat{\beta}_k) k_{1,i}(\hat{\beta}_k - \beta_{k,0}))^2 \\ \leq & (\hat{\beta}_k - \beta_{k,0})^2 \max_{i \leq n} (\partial^1 \tilde{P}_2(\hat{\omega}_{1,i}(\tilde{\beta}_k); \hat{\beta}_k)' \hat{\beta}_g(\hat{\beta}_k))^2 \sum_{i=1}^n k_{1,i}^2 \end{aligned}$$

which together with Assumptions SC2(ii, iii), Lemma SC14 and Lemma SC19 implies that

$$\left\| (\hat{\mathbf{P}}_2(\beta_{k,0}) - \hat{\mathbf{P}}_2^*(\hat{\beta}_k))(B(\beta_{k,0})')^{-1} B(\hat{\beta}_k)' \hat{\beta}_g(\hat{\beta}_k) \right\| = |\hat{\beta}_k - \beta_{k,0}| O_p(n^{1/2}). \tag{SC.191}$$

By (SC.58) and (SC.191)

$$(\hat{\mathbf{P}}_2(\beta_{k,0})' \hat{\mathbf{P}}_2(\beta_{k,0}))^{-1} \hat{\mathbf{P}}_2(\beta_{k,0})' (\hat{\mathbf{P}}_2(\beta_{k,0}) - \hat{\mathbf{P}}_2^*(\hat{\beta}_k))(B(\beta_{k,0})')^{-1} B(\hat{\beta}_k)' \hat{\beta}_g(\hat{\beta}_k) = (\hat{\beta}_k - \beta_{k,0}) O_p(1). \tag{SC.192}$$

Collecting the results in (SC.189), (SC.190) and (SC.192) we get

$$[(\hat{\mathbf{P}}_2^*(\hat{\beta}_k) \hat{\mathbf{P}}_2^*(\hat{\beta}_k)')^{-1} - (\hat{\mathbf{P}}_2(\beta_{k,0})' \hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}] \hat{\mathbf{P}}_2^*(\hat{\beta}_k)' \hat{\mathbf{Y}}_2^*(\hat{\beta}_k) = (\hat{\beta}_k - \beta_{k,0}) O_p(\xi_{1,m_2}). \tag{SC.193}$$

The claim of the lemma follows from (SC.180), (SC.182), (SC.188) and (SC.193). *Q.E.D.*

Lemma SC21. *Under Assumptions SC1, SC2 and SC3, we have*

$$n^{-1} \sum_{i=1}^n u_{2,i} k_{1,i} (\hat{g}_{1,i}(\hat{\beta}_k) - \hat{g}_{1,i}(\beta_{k,0})) = (\hat{\beta}_k - \beta_{k,0}) o_p(1) + o_p(n^{-1/2}).$$

PROOF OF LEMMA SC21. By the definition of $\hat{g}_{1,i}(\hat{\beta}_k)$, we can write

$$\hat{g}_{1,i}(\hat{\beta}_k) = \partial^1 \tilde{P}_2(\hat{\omega}_{1,i}(\hat{\beta}_k); \hat{\beta}_k)' \hat{\beta}_g(\hat{\beta}_k) = \partial^1 \tilde{P}_2(\hat{\omega}_{1,i}(\hat{\beta}_k); \beta_{k,0})' \upsilon_{2,*}$$

where $\upsilon_{2,*} = (\hat{\mathbf{P}}_2^*(\hat{\beta}_k)'\hat{\mathbf{P}}_2^*(\hat{\beta}_k))^{-1}\hat{\mathbf{P}}_2^*(\hat{\beta}_k)'\hat{\mathbf{Y}}_2^*(\hat{\beta}_k)$. Therefore

$$\begin{aligned}
& n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}(\hat{g}_{1,i}(\hat{\beta}_k) - \hat{g}_{1,i}(\beta_{k,0})) \\
&= n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}(\partial^1\tilde{P}_2(\hat{\omega}_{1,i}(\hat{\beta}_k);\beta_{k,0}) - \partial^1\tilde{P}_2(\omega_{1,i}(\beta_{k,0});\beta_{k,0}))'\upsilon_{2,*} \\
&\quad - n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}(\partial^1\tilde{P}_2(\hat{\omega}_{1,i}(\beta_{k,0});\beta_{k,0}) - \partial^1\tilde{P}_2(\omega_{1,i}(\beta_{k,0});\beta_{k,0}))'\hat{\beta}_g(\beta_{k,0}) \\
&\quad + n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}\partial^1\tilde{P}_2(\omega_{1,i}(\beta_{k,0});\beta_{k,0})'(\upsilon_{2,*} - \hat{\beta}_g(\beta_{k,0})). \qquad \text{(SC.194)}
\end{aligned}$$

Since $\upsilon_{2,*} = (B(\beta_{k,0})')^{-1}B(\hat{\beta}_k)'\hat{\beta}_g(\hat{\beta}_k)$, by Assumption SC3(iv), Lemma SC19, Lemma SC20 and (SC.70)

$$\|\upsilon_{2,*}\| = O_p(1). \qquad \text{(SC.195)}$$

By the second order expansion,

$$\begin{aligned}
& n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}(\partial^1\tilde{P}_2(\hat{\omega}_{1,i}(\hat{\beta}_k);\beta_{k,0}) - \partial^1\tilde{P}_2(\omega_{1,i}(\beta_{k,0});\beta_{k,0}))'\upsilon_{2,*} \\
&= n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}(\hat{\omega}_{1,i}(\hat{\beta}_k) - \omega_{1,i})\partial^2\tilde{P}_{2,i}(\beta_{k,0})'\upsilon_{2,*} \\
&\quad + n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}(\hat{\omega}_{1,i}(\hat{\beta}_k) - \omega_{1,i})^2\partial^3\tilde{P}_2(\tilde{\omega}_{1,i};\beta_{k,0})'\upsilon_{2,*} \qquad \text{(SC.196)}
\end{aligned}$$

where $\tilde{\omega}_{1,i}$ lies between $\hat{\omega}_{1,i}(\hat{\beta}_k)$ and $\omega_{1,i}$. Since $\hat{\omega}_{1,i}(\hat{\beta}_k) - \omega_{1,i} = \hat{\phi}_i - \phi_i - k_{1,i}(\hat{\beta}_k - \beta_{k,0})$,

$$\begin{aligned}
& n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}(\hat{\omega}_{1,i}(\hat{\beta}_k) - \omega_{1,i})\partial^2\tilde{P}_2(\omega_{1,i}(\beta_{k,0});\beta_{k,0})'\upsilon_{2,*} \\
&= n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}(\hat{\phi}_i - \phi_i)\partial^2\tilde{P}_{2,i}(\beta_{k,0})'\upsilon_{2,*} \\
&\quad - (\hat{\beta}_k - \beta_{k,0})n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}^2\partial^2\tilde{P}_{2,i}(\beta_{k,0})'\upsilon_{2,*}. \qquad \text{(SC.197)}
\end{aligned}$$

By the Cauchy-Schwarz inequality, Assumptions SC1(i) and SC2(v), (SC.80) and (SC.195),

$$\left|n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}^2\partial^2\tilde{P}_{2,i}(\beta_{k,0})'\upsilon_{2,*}\right| \leq \left\|n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}^2\partial^2\tilde{P}_{2,i}(\beta_{k,0})\right\|\|\upsilon_{2,*}\| = O_p(\xi_{2,m_2}n^{-1/2})$$

which together with Assumption SC3(iv) implies that

$$(\hat{\beta}_k - \beta_{k,0})n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}^2\partial^2\tilde{P}_{2,i}(\beta_{k,0})'\upsilon_{2,*} = (\hat{\beta}_k - \beta_{k,0})o_p(1). \quad \text{(SC.198)}$$

By the Cauchy-Schwarz inequality, Assumptions SC1(i, iv, vi), SC2(v) and SC3(iv), (SC.80) and (SC.195),

$$n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}(\phi_{m_1,i} - \phi_i)\partial^2\tilde{P}_2(\omega_{1,i}(\beta_{k,0});\beta_{k,0})'\upsilon_{2,*} = O_p(\xi_{2,m_2}n^{-1}) = o_p(n^{-1/2}). \quad \text{(SC.199)}$$

By the Cauchy-Schwarz inequality, Assumptions SC1(i, iv, vi), SC2(v) and SC3(iv), (SC.47), (SC.80) and (SC.195),

$$(\hat{\beta}_\phi - \beta_{\phi,m_1})'n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}P_1(x_{1,i})\partial^2\tilde{P}_2(\omega_{1,i}(\beta_{k,0});\beta_{k,0})'\upsilon_{2,*} = O_p(m_1\xi_{2,m_2}n^{-1}) = o_p(n^{-1/2})$$

which together with (SC.197), (SC.198) and (SC.199) implies that

$$n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}(\hat{\omega}_{1,i}(\hat{\beta}_k) - \omega_{1,i})\partial^2\tilde{P}_2(\omega_{1,i}(\beta_{k,0});\beta_{k,0})'\upsilon_{2,*} = (\hat{\beta}_k - \beta_{k,0})o_p(1) + o_p(n^{-1/2}). \quad \text{(SC.200)}$$

Using the similar arguments in showing (SC.84), we can show that

$$n^{-1}\sum_{i=1}^{n} \left|u_{2,i}k_{1,i}(\hat{\phi}_i - \phi_i)^2\right| = O_p(m_1 n^{-1}). \quad \text{(SC.201)}$$

Moreover by the Markov inequality, Assumption SC1(i) and (SC.80)

$$n^{-1}\sum_{i=1}^{n}(\left|u_{2,i}k_{1,i}^2\right| + \left|u_{2,i}k_{1,i}\right|) = O_p(1). \quad \text{(SC.202)}$$

By the Cauchy-Schwarz inequality, Assumption SC3(iv), Lemma SC19, (SC.201) and (SC.202)

$$\begin{aligned} & n^{-1}\sum_{i=1}^{n}\left|u_{2,i}k_{1,i}\right|(\hat{\omega}_{1,i}(\hat{\beta}_k) - \omega_{1,i})^2 \\ \leq\ & 2n^{-1}\sum_{i=1}^{n}\left|u_{2,i}k_{1,i}(\hat{\phi}_i - \phi_i)^2\right| + 2(\hat{\beta}_k - \beta_{k,0})^2 n^{-1}\sum_{i=1}^{n}\left|u_{2,i}k_{1,i}^2\right| = o_p(n^{-1/2}). \quad \text{(SC.203)}\end{aligned}$$

By the Cauchy-Schwarz inequality, Assumptions SC2(ii, iii, v) and SC3(iv), Lemma SC14 and

(SC.203)

$$
\begin{aligned}
&\left|n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(\hat{\omega}_{1,i}(\hat{\beta}_k)-\omega_{1,i})^2\partial^3\tilde{P}_2(\tilde{\omega}_{1,i};\beta_{k,0})'\upsilon_{2,*}\right| \\
\leq\ & \left|n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(\hat{\omega}_{1,i}(\hat{\beta}_k)-\omega_{1,i})^2\partial^3\tilde{P}_2(\tilde{\omega}_{1,i};\hat{\beta}_k)'\tilde{\beta}_{g,m_2}(\hat{\beta}_k)\right| \\
&+\left|n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(\hat{\omega}_{1,i}(\hat{\beta}_k)-\omega_{1,i})^2\partial^3\tilde{P}_2(\tilde{\omega}_{1,i};\hat{\beta}_k)'(\hat{\beta}_g(\hat{\beta}_k)-\tilde{\beta}_{g,m_2}(\hat{\beta}_k))\right| \\
=\ & (\hat{\beta}_k-\beta_{k,0})o_p(1)+o_p(n^{-1/2}) \qquad \text{(SC.204)}
\end{aligned}
$$

which together with (SC.196) and (SC.200) implies that

$$
n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(\partial^1\tilde{P}_2(\hat{\omega}_{1,i}(\hat{\beta}_k);\beta_{k,0})-\partial^1\tilde{P}_{2,i}(\beta_{k,0}))'\upsilon_{2,*}=(\hat{\beta}_k-\beta_{k,0})o_p(1)+o_p(n^{-1/2}). \qquad \text{(SC.205)}
$$

Using similar arguments in proving (SC.205), we can show that

$$
n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(\partial^1\tilde{P}_2(\hat{\omega}_{1,i}(\beta_{k,0});\beta_{k,0})-\partial^1\tilde{P}_{2,i}(\beta_{k,0}))'\hat{\beta}_g(\beta_{k,0})=o_p(n^{-1/2}). \qquad \text{(SC.206)}
$$

By the Cauchy-Schwarz inequality, Assumptions SC1(i), SC2(v) and SC3(iv), Lemma SC20 and (SC.80)

$$
\begin{aligned}
&\left\|n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}\partial^1\tilde{P}_{2,i}(\beta_{k,0})'(\upsilon_{2,*}-\hat{\beta}_g(\beta_{k,0}))\right\| \\
\leq\ & \left\|n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}\partial^1\tilde{P}_{2,i}(\beta_{k,0})\right\|\left\|\upsilon_{2,*}-\hat{\beta}_g(\beta_{k,0})\right\|=(\hat{\beta}_k-\beta_{k,0})o_p(1). \qquad \text{(SC.207)}
\end{aligned}
$$

The claim of the lemma follows from (SC.194), (SC.205), (SC.206) and (SC.207). *Q.E.D.*

Lemma SC22. *Let* $\mathbf{U}_2=(u_{2,1},\ldots,u_{2,n})'$, $\hat{\mathbf{G}}_n=(\hat{g}(\hat{\omega}_{1,1}(\hat{\beta}_k);\hat{\beta}_k),\ldots,\hat{g}(\hat{\omega}_{1,n}(\hat{\beta}_k);\hat{\beta}_k))'$ *and* $\mathbf{G}_n=(g(\omega_{1,1}),\ldots,g(\omega_{1,n}))'$. *Then under Assumptions SC1, SC2 and SC3, we have*

(i) $n^{-1}\mathbf{U}_2'(\hat{\mathbf{P}}_2^*(\hat{\beta}_k)-\hat{\mathbf{P}}_2(\beta_{k,0}))\hat{\beta}_\varphi(\beta_{k,0})=(\hat{\beta}_k-\beta_{k,0})o_p(1)+o_p(n^{-1/2})$

(ii) $n^{-1}\mathbf{L}_2'(\hat{\mathbf{P}}_2^*(\hat{\beta}_k)-\hat{\mathbf{P}}_2(\beta_{k,0}))\hat{\beta}_\varphi(\beta_{k,0})=o_p(1)$*;*

(iii) $n^{-1}\mathbf{K}_2'(\hat{\mathbf{P}}_2^*(\hat{\beta}_k)-\hat{\mathbf{P}}_2(\beta_{k,0}))\hat{\beta}_\varphi(\beta_{k,0})=o_p(1)$*;*

(iv) $n^{-1}(\hat{\mathbf{G}}_n-\mathbf{G}_n)'(\hat{\mathbf{P}}_2^*(\hat{\beta}_k)-\hat{\mathbf{P}}_2(\beta_{k,0}))\hat{\beta}_\varphi(\beta_{k,0})=(\hat{\beta}_k-\beta_{k,0})o_p(1)+o_p(n^{-1/2})$.

Proof of Lemma SC22. (i) By the first order expansion,

$$
\begin{aligned}
& n^{-1}\mathbf{U}_2'(\hat{\mathbf{P}}_2^*(\hat{\beta}_k)-\hat{\mathbf{P}}_2(\beta_{k,0}))(\hat{\beta}_{\varphi}(\beta_{k,0})-\tilde{\beta}_{\varphi,m_2}(\beta_{k,0})) \\
= & -(\hat{\beta}_k-\beta_{k,0})n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}\partial^1\tilde{P}_2(\tilde{\omega}_{1,i};\beta_{k,0})'(\hat{\beta}_{\varphi}(\beta_{k,0})-\tilde{\beta}_{\varphi,m_2}(\beta_{k,0})) \qquad \text{(SC.208)}
\end{aligned}
$$

where $\tilde{\omega}_{1,i}$ lies between $\hat{\omega}_{1,i}(\hat{\beta}_k)$ and $\hat{\omega}_{1,i}(\beta_{k,0})$. By Assumptions SC2(v, vi) and SC3(iv), (SC.117) and (SC.202),

$$
n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}\partial^1\tilde{P}_2(\tilde{\omega}_{1,i};\beta_{k,0})'(\hat{\beta}_{\varphi}(\beta_{k,0})-\tilde{\beta}_{\varphi,m_2}(\beta_{k,0}))=O_p(\xi_{1,m_2}(m_1^{1/2}+m_2^{1/2})n^{-1/2})=o_p(1)
$$

which together with (SC.208) implies that

$$
n^{-1}\mathbf{U}_2'(\hat{\mathbf{P}}_2^*(\hat{\beta}_k)-\hat{\mathbf{P}}_2(\beta_{k,0}))(\hat{\beta}_{\varphi}(\beta_{k,0})-\tilde{\beta}_{\varphi,m_2}(\beta_{k,0}))=(\hat{\beta}_k-\beta_{k,0})o_p(1). \qquad \text{(SC.209)}
$$

By Assumptions SC3(i, ii, iv), Lemma SC4, Lemma SC19, (SC.80) and (SC.202)

$$
\begin{aligned}
& n^{-1}\mathbf{U}_2'(\hat{\mathbf{P}}_2^*(\hat{\beta}_k)-\hat{\mathbf{P}}_2(\beta_{k,0}))\tilde{\beta}_{\varphi,m_2}(\beta_{k,0}) \\
= & n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(\varphi(\hat{\omega}_{1,i}(\hat{\beta}_k))-\varphi(\hat{\omega}_{1,i}(\beta_{k,0})))+o_p(n^{-1/2}) \\
= & n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}\varphi_1(\hat{\omega}_{1,i}(\beta_{k,0}))(\hat{\beta}_k-\beta_{k,0})+o_p(n^{-1/2}) \\
= & n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}\varphi_1(\omega_{1,i})(\hat{\beta}_k-\beta_{k,0})+(\hat{\beta}_k-\beta_{k,0})o_p(1)+o_p(n^{-1/2}) \qquad \text{(SC.210)}
\end{aligned}
$$

By the Markov inequality, Assumptions SC1(i) and SC3(i), (SC.80).

$$
n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}\varphi_1(\omega_{1,i})=O_p(n^{-1/2}) \qquad \text{(SC.211)}
$$

which together with (SC.210) implies that

$$
n^{-1}\mathbf{U}_2'(\hat{\mathbf{P}}_2^*(\hat{\beta}_k)-\hat{\mathbf{P}}_2(\beta_{k,0}))\tilde{\beta}_{\varphi,m_2}(\beta_{k,0})=(\hat{\beta}_k-\beta_{k,0})o_p(1)+o_p(n^{-1/2}). \qquad \text{(SC.212)}
$$

The first claim of the lemma follows by (SC.209) and (SC.212).

(ii) Using the similar arguments in showing (SC.196), we get

$$
n^{-1}\mathbf{L}_2'(\hat{\mathbf{P}}_2^*(\hat{\beta}_k)-\hat{\mathbf{P}}_2(\beta_{k,0}))(\hat{\beta}_{\varphi}(\beta_{k,0})-\tilde{\beta}_{\varphi,m_2}(\beta_{k,0}))=(\hat{\beta}_k-\beta_{k,0})o_p(1). \qquad \text{(SC.213)}
$$

By the mean value expansion, Assumptions SC3(i, ii, iv), Lemma SC4, the consistency of $\hat{\beta}_k$ and the Markov inequality

$$\begin{aligned}
& n^{-1}\mathbf{L}_2'(\hat{\mathbf{P}}_2^*(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\beta_{k,0}))\tilde{\beta}_{\varphi,m_2}(\beta_{k,0}) \\
= \; & n^{-1}\sum_{i=1}^{n} l_{2,i}k_{1,i}(\varphi(\hat{\omega}_{1,i}(\hat{\beta}_k)) - \varphi(\hat{\omega}_{1,i}(\beta_{k,0}))) + o_p(1) \\
= \; & -(\hat{\beta}_k - \beta_{k,0})n^{-1}\sum_{i=1}^{n} l_{2,i}k_{1,i}\varphi_1(\hat{\omega}_{1,i}(\beta_{k,0})) + o_p(1) = o_p(1)
\end{aligned}$$

which together with (SC.213) finishes the proof.

(iii) The third claim of the lemma can be proved the same way as the second one.

(iv) By the first-order expansion,

$$\begin{aligned}
& n^{-1}(\hat{\mathbf{G}}_n - \mathbf{G}_n)'(\hat{\mathbf{P}}_2^*(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\beta_{k,0}))(\hat{\beta}_{\varphi}(\beta_{k,0}) - \tilde{\beta}_{\varphi,m_2}(\beta_{k,0})) \\
= \; & n^{-1}\sum_{i=1}^{n}(\hat{g}_i(\hat{\beta}_k) - g(\omega_{1,i}))(\tilde{P}_2(\hat{\omega}_{1,i}(\hat{\beta}_k);\beta_{k,0}) - \hat{P}_{2,i}(\beta_{k,0}))'(\hat{\beta}_{\varphi}(\beta_{k,0}) - \tilde{\beta}_{\varphi,m_2}(\beta_{k,0})) \\
= \; & -(\hat{\beta}_k - \beta_{k,0})n^{-1}\sum_{i=1}^{n}(\hat{g}_i(\hat{\beta}_k) - g(\omega_{1,i}))\partial^1\tilde{P}_2(\tilde{\omega}_{1,i};\beta_{k,0})'(\hat{\beta}_{\varphi}(\beta_{k,0}) - \tilde{\beta}_{\varphi,m_2}(\beta_{k,0})) \quad \text{(SC.214)}
\end{aligned}$$

where $\tilde{\omega}_{1,i}$ lies between $\hat{\omega}_{1,i}(\hat{\beta}_k)$ and $\hat{\omega}_{1,i}(\beta_{k,0})$. By Assumption SC2(v), Lemma SC15 and (SC.117), we get

$$\begin{aligned}
& n^{-1}\sum_{i=1}^{n}(\hat{g}_i(\hat{\beta}_k) - g(\omega_{1,i}))\partial^1\tilde{P}_2(\tilde{\omega}_{1,i};\beta_{k,0})'(\hat{\beta}_{\varphi}(\beta_{k,0}) - \tilde{\beta}_{\varphi,m_2}(\beta_{k,0})) \\
= \; & \xi_{1,m_2}O_p((m_1^{1/2} + m_2)n^{-1/2})O_p((m_1^{1/2} + m_2^{1/2})n^{-1/2})
\end{aligned}$$

which together with Assumption SC3(iv) and (SC.214) implies that

$$n^{-1}(\hat{\mathbf{G}}_n - \mathbf{G}_n)'(\hat{\mathbf{P}}_2^*(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\beta_{k,0}))(\hat{\beta}_{\varphi}(\beta_{k,0}) - \tilde{\beta}_{\varphi,m_2}(\beta_{k,0})) = (\hat{\beta}_k - \beta_{k,0})o_p(1). \quad \text{(SC.215)}$$

Using Assumptions SC3(i, ii, iv) and Lemma SC15, we get

$$\begin{aligned}
& n^{-1}(\hat{\mathbf{G}}_n - \mathbf{G}_n)'(\hat{\mathbf{P}}_2^*(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\beta_{k,0}))\tilde{\beta}_{\varphi,m_2}(\beta_{k,0}) \\
= \; & n^{-1}\sum_{i=1}^{n}(\hat{g}_i(\hat{\beta}_k) - g(\omega_{1,i}))(\varphi(\hat{\omega}_{1,i}(\hat{\beta}_k)) - \varphi(\hat{\omega}_{1,i}(\beta_{k,0}))) + o_p(n^{-1/2}) \\
= \; & (\hat{\beta}_k - \beta_{k,0})o_p(1) + o_p(n^{-1/2})
\end{aligned}$$

which together with (SC.215) proves the claim. *Q.E.D.*

Lemma SC23. *Under Assumptions SC1, SC2 and SC3, we have*

$$\begin{aligned}
& n^{-1}\sum_{i=1}^{n}(\hat{g}(\hat{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k)-\hat{g}(\hat{\omega}_{1,i}(\beta_{k,0});\beta_{k,0}))(k_{2,i}-k_{1,i}\hat{g}_{1,i}(\hat{\beta}_k)) \\
& = -(\hat{\beta}_k-\beta_{k,0})\left[\mathbb{E}[k_{1,i}g_{1,i}(v_{2,i}-v_{1,i}g_{1,i})]+\mathbb{E}\left[k_{2,i}(a_{2,i}-a_{1,i}g_{1,i})\right]+o_p(1)\right]+o_p(n^{-1/2}).
\end{aligned}$$

Proof of Lemma SC23. By the definition of $\hat{g}(\hat{\omega}_{1,i}(\beta_k);\beta_k)$, we can write

$$\begin{aligned}
\hat{g}(\hat{\omega}_{1,i}(\beta_k);\beta_k) &= \hat{P}_{2,i}(\beta_k)'\hat{\beta}_g(\beta_k)=\hat{P}_{2,i}(\beta_k)'(\hat{\mathbf{P}}_2(\beta_k)'\hat{\mathbf{P}}_2(\beta_k))^{-1}\hat{\mathbf{P}}_2(\beta_k)'\hat{\mathbf{Y}}_2(\beta_k) \\
&= \hat{P}^*_{2,i}(\beta_k)(\hat{\mathbf{P}}^*_2(\beta_k)'\hat{\mathbf{P}}^*_2(\beta_k))^{-1}\hat{\mathbf{P}}^*_2(\beta_k)'\hat{\mathbf{Y}}^*_2(\beta_k)
\end{aligned}$$

and therefore

$$\begin{aligned}
& \hat{g}(\hat{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k)-\hat{g}(\hat{\omega}_{1,i}(\beta_{k,0});\beta_{k,0}) \\
& = (\hat{P}^*_{2,i}(\hat{\beta}_k)-\hat{P}_{2,i}(\beta_{k,0}))'(\hat{\mathbf{P}}^*_2(\hat{\beta}_k)'\hat{\mathbf{P}}^*_2(\hat{\beta}_k))^{-1}\hat{\mathbf{P}}^*_2(\hat{\beta}_k)'\hat{\mathbf{Y}}^*_2(\hat{\beta}_k) \\
& \quad + \hat{P}_{2,i}(\beta_{k,0})'\left[(\hat{\mathbf{P}}^*_2(\hat{\beta}_k)'\hat{\mathbf{P}}^*_2(\hat{\beta}_k))^{-1}-(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}\right]\hat{\mathbf{P}}^*_2(\hat{\beta}_k)'\hat{\mathbf{Y}}^*_2(\hat{\beta}_k) \\
& \quad + \hat{P}_{2,i}(\beta_{k,0})'(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}(\hat{\mathbf{P}}^*_2(\hat{\beta}_k)-\hat{\mathbf{P}}_2(\beta_{k,0}))'\hat{\mathbf{Y}}^*_2(\beta_k) \\
& \quad + \hat{P}_{2,i}(\beta_{k,0})'(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'(\hat{\mathbf{Y}}^*_2(\hat{\beta}_k)-\mathbf{Y}^*_2(\beta_{k,0})). \qquad \text{(SC.216)}
\end{aligned}$$

The proof is divided into 4 steps. The claim of the lemma follows from the results in (SC.165), (SC.217), (SC.230), (SC.240) and (SC.242).

Step 1. In this step, we show that

$$\begin{aligned}
& n^{-1}\sum_{i=1}^{n}\upsilon'_{2,*}(\hat{P}^*_{2,i}(\hat{\beta}_k)-\hat{P}_{2,i}(\beta_{k,0}))(k_{2,i}-k_{1,i}g_{1,i}) \\
= & -(\hat{\beta}_k-\beta_{k,0})\left(\mathbb{E}[k_{1,i}g_{1,i}(k_{2,i}-k_{1,i}g_{1,i})]+o_p(1)\right)+o_p(n^{-1/2}). \qquad \text{(SC.217)}
\end{aligned}$$

where $\upsilon_{2,*}=(\hat{\mathbf{P}}^*_2(\hat{\beta}_k)'\hat{\mathbf{P}}^*_2(\hat{\beta}_k))^{-1}\hat{\mathbf{P}}^*_2(\hat{\beta}_k)'\hat{\mathbf{Y}}^*_2(\hat{\beta}_k)$.

For any $\upsilon_2\in\mathbb{R}^{m_2}$, by the second order expansion,

$$\begin{aligned}
& \upsilon'_2\left(\hat{P}^*_{2,i}(\hat{\beta}_k)-\tilde{P}_{2,i}(\beta_{k,0})-\partial^1\tilde{P}_{2,i}(\beta_{k,0})(\hat{\omega}_{1,i}(\hat{\beta}_k)-\omega_{1,i}(\beta_{k,0}))\right) \\
= & \upsilon'_2\partial^2\tilde{P}_2(\tilde{\omega}_{1,i};\beta_{k,0})(\hat{\omega}_{1,i}(\hat{\beta}_k)-\omega_{1,i}(\beta_{k,0}))^2 \qquad \text{(SC.218)}
\end{aligned}$$

where $\tilde{\omega}_{1,i}$ lies between $\hat{\omega}_{1,i}(\hat{\beta}_k)$ and $\omega_{1,i}(\beta_{k,0})$. Since $\hat{\omega}_{1,i}(\hat{\beta}_k)-\omega_{1,i}(\beta_{k,0})=(\hat{\phi}_i-\phi_i)-k_{1,i}(\hat{\beta}_k-\beta_{k,0})$, we have

$$\left| n^{-1}\sum_{i=1}^{n} \upsilon_2' \left(\begin{array}{c} \hat{P}_{2,i}^{*}(\hat{\beta}_k) - \tilde{P}_{2,i}(\beta_{k,0}) \\ -\partial^1 \tilde{P}_{2,i}(\beta_{k,0})(\hat{\omega}_{1,i}(\hat{\beta}_k) - \omega_{1,i}(\beta_{k,0})) \end{array} \right) (k_{2,i} - k_{1,i}g_{1,i}) \right|$$
$$\leq C \max_{i\leq n} \left\| \upsilon_2' \partial^2 \tilde{P}_2(\tilde{\omega}_{1,i};\beta_{k,0}) \right\| \left(n^{-1}\sum_{i=1}^{n} (\hat{\phi}_i - \phi_i)^2 + (\hat{\beta}_k - \beta_{k,0})^2 n^{-1} \sum_{i=1}^{n} k_{1,i}^2 \right). \quad \text{(SC.219)}$$

By the definition of $\upsilon_{2,*}$, we can write $\upsilon_{2,*}'\partial^2\tilde{P}_2(\tilde{\omega}_{1,i};\beta_{k,0}) = \hat{\beta}_g(\hat{\beta}_k)'\partial^2\tilde{P}_2(\tilde{\omega}_{1,i};\hat{\beta}_k)$ where $\tilde{\omega}_{1,i} \in \Omega_{\varepsilon_n}(\hat{\beta}_k)$ for any $i \leq n$ wpa1. By the triangle inequality, Assumptions SC2(iii, v, vi), Lemma SC14 and Lemma SC19

$$\begin{aligned} & \max_{i\leq n} \left\| \upsilon_{2,*}' \partial^2 \tilde{P}_2(\tilde{\omega}_{1,i};\beta_{k,0}) \right\| \\ \leq \; & \max_{i\leq n} \left\| \tilde{\beta}_{g,m_2}(\hat{\beta}_k)' \partial^2 \tilde{P}_2(\tilde{\omega}_{1,i};\hat{\beta}_k) \right\| + \max_{i\leq n} \left\| (\hat{\beta}_g(\hat{\beta}_k) - \tilde{\beta}_{g,m_2}(\hat{\beta}_k))' \partial^2 \tilde{P}_2(\tilde{\omega}_{1,i};\hat{\beta}_k) \right\| \\ = \; & O_p(1) + O_p((m_2 + m_1^{1/2})\xi_{2,m_2} n^{-1/2}). \end{aligned} \quad \text{(SC.220)}$$

By the Markov inequality, Assumptions SC1(i) and SC3(iv), Lemma SC4, Lemma SC19, (SC.219) and (SC.220), we get

$$\begin{aligned} & n^{-1}\sum_{i=1}^{n} \upsilon_{2,*}' \left(\begin{array}{c} \hat{P}_{2,i}^{*}(\hat{\beta}_k) - \tilde{P}_{2,i}(\beta_{k,0}) \\ -\partial^1 \tilde{P}_{2,i}(\beta_{k,0})(\hat{\omega}_{1,i}(\hat{\beta}_k) - \omega_{1,i}(\beta_{k,0})) \end{array} \right) (k_{2,i} - k_{1,i}g_{1,i}) \\ = \; & (\hat{\beta}_k - \beta_{k,0}) o_p(1) + o_p(n^{-1/2}). \end{aligned} \quad \text{(SC.221)}$$

Similarly, we can show that

$$n^{-1}\sum_{i=1}^{n} \upsilon_{2,*}' \left(\begin{array}{c} \hat{P}_{2,i}(\beta_{k,0}) - \tilde{P}_{2,i}(\beta_{k,0}) \\ -\partial^1 \tilde{P}_{2,i}(\beta_{k,0})(\hat{\omega}_{1,i}(\beta_{k,0}) - \omega_{1,i}(\beta_{k,0}) \end{array} \right) (k_{2,i} - k_{1,i}g_{1,i}) = o_p(n^{-1/2}). \quad \text{(SC.222)}$$

Since $\hat{\omega}_{1,i}(\hat{\beta}_k) - \hat{\omega}_{1,i}(\beta_{k,0}) = -k_{1,i}(\hat{\beta}_k - \beta_{k,0})$, using (SC.221) and (SC.222) we get

$$\begin{aligned} & n^{-1}\sum_{i=1}^{n} \upsilon_{2,*}' (\hat{P}_{2,i}^{*}(\hat{\beta}_k) - \hat{P}_{2,i}(\beta_{k,0}))(k_{2,i} - k_{1,i}g_{1,i}) \\ = \; & -(\hat{\beta}_k - \beta_{k,0}) n^{-1}\sum_{i=1}^{n} \upsilon_{2,*}' \partial^1 \tilde{P}_{2,i}(\beta_{k,0}) k_{1,i}(k_{2,i} - k_{1,i}g_{1,i}) + (\hat{\beta}_k - \beta_{k,0}) o_p(1) + o_p(n^{-1/2}). \end{aligned} \quad \text{(SC.223)}$$

By the definition of $\upsilon_{2,*}$, we can write

$$\upsilon_{2,*}' \partial^1 \tilde{P}_{2,i}(\beta_{k,0}) = \hat{\beta}_g(\hat{\beta}_k)' \partial^1 \tilde{P}_2(\omega_{1,i}(\beta_{k,0});\hat{\beta}_k).$$

Therefore

$$
\begin{aligned}
& n^{-1}\sum_{i=1}^{n}\upsilon_{2,*}'\partial^1\tilde{P}_{2,i}(\beta_{k,0})k_{1,i}(k_{2,i}-k_{1,i}g_{1,i}) \\
= \; & \mathbb{E}[k_{1,i}g_{1,i}(v_{2,i}-v_{1,i}g_{1,i})]+n^{-1}\sum_{i=1}^{n}(g_{1,i}k_{1,i}(k_{2,i}-k_{1,i}g_{1,i})-\mathbb{E}[k_{1,i}g_{1,i}(v_{2,i}-v_{1,i}g_{1,i})]) \\
& +n^{-1}\sum_{i=1}^{n}(\hat{\beta}_g(\hat{\beta}_k)'\partial^1\tilde{P}_2(\omega_{1,i}(\beta_{k,0});\hat{\beta}_k)-g_{1,i})k_{1,i}(k_{2,i}-k_{1,i}g_{1,i}). \qquad \text{(SC.224)}
\end{aligned}
$$

By Assumption SC3(iv) and Lemma SC19, $\omega_{1,i}\in\Omega_{\varepsilon_n}(\hat{\beta}_k)$ for any $i\leq n$ wpa1. Therefore by Assumption SC2(v),

$$
\max_{i\leq n}\left\|\partial^1\tilde{P}_2(\omega_{1,i}(\beta_{k,0});\hat{\beta}_k)\right\|=O_p(\xi_{1,m_2}). \qquad \text{(SC.225)}
$$

By the triangle inequality and the Cauchy-Schwarz inequality, Assumptions SC2(ii, iii, iv) and SC3(iv), Lemma SC14 and Lemma SC19, and (SC.225),

$$
\begin{aligned}
& n^{-1}\sum_{i=1}^{n}\left|\hat{\beta}_g(\hat{\beta}_k)'\partial^1\tilde{P}_2(\omega_{1,i}(\beta_{k,0});\hat{\beta}_k)-g_1(\omega_{1,i}(\beta_{k,0});\beta_{k,0})\right| \\
\leq \; & n^{-1}\sum_{i=1}^{n}\left|(\hat{\beta}_g(\hat{\beta}_k)-\tilde{\beta}_{g,m_2}(\hat{\beta}_k))'\partial^1\tilde{P}_2(\omega_{1,i}(\beta_{k,0});\hat{\beta}_k)\right| \\
& +n^{-1}\sum_{i=1}^{n}\left|\tilde{\beta}_{g,m_2}(\hat{\beta}_k)'\partial^1\tilde{P}_2(\omega_{1,i}(\beta_{k,0});\hat{\beta}_k)-g_1(\omega_{1,i}(\beta_{k,0});\hat{\beta}_k)\right| \\
& +n^{-1}\sum_{i=1}^{n}\left|g_1(\omega_{1,i}(\beta_{k,0});\hat{\beta}_k)-g_1(\omega_{1,i}(\beta_{k,0});\beta_{k,0})\right| \\
= \; & O_p(\xi_{1,m_2}(m_2+m_1^{1/2})n^{-1/2})+O_p(m^{-r_g})+(\hat{\beta}_k-\beta_{k,0})O_p(1)=o_p(1) \qquad \text{(SC.226)}
\end{aligned}
$$

which together with Assumptions SC1(i) and SC2(ii) implies that

$$
n^{-1}\sum_{i=1}^{n}(\hat{\beta}_g(\hat{\beta}_k)'\partial^1\tilde{P}_2(\omega_{1,i}(\beta_{k,0});\hat{\beta}_k)-g_{1,i})k_{1,i}(k_{2,i}-k_{1,i}g_{1,i})=o_p(1). \qquad \text{(SC.227)}
$$

By Assumptions SC1(i), SC2(ii, iii, vi) and SC3(iv), Lemma SC4 and Lemma SC9

$$
n^{-1}\sum_{i=1}^{n}(\hat{\beta}_g(\beta_{k,0})'\partial^1\tilde{P}_2(\omega_{1,i}(\beta_{k,0});\beta_{k,0})-g_{1,i})k_{1,i}(k_{2,i}-k_{1,i}g_{1,i})=o_p(1). \qquad \text{(SC.228)}
$$

By Assumptions SC1(i) and SC2(ii), and the Markov inequality,

$$
n^{-1}\sum_{i=1}^{n}(g_{1,i}k_{1,i}(k_{2,i}-k_{1,i}g_{1,i})-\mathbb{E}[k_{1,i}g_{1,i}(v_{2,i}-v_{1,i}g_{1,i})])=O_p(n^{-1/2})
$$

which together with (SC.224), (SC.227) and (SC.228) implies that

$$n^{-1}\sum_{i=1}^{n}\upsilon_{2,*}'\partial^{1}\tilde{P}_{2,i}(\beta_{k,0})k_{1,i}(k_{2,i}-k_{1,i}g_{1,i})=\mathbb{E}[k_{1,i}g_{1,i}(v_{2,i}-v_{1,i}g_{1,i})]+o_p(1). \tag{SC.229}$$

The claim in (SC.217) follows from (SC.223) and (SC.229).

Step 2. In this step, we show that

$$\begin{aligned}&\upsilon_{2,*}'\left[\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0})-\hat{\mathbf{P}}_2^{*}(\hat{\beta}_k)'\hat{\mathbf{P}}_2^{*}(\hat{\beta}_k)\right]\hat{\beta}_{\varphi}(\beta_{k,0})\\&=(\hat{\beta}_k-\beta_{k,0})\left(\mathbb{E}[g_{1,i}k_{1,i}(a_{2,i}-a_{1,i}g_{1,i})]+o_p(1)\right)\\&\quad+n^{-1}\upsilon_{2,*}'\hat{\mathbf{P}}_2^{*}(\hat{\beta}_k)'(\hat{\mathbf{P}}_2(\beta_{k,0})-\hat{\mathbf{P}}_2^{*}(\hat{\beta}_k))\hat{\beta}_{\varphi}(\beta_{k,0})+o_p(n^{-1/2})\end{aligned} \tag{SC.230}$$

where $\hat{\beta}_{\varphi}(\beta_{k,0})=(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_{k,0})(k_{2,i}-k_{1,i}g_{1,i})$.

Since we can write

$$\begin{aligned}&\upsilon_{2,*}'\left[\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0})-\hat{\mathbf{P}}_2^{*}(\hat{\beta}_k)'\hat{\mathbf{P}}_2^{*}(\hat{\beta}_k)\right]\hat{\beta}_{\varphi}(\beta_{k,0})\\&=\upsilon_{2,*}'(\hat{\mathbf{P}}_2(\beta_{k,0})-\hat{\mathbf{P}}_2^{*}(\hat{\beta}_k))'\hat{\mathbf{P}}_2(\beta_{k,0})\hat{\beta}_{\varphi}(\beta_{k,0})+\upsilon_{2,*}'\hat{\mathbf{P}}_2^{*}(\hat{\beta}_k)'(\hat{\mathbf{P}}_2(\beta_{k,0})-\hat{\mathbf{P}}_2^{*}(\hat{\beta}_k))\hat{\beta}_{\varphi}(\beta_{k,0}),\end{aligned}$$

to prove (SC.230) it is sufficient to show that

$$n^{-1}\upsilon_{2,*}'(\hat{\mathbf{P}}_2(\beta_{k,0})-\hat{\mathbf{P}}_2^{*}(\hat{\beta}_k))'\hat{\mathbf{P}}_2(\beta_{k,0})\hat{\beta}_{\varphi}(\beta_{k,0})=(\hat{\beta}_k-\beta_{k,0})\left(\mathbb{E}\left[g_{1,i}k_{1,i}\varphi_i\right]+o_p(1)\right)+o_p(n^{-1/2}). \tag{SC.231}$$

where $\varphi_i=a_{2,i}-a_{1,i}g_{1,i}$.

By the Cauchy-Schwarz inequality, Assumption SC3(iv), (SC.58), (SC.117) and (SC.191),

$$\begin{aligned}&n^{-1}\left|\upsilon_{2,*}'(\hat{\mathbf{P}}_2(\beta_{k,0})-\hat{\mathbf{P}}_2^{*}(\hat{\beta}_k))'\hat{\mathbf{P}}_2(\beta_{k,0})(\hat{\beta}_{\varphi}(\beta_{k,0})-\tilde{\beta}_{\varphi,m_2}(\beta_{k,0}))\right|\\\leq\;&n^{-1}\left\|\upsilon_{2,*}'(\hat{\mathbf{P}}_2(\beta_{k,0})-\hat{\mathbf{P}}_2^{*}(\hat{\beta}_k))\right\|\left\|\hat{\mathbf{P}}_2(\beta_{k,0})(\hat{\beta}_{\varphi}(\beta_{k,0})-\tilde{\beta}_{\varphi,m_2}(\beta_{k,0}))\right\|\\\leq\;&\frac{\left\|\upsilon_{2,*}'(\hat{\mathbf{P}}_2(\beta_{k,0})-\hat{\mathbf{P}}_2^{*}(\hat{\beta}_k))\right\|\left\|\hat{\beta}_{\varphi}(\beta_{k,0})-\tilde{\beta}_{\varphi,m_2}(\beta_{k,0})\right\|}{(\lambda_{\max}(n^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0})))^{-1/2}n^{1/2}}\\=\;&|\hat{\beta}_k-\beta_{k,0}|O_p((m_1^{1/2}+m_2^{1/2})n^{-1/2})=|\hat{\beta}_k-\beta_{k,0}|o_p(1).\end{aligned} \tag{SC.232}$$

By the Cauchy-Schwarz inequality, Assumptions SC3(i, ii, iv), Lemma SC4, (SC.53) and (SC.191),

$$\begin{aligned}&n^{-1}\left|\upsilon_{2,*}'(\hat{\mathbf{P}}_2(\beta_{k,0})-\hat{\mathbf{P}}_2^{*}(\hat{\beta}_k))'(\hat{\mathbf{P}}_2(\beta_{k,0})-\tilde{\mathbf{P}}_2(\beta_{k,0}))'\tilde{\beta}_{\varphi,m_2}(\beta_{k,0})\right|\\&\leq n^{-1}\left\|(\hat{\mathbf{P}}_2(\beta_{k,0})-\hat{\mathbf{P}}_2^{*}(\hat{\beta}_k))\upsilon_{2,*}\right\|\left\|(\hat{\mathbf{P}}_2(\beta_{k,0})-\tilde{\mathbf{P}}_2(\beta_{k,0}))\tilde{\beta}_{\varphi,m_2}(\beta_{k,0})\right\|_S\\&=|\hat{\beta}_k-\beta_{k,0}|O_p(m_1^{1/2}n^{-1/2})=|\hat{\beta}_k-\beta_{k,0}|o_p(1).\end{aligned} \tag{SC.233}$$

By (SC.191) and Assumption SC3(ii, iv),

$$\begin{aligned} & n^{-1}\left|\upsilon_{2,*}'(\hat{\mathbf{P}}_2(\beta_{k,0})-\hat{\mathbf{P}}_2(\hat{\beta}_k))'(\tilde{\mathbf{P}}_2(\beta_{k,0})'\tilde{\beta}_{\varphi,m_2}(\beta_{k,0})-\boldsymbol{\varphi}_n)\right| \\ & \quad \le n^{-1}\left\|(\hat{\mathbf{P}}_2(\beta_{k,0})-\hat{\mathbf{P}}_2^*(\hat{\beta}_k))\upsilon_{2,*}\right\|_S \left\|\tilde{\mathbf{P}}_2(\beta_{k,0})'\tilde{\beta}_{\varphi,m_2}(\beta_{k,0})-\boldsymbol{\varphi}_n\right\| \\ & \quad = \left|\hat{\beta}_k-\beta_{k,0}\right| O_p(n^{-1/2}) = |\hat{\beta}_k-\beta_{k,0}|o_p(1) \end{aligned}$$

where $\boldsymbol{\varphi}_n = (\varphi_1, \ldots, \varphi_n)'$, which together with (SC.232) and (SC.233) implies that

$$\begin{aligned} & n^{-1}\upsilon_{2,*}'(\hat{\mathbf{P}}_2(\beta_{k,0})-\hat{\mathbf{P}}_2^*(\hat{\beta}_k))'\hat{\mathbf{P}}_2(\beta_{k,0})\hat{\beta}_\varphi(\beta_{k,0}) \\ = \; & n^{-1}\upsilon_{2,*}'(\hat{\mathbf{P}}_2(\beta_{k,0})-\hat{\mathbf{P}}_2^*(\hat{\beta}_k))'\boldsymbol{\varphi}_n+(\hat{\beta}_k-\beta_{k,0})o_p(1). \end{aligned} \quad \text{(SC.234)}$$

Since $\upsilon_{2,*} = (B(\beta_{k,0})')^{-1}B(\hat{\beta}_k)'\hat{\beta}_g(\hat{\beta}_k)$, we can write

$$\upsilon_{2,*}'(\hat{\mathbf{P}}_2(\beta_{k,0})-\hat{\mathbf{P}}_2^*(\hat{\beta}_k))'\boldsymbol{\varphi}_n = \sum_{i=1}^n \hat{\beta}_g(\hat{\beta}_k)'(\tilde{P}_2(\hat{\omega}_{1,i}(\beta_{k,0});\hat{\beta}_k)-\tilde{P}_2(\hat{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k))\varphi_i. \quad \text{(SC.235)}$$

By the first-order expansion, the triangle inequality and the Cauchy-Schwarz inequality, Assumptions SC1(i) and SC3(i, iv), Lemma SC14 and Lemma SC19, we have

$$\begin{aligned} & n^{-1}\sum_{i=1}^n(\hat{\beta}_g(\hat{\beta}_k)-\tilde{\beta}_{g,m_2}(\hat{\beta}_k))'(\tilde{P}_2(\hat{\omega}_{1,i}(\beta_{k,0});\hat{\beta}_k)-\tilde{P}_2(\hat{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k))\varphi_i \\ = \; & (\hat{\beta}_k-\beta_{k,0})n^{-1}\sum_{i=1}^n(\hat{\beta}_g(\hat{\beta}_k)-\tilde{\beta}_{g,m_2}(\hat{\beta}_k))'\partial^1\tilde{P}_2(\tilde{\omega}_{1,i};\hat{\beta}_k)k_{1,i}\varphi_i \\ = \; & (\hat{\beta}_k-\beta_{k,0})O_p((m_1^{1/2}+m_2)n^{-1/2})O_p(\xi_{1,m_2}) = (\hat{\beta}_k-\beta_{k,0})o_p(1). \end{aligned} \quad \text{(SC.236)}$$

By Assumptions SC1(i), SC2(iii) and SC3(i),

$$n^{-1}\sum_{i=1}^n(\tilde{\beta}_{g,m_2}(\hat{\beta}_k)'\tilde{P}_2(\hat{\omega}_{1,i}(\beta_{k,0});\hat{\beta}_k)-g(\hat{\omega}_{1,i}(\beta_{k,0});\hat{\beta}_k))k_{1,i}\varphi_i = o_p(n^{-1/2})$$

and

$$n^{-1}\sum_{i=1}^n(\tilde{\beta}_{g,m_2}(\hat{\beta}_k)'\tilde{P}_2(\hat{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k)-g(\hat{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k))k_{1,i}\varphi_i = o_p(n^{-1/2})$$

which together with (SC.235) and (SC.236) implies that

$$\begin{aligned} & n^{-1}\upsilon_{2,*}'(\hat{\mathbf{P}}_2(\beta_{k,0})-\hat{\mathbf{P}}_2^*(\hat{\beta}_k))'\boldsymbol{\varphi}_n \\ = \; & n^{-1}\sum_{i=1}^n(g(\hat{\omega}_{1,i}(\beta_{k,0});\hat{\beta}_k)-g(\hat{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k))k_{1,i}\varphi_i+(\hat{\beta}_k-\beta_{k,0})o_p(1)+o_p(n^{-1/2}). \end{aligned} \quad \text{(SC.237)}$$

By Assumptions SC1(i), SC2(ii) and SC3(i), Lemma SC4 and Lemma SC19

$$n^{-1}\sum_{i=1}^{n} k_{1,i}\varphi_i(g_1(\hat{\omega}_{1,i}(\beta_{k,0});\hat{\beta}_k) - g_1(\omega_{1,i}(\beta_{k,0});\beta_{k,0})) = o_p(1). \tag{SC.238}$$

By Assumptions SC1(i), SC2(ii) and SC3(i), and Lemma SC19

$$n^{-1}\sum_{i=1}^{n} k_{1,i}\varphi_i(g(\hat{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k) - g(\hat{\omega}_{1,i}(\beta_{k,0});\hat{\beta}_k) + g_1(\hat{\omega}_{1,i}(\beta_{k,0});\hat{\beta}_k)(\hat{\beta}_k - \beta_{k,0})) = (\hat{\beta}_k - \beta_{k,0})o_p(1)$$

which together with (SC.238) implies that

$$\begin{aligned} & n^{-1}\sum_{i=1}^{n}(g(\hat{\omega}_{1,i}(\beta_{k,0});\hat{\beta}_k) - g(\hat{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k))k_{1,i}\varphi_i \\ = \;& (\hat{\beta}_k - \beta_{k,0})n^{-1}\sum_{i=1}^{n} g_{1,i}k_{1,i}\varphi_i + (\hat{\beta}_k - \beta_{k,0})o_p(1) \\ = \;& (\hat{\beta}_k - \beta_{k,0})\mathbb{E}[g_{1,i}k_{1,i}\varphi_i] + (\hat{\beta}_k - \beta_{k,0})o_p(1) + o_p(n^{-1/2}) \end{aligned} \tag{SC.239}$$

where the second equality is by the Markov inequality. The claim in (SC.231) now follows from (SC.234), (SC.237) and (SC.239).

Step 3. In this step, we show that

$$n^{-1}(\hat{\mathbf{Y}}_2^*(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\hat{\beta}_k)\hat{\beta}_g(\hat{\beta}_k))'(\hat{\mathbf{P}}_2^*(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\beta_{k,0}))\hat{\beta}_\varphi(\beta_{k,0}) = (\hat{\beta}_k - \beta_{k,0})o_p(1) + o_p(n^{-1/2}). \tag{SC.240}$$

By definition $\hat{y}_2^*(\hat{\beta}_k) = \hat{y}_2^* - k_{2,i}\hat{\beta}_k$, we can write

$$\begin{aligned} \hat{y}_2^*(\hat{\beta}_k) - \hat{P}_2(\hat{\beta}_k)'\hat{\beta}_g(\hat{\beta}_k) &= y_2^* - l_{2,i}\hat{\beta}_l - k_{2,i}\hat{\beta}_k - \hat{g}(\hat{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k) \\ &= u_{2,i} - l_{2,i}(\hat{\beta}_l - \beta_{l,o}) - k_{2,i}(\hat{\beta}_k - \beta_{k,o}) - (\hat{g}(\hat{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k) - g(\omega_{1,i})). \end{aligned}$$

Therefore,

$$\begin{aligned} & n^{-1}(\hat{\mathbf{Y}}_2^*(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\hat{\beta}_k)\hat{\beta}_g(\hat{\beta}_k))'(\hat{\mathbf{P}}_2^*(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\beta_{k,0}))\hat{\beta}_\varphi(\beta_{k,0}) \\ = \;& n^{-1}\mathbf{U}_2'(\hat{\mathbf{P}}_2^*(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\beta_{k,0}))\hat{\beta}_\varphi(\beta_{k,0}) \\ & -(\hat{\beta}_l - \beta_{l,o})n^{-1}\mathbf{L}_2'(\hat{\mathbf{P}}_2^*(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\beta_{k,0}))\hat{\beta}_\varphi(\beta_{k,0}) \\ & -n^{-1}(\hat{\beta}_k - \beta_{l,o})\mathbf{K}_2'(\hat{\mathbf{P}}_2^*(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\beta_{k,0}))\hat{\beta}_\varphi(\beta_{k,0}) \\ & -n^{-1}(\hat{\mathbf{G}}_2 - \mathbf{G}_2)'(\hat{\mathbf{P}}_2^*(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\beta_{k,0}))\hat{\beta}_\varphi(\beta_{k,0}) \end{aligned} \tag{SC.241}$$

which combined with Lemma SC22 proves (SC.240).

Step 4. In this step, we show that

$$n^{-1}(\hat{\mathbf{Y}}_2^*(\hat{\beta}_k)-\hat{\mathbf{Y}}_2^*(\beta_{k,0}))'\hat{\mathbf{P}}_2(\beta_{k,0})\hat{\beta}_\varphi(\beta_{k,0}) = -(\hat{\beta}_k-\beta_{k,0})\left(\mathbb{E}[k_{2,i}(a_{2,i}-a_{1,i}g_{1,i})]+o_p(1)\right). \quad \text{(SC.242)}$$

Since $\hat{y}_2^*(\hat{\beta}_k) - \hat{y}_2^*(\beta_{k,0}) = -k_{2,i}(\hat{\beta}_k - \beta_{k,0})$, we have

$$n^{-1}(\hat{\mathbf{Y}}_2^*(\hat{\beta}_k) - \hat{\mathbf{Y}}_2^*(\beta_{k,0}))'\hat{\mathbf{P}}_2(\beta_{k,0})\hat{\beta}_\varphi(\beta_{k,0}) = -n^{-1}(\hat{\beta}_k - \beta_{k,0})'\mathbf{K}_2'\hat{\mathbf{P}}_2(\beta_{k,0})\hat{\beta}_\varphi(\beta_{k,0}) \quad \text{(SC.243)}$$

and

$$\begin{aligned} n^{-1}\mathbf{K}_2'\hat{\mathbf{P}}_2(\beta_{k,0})\hat{\beta}_\varphi(\beta_{k,0}) &= \mathbb{E}[k_{2,i}\varphi(\omega_{1,i})] + n^{-1}\sum_{i=1}^n(k_{2,i}\varphi(\omega_{1,i}) - \mathbb{E}[k_{2,i}\varphi(\omega_{1,i})]) \\ &+n^{-1}\sum_{i=1}^n k_{2,i}(\varphi(\omega_{1,i}) - \varphi(\hat{\omega}_{1,i}(\beta_{k,0}))) \\ &+n^{-1}\sum_{i=1}^n k_{2,i}(\varphi(\hat{\omega}_{1,i}(\beta_{k,0})) - \hat{P}_{2,i}(\beta_{k,0})'\tilde{\beta}_{\varphi,m_2}(\beta_{k,0})) \\ &+n^{-1}\sum_{i=1}^n k_{2,i}\hat{P}_{2,i}(\beta_{k,0})'(\hat{\beta}_\varphi(\beta_{k,0}) - \tilde{\beta}_{\varphi,m_2}(\beta_{k,0})). \end{aligned} \quad \text{(SC.244)}$$

By Assumptions SC1(i) and SC3(i, ii), and the Markov inequality and Lemma SC4, we have

$$n^{-1}\sum_{i=1}^n(k_{2,i}\varphi(\omega_{1,i}) - \mathbb{E}[k_{2,i}\varphi(\omega_{1,i})]) = o_p(1) \quad \text{(SC.245)}$$

and

$$n^{-1}\sum_{i=1}^n k_{2,i}(\varphi(\hat{\omega}_{1,i}(\beta_{k,0})) - \hat{P}_{2,i}(\beta_{k,0})'\tilde{\beta}_{\varphi,m_2}(\beta_{k,0})) = o_p(1) \quad \text{(SC.246)}$$

and

$$n^{-1}\sum_{i=1}^n k_{2,i}(\varphi(\omega_{1,i}) - \varphi(\hat{\omega}_{1,i}(\beta_{k,0}))) = o_p(1). \quad \text{(SC.247)}$$

By the Cauchy-Schwarz inequality, (SC.58) and (SC.115)

$$\begin{aligned} &\left|n^{-1}\sum_{i=1}^n k_{2,i}\hat{P}_{2,i}(\beta_{k,0})'(\hat{\beta}_\varphi(\beta_{k,0}) - \tilde{\beta}_{\varphi,m_2}(\beta_{k,0}))\right| \\ \leq\ & C\lambda_{\max}(n^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))\left\|\hat{\beta}_\varphi(\beta_{k,0}) - \tilde{\beta}_{\varphi,m_2}(\beta_{k,0})\right\| = o_p(1). \end{aligned} \quad \text{(SC.248)}$$

The claim in (SC.242) follows from (SC.243), (SC.244), (SC.245), (SC.246) and (SC.247). *Q.E.D.*

SC.4 Auxiliary results for the standard error estimation

Assumption SC4. *(i) There exists $\hat{\varepsilon}_{1,i}$ for $i = 1, \ldots, n$ such that $n^{-1}\sum_{i=1}^{n}(\hat{\varepsilon}_{1,i} - \varepsilon_{1,i})^4 = o_p(1)$; (ii) there exist $r_h > 1$ and $\beta_{h,m} \in \mathbb{R}^m$ such that $\sup_{x\in\mathcal{X}} |h_m(x) - h(x)| = O(m^{-r_h})$ where $h_m(x) \equiv P_1(x)' \beta_{h,m}$ and $\xi_{0,m_1} m^{-r_h} = o(1)$; (iii) $\Omega > 0$.*

The following lemma is useful to show the consistency of the estimator of the asymptotic variance.

Lemma SC24. *Under Assumptions SC1, SC2 and SC3, we have*
(i) $n^{-1}\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k) - n^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}) = O_p(\xi_{1,m_2}n^{-1/2})$;
(ii) $\max_{i\leq n} |\hat{g}_{1,i} - g_{1,i}| = O_p(\xi_{1,m_2}(m_2 + m_1^{1/2})n^{-1/2})$;
(iii) $n^{-1}\sum_{i=1}^{n}(\hat{\varsigma}_i - v_{2,i} + v_{1,i}g_{1,i})^4 = O_p\left((m_2^4 + m_1^2)\xi_{1,m_2}^4\xi_{0,m_2}^2 n^{-2}\right)$;
(iv) $n^{-1}\sum_{i=1}^{n}(\hat{u}_{2,i} - u_{2,i})^4 = O_p((m_1^2 + m_2^4)\xi_{0,m_2}^2 n^{-2})$;
(v) $\max_{i\leq n} |\hat{h}_i - h_i| = o_p(1)$.

Proof of Lemma SC24. (i) For any $\upsilon_2 \in \mathbb{R}^{m_2}$, by Assumption SC2(v) and (SB.12) in Theorem SB1

$$\begin{aligned} & \upsilon_2'(\hat{\mathbf{P}}_2(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\beta_{k,0}))'(\hat{\mathbf{P}}_2(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\beta_{k,0}))\upsilon_2 \\ = & \sum_{i=1}^{n}(\upsilon_2'((\hat{P}_{2,i}(\hat{\beta}_k) - \hat{P}_{2,i}(\beta_{k,0})))^2 \\ = & (\hat{\beta}_k - \beta_{k,0})^2\sum_{i=1}^{n}(\upsilon_2'\partial\tilde{P}_2(\hat{\omega}_{1,i}(\tilde{\beta}_k);\tilde{\beta}_k)/\partial\beta_k)^2 = \|\upsilon_2\|^2 O_p(\xi_{1,m_2}^2 n^{-1}) \end{aligned}$$

which implies that

$$\left\|\hat{\mathbf{P}}_2(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\beta_{k,0})\right\|_S = O_p(\xi_{1,m_2}n^{-1/2}). \quad \text{(SC.249)}$$

By the triangle inequality and the Cauchy-Schwarz inequality, Assumption SC3(iv), (SC.58) and (SC.249)

$$\begin{aligned} & n^{-1}\left\|\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0})\right\|_S \\ \leq & n^{-1}\left\|(\hat{\mathbf{P}}_2(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\beta_{k,0}))'\hat{\mathbf{P}}_2(\beta_{k,0})\right\|_S \\ & + n^{-1}\left\|\hat{\mathbf{P}}_2(\beta_{k,0})'(\hat{\mathbf{P}}_2(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\beta_{k,0}))\right\|_S \\ & + n^{-1}\left\|(\hat{\mathbf{P}}_2(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\beta_{k,0}))'(\hat{\mathbf{P}}_2(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\beta_{k,0}))\right\|_S = O_p(\xi_{1,m_2}n^{-1/2}) \quad \text{(SC.250)} \end{aligned}$$

which proves the claim.

(ii) Using the similar arguments in deriving (SC.157), we can show that

$$\begin{aligned}
&\max_{i\leq n}\left|\partial^1\hat{P}_{2,i}(\hat{\beta}_k)'\hat{\beta}_g(\hat{\beta}_k)-g_1(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k)\right| \\
\leq\ &\max_{i\leq n}\left|\partial^1\hat{P}_{2,i}(\hat{\beta}_k)'\hat{\beta}_g(\hat{\beta}_k)-g_{1,m_2}(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k)\right| \\
&+\max_{i\leq n}\left|g_{1,m_2}(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k)-g_1(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k)\right| \\
\leq\ &2\xi_{1,m_2}||\hat{\beta}_g(\hat{\beta}_k)-\tilde{\beta}_{g,m_2}(\hat{\beta}_k)||+Cm_2^{-r_g}
\end{aligned}$$

which together with (SB.12) and Lemma SC14 implies that

$$\max_{i\leq n}\left|\partial^1\hat{P}_{2,i}(\hat{\beta}_k)'\hat{\beta}_g(\hat{\beta}_k)-g_1(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k)\right|=O_p(\xi_{1,m_2}(m_2+m_1^{1/2})n^{-1/2}). \tag{SC.251}$$

Using similar arguments in showing (SC.160), we get

$$\begin{aligned}
&\max_{i\leq n}\left|(\partial^1\hat{P}_{2,i}(\hat{\beta}_k)-\partial^1\tilde{P}_2(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k))'\hat{\beta}_g(\hat{\beta}_k)\right| \\
\leq\ &\max_{i\leq n}\left|(\partial^1\hat{P}_{2,i}(\hat{\beta}_k)-\partial^1\tilde{P}_2(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k))'\tilde{\beta}_{g,m_2}(\hat{\beta}_k)\right| \\
&+\max_{i\leq n}\left|(\partial^1\hat{P}_{2,i}(\hat{\beta}_k)-\partial^1\tilde{P}_2(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k))'(\hat{\beta}_g(\hat{\beta}_k)-\tilde{\beta}_{g,m_2}(\hat{\beta}_k))\right| \\
\leq\ &\max_{i\leq n}\left|g_1(\hat{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k)-g_1(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k)\right|+O_p(\xi_{2,m_2}\xi_{1,m_1}m_1^{1/2}(m_2+m_1^{1/2})n^{-1}) \\
=\ &O_p(\xi_{1,m_2}(m_2+m_1^{1/2})n^{-1/2}).
\end{aligned} \tag{SC.252}$$

By Assumption SC2(ii) and (SB.12), we have

$$\max_{i\leq n}\left|g_1(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k)-g_1(\omega_{1,i}(\beta_{k,0});\beta_{k,0})\right|=O_p(n^{-1/2})$$

which together with (SC.251) and (SC.252) proves the second claim of the lemma.

(iii) Define $\hat{\varphi}_i=\hat{P}_{2,i}(\hat{\beta}_k)'\hat{\beta}_\varphi(\hat{\beta}_k)$ for $i\leq n$, where

$$\hat{\beta}_\varphi(\hat{\beta}_k)=(\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k))^{-1}\sum_{i=1}^n\hat{P}_{2,i}(\hat{\beta}_k)(k_{2,i}-k_{1,i}\hat{g}_{1,i}(\hat{\beta}_k)).$$

Let $\Delta k_{2,i}=k_{2,i}-k_{1,i}g_{1,i}$ and $\Delta\hat{k}_{2,i}=k_{2,i}-k_{1,i}\hat{g}_{1,i}(\hat{\beta}_k)$. Since $v_{2,i}-v_{1,i}g_{1,i}=\Delta k_{2,i}-\varphi_i$ and $\hat{\varsigma}_i=\Delta\hat{k}_{2,i}-\hat{\varphi}_i$, we have

$$n^{-1}\sum_{i=1}^n(\hat{\varsigma}_i-v_{2,i}+v_{1,i}g_{1,i})^4\leq Cn^{-1}\sum_{i=1}^n(\Delta\hat{k}_{2,i}-\Delta k_{2,i})^4+Cn^{-1}\sum_{i=1}^n(\hat{\varphi}_i-\varphi_i)^4. \tag{SC.253}$$

By Assumption SC3(ii), Lemma SC16, Lemma SC24(ii) and (SB.12),

$$\begin{aligned} n^{-1}\sum_{i=1}^{n}(\Delta\hat{k}_{2,i}-\Delta k_{2,i})^4 &= n^{-1}\sum_{i=1}^{n}k_{1,i}^4(\hat{g}_{1,i}(\hat{\beta}_k)-g_1(\omega_{1,i}(\beta_{k,0});\beta_{k,0}))^4 \\ &\leq Cn^{-1}\sum_{i=1}^{n}(\hat{g}_{1,i}(\hat{\beta}_k)-g_1(\omega_{1,i}(\beta_{k,0});\beta_{k,0}))^4 \\ &= O_p((m_2^4+m_1^2)\xi_{1,m_2}^4 n^{-2}). \end{aligned} \quad \text{(SC.254)}$$

Similarly we can show that

$$n^{-1}\sum_{i=1}^{n}(\Delta\hat{k}_{2,i}-\Delta k_{2,i})^2 = O_p(\xi_{1,m_2}^2(m_2^2+m_1)n^{-1}). \quad \text{(SC.255)}$$

By the definition of $\hat{\varphi}_i$, we can write

$$\begin{aligned} \hat{\varphi}_i-\varphi_i &= \hat{P}_{2,i}(\hat{\beta}_k)'(\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k))^{-1}\hat{\mathbf{P}}_2(\hat{\beta}_k)(\Delta\hat{\mathbf{K}}_2-\Delta\mathbf{K}_2) \\ &\quad +\hat{P}_{2,i}(\hat{\beta}_k)'(\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k))^{-1}(\hat{\mathbf{P}}_2(\hat{\beta}_k)-\hat{\mathbf{P}}_2(\beta_{k,0}))\Delta\mathbf{K}_2 \\ &\quad +\hat{P}_{2,i}(\hat{\beta}_k)'[(\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k))^{-1}-(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}]\hat{\mathbf{P}}_2(\beta_{k,0})\Delta\mathbf{K}_2 \\ &\quad +(\hat{P}_{2,i}(\hat{\beta}_k)-\hat{P}_{2,i}(\beta_{k,0}))'(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})\Delta\mathbf{K}_2 \\ &\quad +\hat{P}_{2,i}(\beta_{k,0})'(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})\Delta\mathbf{K}_2-\varphi_i. \end{aligned} \quad \text{(SC.256)}$$

where $\Delta\hat{\mathbf{K}}_2=(\Delta\hat{k}_{2,1},\ldots,\Delta\hat{k}_{2,n})'$ and $\Delta\mathbf{K}_2=(\Delta k_{2,1},\ldots,\Delta k_{2,n})'$. By Assumption SC2(v), (SC.58) and (SC.255),

$$\begin{aligned} & n^{-1}\sum_{i=1}^{n}(\hat{P}_{2,i}(\hat{\beta}_k)'(\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k))^{-1}\hat{\mathbf{P}}_2(\hat{\beta}_k)(\Delta\hat{\mathbf{K}}_2-\Delta\mathbf{K}_2))^4 \\ \leq\ & \xi_{0,m_2}^2\left\|(\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k))^{-1}\hat{\mathbf{P}}_2(\hat{\beta}_k)(\Delta\hat{\mathbf{K}}_2-\Delta\mathbf{K}_2)\right\|^2 \\ & \times n^{-1}\sum_{i=1}^{n}(\hat{P}_{2,i}(\hat{\beta}_k)'(\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k))^{-1}\hat{\mathbf{P}}_2(\hat{\beta}_k)(\Delta\hat{\mathbf{K}}_2-\Delta\mathbf{K}_2))^2 \\ \leq\ & (\lambda_{\min}(n^{-1}\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k)))^{-1}\xi_{0,m_2}^2\left(n^{-1}\sum_{i=1}^{n}(\Delta\hat{k}_{2,i}-\Delta k_{2,i})^2\right)^2 \\ =\ & O_p\left((m_2^4+m_1^2)\xi_{1,m_2}^4\xi_{0,m_2}^2n^{-2}\right). \end{aligned} \quad \text{(SC.257)}$$

By the first order expansion, the triangle inequality and the Cauchy-Schwarz inequality,

$$\begin{aligned}
& n^{-1}\sum_{i=1}^{n}\left|\hat{P}_{2,i}(\hat{\beta}_k)'(\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k))^{-1}(\hat{\mathbf{P}}_2(\hat{\beta}_k)-\hat{\mathbf{P}}_2(\beta_{k,0}))\Delta\mathbf{K}_2\right|^4 \\
= \; & (\hat{\beta}_k-\beta_{k,0})^4 n^{-1}\sum_{i=1}^{n}\left|\hat{P}_{2,i}(\hat{\beta}_k)'(\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k))^{-1}\sum_{i=1}^{n}\partial^1\tilde{P}_2(\tilde{\omega}_{1,i};\beta_{k,0})k_{1,i}\Delta k_{2,i}\right|^4 \\
\leq \; & (\hat{\beta}_k-\beta_{k,0})^4\xi_{0,m_2}^2\left\|(\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k))^{-1}\sum_{i=1}^{n}\partial^1\tilde{P}_2(\tilde{\omega}_{1,i};\beta_{k,0})k_{1,i}\Delta k_{2,i}\right\|^2 \\
& \times n^{-1}\sum_{i=1}^{n}\left|\hat{P}_{2,i}(\hat{\beta}_k)'(\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k))^{-1}\sum_{i=1}^{n}\partial^1\tilde{P}_2(\tilde{\omega}_{1,i};\beta_{k,0})k_{1,i}\Delta k_{2,i}\right|^2 \\
\leq \; & \frac{(\hat{\beta}_k-\beta_{k,0})^4\xi_{0,m_2}^2\xi_{1,m_2}^4}{(\lambda_{\min}(n^{-1}\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k)))^3}\left|n^{-1}\sum_{i=1}^{n}k_{1,i}^2(\Delta k_{2,i})^2\right|^2
\end{aligned}$$

where $\tilde{\omega}_{1,i}$ lies between $\hat{\omega}_{1,i}(\hat{\beta}_k)$ and $\hat{\omega}_{1,i}(\beta_{k,0})$, which together with Assumptions SC1(i) and SC2(ii), (SB.12) and (SC.58) implies that

$$n^{-1}\sum_{i=1}^{n}\left|\hat{P}_{2,i}(\hat{\beta}_k)'(\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k))^{-1}(\hat{\mathbf{P}}_2(\hat{\beta}_k)-\hat{\mathbf{P}}_2(\beta_{k,0}))\Delta\mathbf{K}_2\right|^4 = O_p(\xi_{1,m_2}^4\xi_{0,m_2}^2 n^{-2}). \quad \text{(SC.258)}$$

By Assumptions SC2(iv, vi) and SC3(i, ii) and (SC.117),

$$\begin{aligned}
\left\|\hat{\beta}_{\varphi}(\beta_{k,0})\right\| \leq \; & \left\|\tilde{\beta}_{\varphi,m_2}(\beta_{k,0})\right\| + \left\|\hat{\beta}_{\varphi}(\beta_{k,0})-\tilde{\beta}_{\varphi,m_2}(\beta_{k,0})\right\| \\
\leq \; & (\lambda_{\min}(Q_{m_2}(\beta_{k,0})))^{-1/2}\mathbb{E}\left\|\varphi_{m_2}\right\|_2 + O_p((m_1^{1/2}+m_2^{1/2})n^{-1/2}) \\
\leq \; & C\mathbb{E}\left\|\varphi_{m_2}\right\|_2 + C\mathbb{E}\left\|\varphi_{m_2}-\varphi(\omega_{1,i})\right\|_2 + o_p(1) = O_p(1). \qquad \text{(SC.259)}
\end{aligned}$$

By the Cauchy-Schwarz inequality, Lemma SC24(i), (SC.70) and (SC.259),

$$\begin{aligned}
& n^{-1}\sum_{i=1}^{n}\left|\hat{P}_{2,i}(\hat{\beta}_k)'[(\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k))^{-1}-(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}]\hat{\mathbf{P}}_2(\beta_{k,0})\Delta\mathbf{K}_2\right|^4 \\
= \; & n^{-1}\sum_{i=1}^{n}\left|\hat{P}_{2,i}(\hat{\beta}_k)'(\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k))^{-1}[\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k)-\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0})]\hat{\beta}_{\varphi}(\beta_{k,0})\right|^4 \\
\leq \; & \xi_{0,m_2}^2\left\|(\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k))^{-1}[\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k)-\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0})]\hat{\beta}_{\varphi}(\beta_{k,0})\right\|^2 \\
& \times n^{-1}\sum_{i=1}^{n}\left|\hat{P}_{2,i}(\hat{\beta}_k)'(\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k))^{-1}[\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k)-\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0})]\hat{\beta}_{\varphi}(\beta_{k,0})\right|^2 \\
\leq \; & \frac{\xi_{0,m_2}^2\left\|\hat{\beta}_{\varphi}(\beta_{k,0})\right\|^4}{(\lambda_{\min}(n^{-1}\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k)))^3}\left\|n^{-1}\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k)-n^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0})\right\|_S^4 \\
= \; & O_p(\xi_{0,m_2}^2\xi_{1,m_2}^4 n^{-2}). \qquad \text{(SC.260)}
\end{aligned}$$

By the first order expansion, (SB.12) in Theorem SB1, Assumption SC3(ii) and (SC.259),

$$\begin{aligned}
& n^{-1}\sum_{i=1}^{n}((\hat{P}_{2,i}(\hat{\beta}_k)-\hat{P}_{2,i}(\beta_{k,0}))'\hat{\beta}_{\varphi}(\beta_{k,0}))^4 \\
= \; & (\hat{\beta}_k-\beta_{k,0})^4 n^{-1}\sum_{i=1}^{n}(\partial\tilde{P}_2(\hat{\omega}_{1,i}(\tilde{\beta}_k);\tilde{\beta}_k)/\partial\beta_k)'\hat{\beta}_{\varphi}(\beta_{k,0}))^4 \\
\leq \; & (\hat{\beta}_k-\beta_{k,0})^4\xi_{1,m_2}^4\left\|\hat{\beta}_{\varphi}(\beta_{k,0})\right\|^4 = O_p(\xi_{1,m_2}^4 n^{-2}). \qquad \text{(SC.261)}
\end{aligned}$$

By Assumptions SC2(v) and SC3(i, iv), Lemma SC4, (SC.54), (SC.117), and (SC.259)

$$\begin{aligned}
& n^{-1}\sum_{i=1}^{n}(\hat{P}_{2,i}(\beta_{k,0})'(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})\Delta\mathbf{K}_2-\varphi_i)^4 \\
\leq \; & Cn^{-1}\sum_{i=1}^{n}((\hat{P}_{2,i}(\beta_{k,0})-\tilde{P}_{2,i}(\beta_{k,0}))'\hat{\beta}_{\varphi}(\beta_{k,0}))^4 \\
& +Cn^{-1}\sum_{i=1}^{n}(\tilde{P}_{2,i}(\beta_{k,0})'(\hat{\beta}_{\varphi}(\beta_{k,0})-\tilde{\beta}_{\varphi,m_2}(\beta_{k,0})))^4 \\
& +Cn^{-1}\sum_{i=1}^{n}(\tilde{P}_{2,i}(\beta_{k,0})'\tilde{\beta}_{\varphi,m_2}(\beta_{k,0})-\varphi_i)^4 \\
\leq \; & C\left\|\hat{\beta}_{\varphi}(\beta_{k,0})\right\|^4\xi_{1,m_2}^4 n^{-1}\sum_{i=1}^{n}(\hat{\phi}_i-\phi_i)^4 \\
& +\xi_{1,m_2}^2\lambda_{\max}(n^{-1}\tilde{\mathbf{P}}_2(\beta_{k,0})'\tilde{\mathbf{P}}_2(\beta_{k,0}))\left\|\hat{\beta}_{\varphi}(\beta_{k,0})-\tilde{\beta}_{\varphi,m_2}(\beta_{k,0})\right\|^4+O(m_2^{-4r_{\varphi}}) \\
= \; & O_p((\xi_{1,m_2}^4\xi_{0,m_2}^2 m_1^2+\xi_{1,m_2}^2 m_2^2)n^{-2}). \qquad \text{(SC.262)}
\end{aligned}$$

Collecting the results in (SC.256), (SC.257), (SC.258), (SC.260), (SC.261) and (SC.262), we get

$$n^{-1}\sum_{i=1}^{n}(\hat{\varphi}_i-\varphi_i)^4 = O_p\left((m_2^4+m_1^2)\xi_{1,m_2}^4\xi_{0,m_2}^2 n^{-2}\right)$$

which together with (SC.253), (SC.254) and (SC.256) proves the third claim of the lemma.

(iv) By the definition of $\hat{u}_{2,i}$, we can write

$$\hat{u}_{2,i}-u_{2,i} = -l_{2,i}(\hat{\beta}_l-\beta_{l,0})-k_{2,i}(\hat{\beta}_k-\beta_{k,0})-(\hat{g}_i(\hat{\beta}_k)-g(\omega_{1,i}(\beta_{k,0});\beta_{k,0}))$$

which implies that

$$\begin{aligned} n^{-1}\sum_{i=1}^{n}(\hat{u}_{2,i}-u_{2,i})^4 &\leq C(\hat{\beta}_l-\beta_{l,0})^4 n^{-1}\sum_{i=1}^{n} l_{2,i}^4 + C(\hat{\beta}_k-\beta_{k,0})^4 n^{-1}\sum_{i=1}^{n} k_{2,i}^4 \\ &\quad + Cn^{-1}\sum_{i=1}^{n}(\hat{g}_i(\hat{\beta}_k)-g(\omega_{1,i}(\beta_{k,0});\beta_{k,0}))^4 \\ &= Cn^{-1}\sum_{i=1}^{n}(\hat{g}_i(\hat{\beta}_k)-g(\omega_{1,i}(\beta_{k,0});\beta_{k,0}))^4 + O_p(n^{-2}) \end{aligned} \quad \text{(SC.263)}$$

where the equality is by Assumptions SC1(i, iii) and SC2(i, ii), and Lemma SC19. Using similar arguments in showing Lemma SC15, we can show that

$$\begin{aligned} & \max_{i\leq n}\left|\hat{g}_i(\hat{\beta}_k)-g(\omega_{1,i}(\beta_{k,0});\beta_{k,0})\right|^2 \\ = \; & \xi_{0,m_2}^2||\hat{\beta}_g(\hat{\beta}_k)-\tilde{\beta}_{g,m_2}(\hat{\beta}_k)||^2 + O_p(m_2^{-2r_g}+n^{-1}) = O_p((m_1+m_2^2)\xi_{0,m_2}^2 n^{-1}) \end{aligned}$$

which together with (SB.12) and Lemma SC15 shows that

$$n^{-1}\sum_{i=1}^{n}(\hat{g}_i(\hat{\beta}_k)-g(\omega_{1,i}(\beta_{k,0});\beta_{k,0}))^4 = O_p((m_1^2+m_2^4)\xi_{0,m_2}^2 n^{-2}). \quad \text{(SC.264)}$$

The claim of the lemma follows from (SC.263) and (SC.264).

(v) Let $\hat{\beta}_h \equiv (\mathbf{P}_1'\mathbf{P}_1)^{-1}\sum_{i=1}^{n} P_1(x_{1,i})l_{1,i}$. By Assumptions SC1 and SC4(ii), we can use similar arguments in showing (SC.47) to get

$$\hat{\beta}_h-\beta_{h,m} = O_p(m_1^{1/2}n_1^{-1/2}+m_1^{-r_{h_1}}). \quad \text{(SC.265)}$$

Therefore by the triangle inequality, Assumption SC1(vi) and (SC.265),

$$\begin{aligned}\max_{i\leq n}|\hat{h}_i - h_i| &\leq \xi_{0,m_1}\left\|\hat{\beta}_h - \beta_{h,m}\right\| + \max_{i\leq n}|h_m(x_{1,i}) - h_i| \\ &= O_p(\xi_{0,m_1}m_1^{1/2}n_1^{-1/2} + \xi_{0,m_1}m_1^{-r_{h_1}}) = o_p(1)\end{aligned}$$

where the second equality is by Assumptions SC1(vi) and SC4(ii). *Q.E.D.*

Lemma SC25. *Under Assumptions SC1, SC2, SC3 and SC4, we have*

(i) $\hat{\Upsilon}_n - \Upsilon = o_p(1)$;

(ii) $\hat{\Gamma}_n - \Gamma = o_p(1)$;

(iii) $\hat{\Omega}_n - \Omega = o_p(1)$.

Proof of Lemma SC27. (i) By Assumptions SC1(i) and SC2(ii), and the Markov inequality

$$n^{-1}\sum_{i=1}^{n}(v_{2,i} - v_{1,i}g_{1,i})^2 = \Upsilon + O_p(n^{-1/2}) = O_p(1) \tag{SC.266}$$

which together with Assumption SC4(iv) and Lemma SC24(iii) proves the first claim of the lemma.

(ii) Let $\tilde{\Gamma}_n = \sum_{i=1}^{n}(l_{2,i} - h_{1,i}g_{1,i})(v_{2,i} - v_{1,i}g_{1,i})$. Then by Assumptions SC1(i, ii, iii) and SC2(i, ii), and the Slutsky Theorem, we have

$$n^{-1}\sum_{i=1}^{n}\varepsilon_{1,i}^2 = \mathbb{E}\left[\varepsilon_{1,i}^2\right] + O_p(n^{-1/2}) \tag{SC.267}$$

and

$$n^{-1}\sum_{i=1}^{n}(l_{2,i} - h_{1,i}g_{1,i})(v_{2,i} - v_{1,i}g_{1,i}) = \mathbb{E}\left[(l_{2,i} - h_{1,i}g_{1,i})(v_{2,i} - v_{1,i}g_{1,i})\right] + O_p(n^{-1/2}) \tag{SC.268}$$

which implies that

$$\tilde{\Gamma}_n = \Gamma + O_p(n^{-1/2}). \tag{SC.269}$$

By the definition of $\hat{\Gamma}_n$, we can write

$$\begin{aligned}\hat{\Gamma}_n - \tilde{\Gamma}_n &= n^{-1}\sum_{i=1}^{n}\Big[(l_{2,i} - \hat{h}_i\hat{g}_{1,i})(\hat{v}_{2,i} - \hat{v}_{1,i}\hat{g}_{1,i}) - (l_{2,i} - h_i g_{1,i})(v_{2,i} - v_{1,i}g_{1,i})\Big] \\ &= -n^{-1}\sum_{i=1}^{n}(\hat{h}_i\hat{g}_{1,i} - h_i g_{1,i})(\hat{v}_{2,i} - \hat{v}_{1,i}\hat{g}_{1,i} - v_{2,i} + v_{1,i}g_{1,i}) \\ &\quad -n^{-1}\sum_{i=1}^{n}(\hat{h}_i\hat{g}_{1,i} - h_i g_{1,i})(v_{2,i} - v_{1,i}g_{1,i}) \\ &\quad +n^{-1}\sum_{i=1}^{n}(l_{2,i} - h_i g_{1,i})(\hat{v}_{2,i} - \hat{v}_{1,i}\hat{g}_{1,i} - v_{2,i} + v_{1,i}g_{1,i}). \qquad \text{(SC.270)}\end{aligned}$$

The second claim of the lemma follows from (SC.270), Assumption SC3(iv) and Lemma SC24(ii, iii, v).

(iii) Since $\hat{\eta}_{1,i} = \eta_{1,i} - l_{1,i}(\hat{\beta}_l - \beta_{l,0}) - (\hat{\phi}_i - \phi_i)$, by Assumptions SC1(ii, iii) and SC3(iv), the Markov inequality and Lemma SC4

$$\begin{aligned}n^{-1}\sum_{i=1}^{n}(\hat{\eta}_{1,i} - \eta_{1,i})^4 &\leq C(\hat{\beta}_l - \beta_{l,0})^4 n^{-1}\sum_{i=1}^{n} l_{1,i}^4 + \max_{i\leq n}(\hat{\phi}_i - \phi_i)^2 n^{-1}\sum_{i=1}^{n}(\hat{\phi}_i - \phi_i)^2 \\ &= O_p(n^{-2}) + O_p(\xi_{0,m_1}^2 m_1^2 n^{-2}) = O_p(\xi_{0,m_1}^2 m_1^2 n^{-2}) = o_p(1). \qquad \text{(SC.271)}\end{aligned}$$

By Assumptions SC2(ii) and SC3(iv), and Lemma SC24(ii)

$$\max_{i\leq n}\hat{g}_{1,i}^4 \leq C\max_{i\leq n}(\hat{g}_{1,i} - g_{1,i})^4 + C\max_{i\leq n} g_{1,i}^4 = O_p(1). \qquad \text{(SC.272)}$$

By Assumptions SC1(i, ii) and SC3(iv), Lemma SC24(ii, iv), (SC.271) and (SC.272), we get

$$\begin{aligned}& n^{-1}\sum_{i=1}^{n}(\hat{u}_{2,i} - \hat{\eta}_{1,i}\hat{g}_{1,i} - u_{2,i} + \eta_{1,i}g_{1,i})^4 \\ \leq\ & Cn^{-1}\sum_{i=1}^{n}(\hat{u}_{2,i} - u_{2,i})^4 + C\max_{i\leq n}\hat{g}_{1,i}^4 n^{-1}\sum_{i=1}^{n}(\hat{\eta}_{1,i} - \eta_{1,i})^4 \\ & +C\max_{i\leq n}(\hat{g}_{1,i} - g_{1,i})^4 n^{-1}\sum_{i=1}^{n}\eta_{1,i}^4 = o_p(1) \qquad \text{(SC.273)}\end{aligned}$$

which together with Lemma SC24(iii), Assumptions SC1(i, ii), SC2(i, ii) and SC3(iv) implies that

$$n^{-1}\sum_{i=1}^{n}\begin{pmatrix} (\hat{u}_{2,i} - \hat{\eta}_{1,i}\hat{g}_{1,i})(\hat{v}_{2,i} - \hat{v}_{1,i}\hat{g}_{1,i}) \\ -(u_{2,i} - \eta_{1,i}g_{1,i})(v_{2,i} - v_{1,i}g_{1,i}) \end{pmatrix}^2 = o_p(1). \qquad \text{(SC.274)}$$

By Assumptions SC1(i, ii, iii) and SC4, and (SC.271), we have

$$n^{-1}\sum_{i=1}^{n}\hat{\varepsilon}_{1,i}^{4}+n^{-1}\sum_{i=1}^{n}\hat{\eta}_{1,i}^{4}=O_p(1) \tag{SC.275}$$

which combined with Lemma SC25(ii), (SC.271) and Assumption SC4 implies that

$$\begin{aligned} n^{-1}\sum_{i=1}^{n}(\hat{\Gamma}_n\hat{\varepsilon}_{1,i}\hat{\eta}_{1,i}-\Gamma\varepsilon_{1,i}\eta_{1,i})^2 &\leq C(\hat{\Gamma}_n-\Gamma)^2 n^{-1}\sum_{i=1}^{n}\hat{\varepsilon}_{1,i}^{2}\hat{\eta}_{1,i}^{2} \\ &+C\Gamma^2 n^{-1}\sum_{i=1}^{n}(\hat{\varepsilon}_{1,i}-\varepsilon_{1,i})^2\hat{\eta}_{1,i}^{2} \\ &+C\Gamma^2 n^{-1}\sum_{i=1}^{n}\varepsilon_{1,i}^{2}(\hat{\eta}_{1,i}-\eta_{1,i})^2=o_p(1). \end{aligned} \tag{SC.276}$$

Let $\tilde{\Omega}_n=n^{-1}\sum_{i=1}^{n}\left((u_{2,i}-\eta_{1,i}g_{1,i})(v_{2,i}-v_{1,i}g_{1,i})-\Gamma\varepsilon_{1,i}\eta_{1,i}\right)^2$. Then by Assumptions SC1(i) and SC2(ii), and the Markov inequality

$$\tilde{\Omega}_n=\Omega+O_p(n^{-1/2}). \tag{SC.277}$$

By the definition of $\tilde{\Omega}_n$ and $\hat{\Omega}_n$, the triangle inequality and the Cauchy-Schwarz inequality, (SC.274), (SC.276) and (SC.277), we get

$$\begin{aligned} \left|\hat{\Omega}_n-\tilde{\Omega}_n\right| &\leq Cn^{-1}\sum_{i=1}^{n}\begin{pmatrix} (\hat{u}_{2,i}-\hat{\eta}_{1,i}\hat{g}_{1,i})(\hat{v}_{2,i}-\hat{v}_{1,i}\hat{g}_{1,i}) \\ -(u_{2,i}-\eta_{1,i}g_{1,i})(v_{2,i}-v_{1,i}g_{1,i}) \end{pmatrix}^2 \\ &+Cn^{-1}\sum_{i=1}^{n}\left(\hat{\Gamma}_n\hat{\varepsilon}_{1,i}\hat{\eta}_{1,i}-\Gamma\varepsilon_{1,i}\eta_{1,i}\right)^2 \\ &+C\tilde{\Omega}_n^{1/2}\left(n^{-1}\sum_{i=1}^{n}\begin{pmatrix} (\hat{u}_{2,i}-\hat{\eta}_{1,i}\hat{g}_{1,i})(\hat{v}_{2,i}-\hat{v}_{1,i}\hat{g}_{1,i}) \\ -(u_{2,i}-\eta_{1,i}g_{1,i})(v_{2,i}-v_{1,i}g_{1,i}) \end{pmatrix}^2\right)^{1/2} \\ &+C\tilde{\Omega}_n^{1/2}\left(n^{-1}\sum_{i=1}^{n}(\hat{\Gamma}_n\hat{\varepsilon}_{1,i}\hat{\eta}_{1,i}-\Gamma\varepsilon_{1,i}\eta_{1,i})^2\right)^{1/2}=o_p(1) \end{aligned}$$

which together with (SC.277) proves the third claim of the Lemma. *Q.E.D.*

SC.5 Preliminary results

Lemma SC26 (Matrix Bernstein)**.** *Consider a finite sequence $\{d_i\}$ of independent, random matrices with dimension $m_1 \times m_2$. Assume that*

$$\mathbb{E}[d_i] = 0 \text{ and } \|d_i\|_S \leq \xi$$

where ξ is a finite constant. Introduce the random matrix $D_n = \sum_{i=1}^n d_i$. Compute the variance parameter

$$\sigma^2 = \max\left\{\left\|\sum_{i=1}^n \mathbb{E}\left[d_i d_i'\right]\right\|_S, \left\|\sum_{i=1}^n \mathbb{E}\left[d_i' d_i\right]\right\|_S\right\}.$$

Then for any $t \geq 0$

$$\mathbb{P}(\|D_n\|_S \geq t) \leq (m_1 + m_2)\exp\left(-\frac{t^2/2}{\sigma^2 + \xi t/3}\right).$$

The proof of the above lemma can be found in Tropp (2012).

Lemma SC27. *Let $S_{2,i}(\beta_k) = \tilde{P}_{2,i}(\beta_k)\tilde{P}_{2,i}(\beta_k)'$ where $\tilde{P}_{2,i}(\beta_k) = \tilde{P}_2(\omega_{1,i}(\beta_k), \beta_k)$ for any $\beta_k \in \Theta_k$. Then under Assumptions SC1(i) and SC2(iv, v, vi), we have*

$$\sup_{\beta_k \in \Theta_k}\left\|n^{-1}\sum_{i=1}^n S_{2,i}(\beta_k) - \mathbb{E}[S_{2,i}(\beta_k)]\right\|_S = O_p((\log(n))^{1/2}\xi_{0,m_2}n^{-1/2}).$$

Proof of Lemma SC27. For any $\beta_k \in \Theta_k$, by the triangle inequality and Assumptions SC2(iv, v),

$$\|S_{2,i}(\beta_k) - \mathbb{E}[S_{2,i}(\beta_k)]\|_S \leq \|S_{2,i}(\beta_k)\|_S + \|\mathbb{E}[S_{2,i}(\beta_k)]\|_S \leq C\xi_{0,m_2}^2. \tag{SC.278}$$

By Assumptions SC1(i) and SC2(iv, v),

$$\left\|\sum_{i=1}^n \mathbb{E}\left[(S_{2,i}(\beta_k) - \mathbb{E}[S_{2,i}(\beta_k)])^2\right]\right\|_S \leq n\left(\left\|\mathbb{E}\left[(S_{2,i}(\beta_k))^2\right]\right\|_S + \left\|(\mathbb{E}[S_{2,i}(\beta_k)])^2\right\|_S\right) \leq Cn\xi_{0,m_2}^2. \tag{SC.279}$$

Therefore we can use Lemma SC26 to deduce that

$$\mathbb{P}\left(\left\|n^{-1}\sum_{i=1}^n S_{2,i}(\beta_k) - \mathbb{E}[S_{2,i}(\beta_k)]\right\|_S \geq t\right) \leq 2m_2\exp\left(-\frac{1}{C}\frac{nt^2/2}{\xi_{0,m_2}^2(1+t/3)}\right) \tag{SC.280}$$

for any $\beta_k \in \Theta_k$ and any $t \geq 0$.

Since $k_{1,i}$ has bounded support, there exists a finite constant C_k such that $|k_{1,i}| \leq C_k$ for any *i*. Consider any $\beta_{k,1}, \beta_{k,2} \in \Theta_k$ and any $\gamma \in \mathbb{R}^{m_2}$ with $\|\gamma\| = 1$. By the triangle inequality,

$$\begin{aligned} \left\|S_{2,i}(\beta_{k,1}) - S_{2,i}(\beta_{k,2})\right\|_S \leq & \left\|S_{2,i}(\beta_{k,1}) - \tilde{P}_{2,i}(\beta_{k,2})\tilde{P}_{2,i}(\beta_{k,1})'\right\|_S \\ & + \left\|\tilde{P}_{2,i}(\beta_{k,2})\tilde{P}_{2,i}(\beta_{k,1})' - S_{2,i}(\beta_{k,2})\right\|_S . \end{aligned} \tag{SC.281}$$

By the mean value expansion and the Cauchy-Schwarz inequality, and Assumption SC2(v)

$$\begin{aligned} & \left|\gamma'(\tilde{P}_{2,i}(\beta_{k,1})\tilde{P}_{2,i}(\beta_{k,1})' - \tilde{P}_{2,i}(\beta_{k,2})\tilde{P}_{2,i}(\beta_{k,1})')\right|^2 \\ = & \left\|\tilde{P}_{2,i}(\beta_{k,1})\right\|^2 \left|\gamma'(\tilde{P}_{2,i}(\beta_{k,1}) - \tilde{P}_{2,i}(\beta_{k,2}))\right|^2 \\ = & \left\|\tilde{P}_{2,i}(\beta_{k,1})\right\|^2 \left|\gamma'\partial\tilde{P}_2\left(\omega_{1,i}(\tilde{\beta}_{k,12}); \tilde{\beta}_{k,12}\right)/\partial\beta_k\right|^2 (\beta_{k,1} - \beta_{k,2})^2 \\ \leq & \|\gamma\|^2 \xi_{0,m_2}^2 \xi_{1,m_2}^2 (\beta_{k,1} - \beta_{k,2})^2 \end{aligned}$$

where $\tilde{\beta}_{k,12}$ lies between $\beta_{k,1}$ and $\beta_{k,2}$, which together with Assumption SC2(vi) implies that

$$\left\|S_{2,i}(\beta_{k,1}) - \tilde{P}_{2,i}(\beta_{k,2})\tilde{P}_{2,i}(\beta_{k,1})'\right\|_S \leq C m_2^3 \left|\beta_{k,2} - \beta_{k,1}\right| . \tag{SC.282}$$

The same upper bound can be established for the second term in the right hand side of the inequality of (SC.281). Therefore,

$$\left\|S_{2,i}(\beta_{k,1}) - S_{2,i}(\beta_{k,2})\right\|_S \leq C m_2^3 \left|\beta_{k,2} - \beta_{k,1}\right| . \tag{SC.283}$$

Similarly, we can show that

$$\left\|\mathbb{E}\left[S_{2,i}(\beta_{k,1})\right] - \mathbb{E}\left[S_{2,i}(\beta_{k,2})\right]\right\|_S \leq C m_2^3 \left|\beta_{k,2} - \beta_{k,1}\right| . \tag{SC.284}$$

Combining the results in (SC.283) and (SC.284), and applying the triangle inequality, we get

$$\left\| \begin{array}{c} n^{-1}\sum_{i=1}^n \left(S_{2,i}(\beta_{k,1}) - \mathbb{E}\left[S_{2,i}(\beta_{k,1})\right]\right) \\ -n^{-1}\sum_{i=1}^n \left(S_{2,i}(\beta_{k,2}) - \mathbb{E}\left[S_{2,i}(\beta_{k,2})\right]\right) \end{array} \right\|_S \leq C_S m_2^3 \left|\beta_{k,2} - \beta_{k,1}\right| \tag{SC.285}$$

where C_S is a finite fixed constant. Since the parameter space Θ_k is compact, there exist $\{\beta_k(l)\}_{l=1,\ldots,K_n}$ such that for any $\beta_k \in \Theta_k$

$$\min_{l=1,\ldots,K_n} \left|\beta_k - \beta_k(l)\right| \leq (C_S m_2^3 n^{1/2})^{-1} \tag{SC.286}$$

where $K_n \leq 2C_S m_2^3 n^{1/2}$. For any $\beta_k \in \Theta_k$, by (SC.285) and (SC.286)

$$\left\| n^{-1}\sum_{i=1}^{n} S_{2,i}(\beta_k) - \mathbb{E}\left[S_{2,i}(\beta_k)\right] \right\|_S \leq \max_{l=1,\ldots,K_n} \left\| n^{-1}\sum_{i=1}^{n} S_{2,i}(\beta_k(l)) - \mathbb{E}\left[S_{2,i}(\beta_k(l))\right] \right\|_S + n^{-1/2}. \tag{SC.287}$$

Therefore for any $B > 1$,

$$\begin{aligned}
&\mathbb{P}\left(\sup_{\beta_k \in \Theta_k} \left\| n^{-1}\sum_{i=1}^{n} S_{2,i}(\beta_k) - \mathbb{E}\left[S_{2,i}(\beta_k)\right] \right\|_S \geq B(\xi_{0,m_2}^2 \log(n) n^{-1})^{1/2} \right) \\
&\leq \mathbb{P}\left(\max_{l=1,\ldots,K_n} \left\| n^{-1}\sum_{i=1}^{n} S_{2,i}(\beta_k(l)) - \mathbb{E}\left[S_{2,i}(\beta_k(l))\right] \right\|_S \geq (B-1)(\xi_{0,m_2}^2 \log(n) n^{-1})^{1/2} \right) \\
&\leq \sum_{l=1}^{K_n} \mathbb{P}\left(\left\| n^{-1}\sum_{i=1}^{n} S_{2,i}(\beta_k(l)) - \mathbb{E}\left[S_{2,i}(\beta_k(l))\right] \right\|_S \geq (B-1)(\xi_{0,m_2}^2 \log(n) n^{-1})^{1/2} \right) \\
&\leq 2K_n m_2 \exp\left(-\frac{B}{C} \frac{\log(n)}{1 + (\xi_{0,m_2}^2 \log(n) n^{-1})^{1/2}} \right)
\end{aligned} \tag{SC.288}$$

where the last inequality is by (SC.280). The claim of the theorem follows from (SC.288) and Assumption SC2(vi). *Q.E.D.*

Lemma SC28. *Let* $u_{2,i}(\beta_k) = y_{2,i}^* - k_{2,i}\beta_k - g(\omega_{1,i}(\beta_k), \beta_k)$*. Then under Assumptions SC1 and SC2(ii, iii, v, vi), we have*

$$\sup_{\beta_k \in \Theta_k} \left\| n^{-1}\sum_{i=1}^{n} \tilde{P}_2(\omega_{1,i}(\beta_k), \beta_k) u_{2,i}(\beta_k) \right\| = O_p(m_2 n^{-1/2}).$$

Proof of Lemma SC28. Define $\pi_n(\beta_k) = n^{-1/2}\sum_{i=1}^{n} \tilde{P}_2(\omega_{1,i}(\beta_k), \beta_k) u_{2,i}(\beta_k)$. For any $\beta_k \in \Theta_k$, by Assumption SC2(i) and (SC.60),

$$\mathbb{E}\left[(u_{2,i}(\beta_k))^4 \middle| \omega_{1,i}(\beta_k) \right] \leq C\mathbb{E}\left[(y_{2,i}^*)^4 + k_{2,i}^4 \middle| \omega_{1,i}(\beta_k) \right] + C\left| g(\omega_{1,i}(\beta_k); \beta_k) \right|^4 \leq C. \tag{SC.289}$$

For any $\beta_{k,1}, \beta_{k,2} \in \Theta_k$, by the i.i.d. assumption and the Cauchy-Schwarz inequality

$$\begin{aligned}
&\mathbb{E}\left[\|\pi_n(\beta_{k,1}) - \pi_n(\beta_{k,2})\|^2 \right] \\
&= \mathbb{E}\left[\left\| \tilde{P}_2(\omega_{1,i}(\beta_{k,1}), \beta_{k,1}) u_{2,i}(\beta_{k,1}) - \tilde{P}_2(\omega_{1,i}(\beta_{k,2}), \beta_{k,2}) u_{2,i}(\beta_{k,2}) \right\|^2 \right] \\
&\leq 2\mathbb{E}\left[(u_{2,i}(\beta_{k,2}))^2 \left\| \tilde{P}_2(\omega_{1,i}(\beta_{k,1}), \beta_{k,1}) - \tilde{P}_2(\omega_{1,i}(\beta_{k,2}), \beta_{k,2}) \right\|^2 \right] \\
&\quad + 2\mathbb{E}\left[\left\| \tilde{P}_2(\omega_{1,i}(\beta_{k,1}), \beta_{k,1}) \right\|^2 (u_{2,i}(\beta_{k,2}) - u_{2,i}(\beta_{k,1}))^2 \right].
\end{aligned} \tag{SC.290}$$

Consider any $\gamma \in \mathbb{R}^{m_2}$. By the mean value expansion and Assumption SC2(v)

$$\begin{aligned} & \left|\gamma'(\tilde{P}_2(\omega_1(\beta_{k,1}),\beta_{k,1}) - \tilde{P}_2(\omega_1(\beta_{k,2}),\beta_{k,2}))\right|^2 \\ = & \left|\gamma'\partial\tilde{P}_2\left(\omega_{1,i}(\tilde{\beta}_{k,12});\tilde{\beta}_{k,12}\right)/\partial\beta_k\right|^2(\beta_{k,1}-\beta_{k,2})^2 \leq \|\gamma\|^2\,\xi_{1,m_2}^2(\beta_{k,1}-\beta_{k,2})^2 \end{aligned}$$

where $\tilde{\beta}_{k,12}$ lies between $\beta_{k,1}$ and $\beta_{k,2}$, which implies that

$$\left\|\tilde{P}_2(\omega_{1,i}(\beta_{k,1}),\beta_{k,1}) - \tilde{P}_2(\omega_{1,i}(\beta_{k,2}),\beta_{k,2})\right\|^2 \leq \xi_{1,m_2}^2(\beta_{k,1}-\beta_{k,2})^2. \tag{SC.291}$$

Therefore, by (SC.289) and (SC.291),

$$\mathbb{E}\left[(u_{2,i}(\beta_{k,2}))^2\left\|\tilde{P}_2(\omega_{1,i}(\beta_{k,1}),\beta_{k,1}) - \tilde{P}_2(\omega_{1,i}(\beta_{k,2}),\beta_{k,2})\right\|^2\right] \leq C\xi_{1,m_2}^2(\beta_{k,2}-\beta_{k,1})^2. \tag{SC.292}$$

By the definition of $u_{2,i}(\beta_k)$, we can write

$$u_{1,i}(\beta_{k,2}) - u_{1,i}(\beta_{k,1}) = g(\omega_{1,i}(\beta_{k,1}),\beta_{k,1}) - g(\omega_{1,i}(\beta_{k,2}),\beta_{k,2}) + k_{2,i}(\beta_{k,2}-\beta_{k,1}).$$

Therefore, by Assumptions SC2(ii, iv), and the assumption that $k_{2,i}$ has bounded support, we have

$$\mathbb{E}\left[\left\|\tilde{P}_2(\omega_{1,i}(\beta_{k,1}),\beta_{k,1})\right\|^2(u_{1,i}(\beta_{k,2}) - u_{1,i}(\beta_{k,1}))^2\right] \leq Cm_2(\beta_{k,2}-\beta_{k,1})^2 \tag{SC.293}$$

which together with Assumption SC2(vi), (SC.290) and (SC.292) implies that

$$\|\|\pi_n(\beta_{k,1}) - \pi_n(\beta_{k,2})\|\|_2 \leq Cm_2^2\left|\beta_{k,2}-\beta_{k,1}\right| \tag{SC.294}$$

for any $\beta_{k,1}, \beta_{k,2} \in \Theta_k$.

We next use the chaining technique to prove the theorem. The proof follows similar arguments of proving Theorem 2.2.4 in van der Vaart and Wellner (1996). Construct nested sets $\Theta_{k,1} \subset \Theta_{k,2} \subset \cdots \subset \Theta_k$ such that $\Theta_{k,j}$ is a maximal set of points in the sense that for every $\beta_{k,j}$, $\beta'_{k,j} \in \Theta_{k,j}$ there is $\left|\beta_{k,j} - \beta'_{k,j}\right| > 2^{-j}$. Since Θ_k is a compact set, the number of the points in $\Theta_{k,j}$ is less than $C2^j$. Link every point $\beta_{k,j+1} \in \Theta_{k,j+1}$ to a unique $\beta_{k,j} \in \Theta_{k,j}$ such that $|\beta_{k,j+1} - \beta_{k,j}| \leq 2^{-j}$. Let $J_n = \min\{j : 2^{-j} \leq Cm_2^{-1}\}$. Consider any positive integer $J > J_n$. Obtain for every $\beta_{k,J+1}$ a chain $\beta_{k,J+1}, \ldots, \beta_{k,J_n}$ that connects it to a point β_{k,J_n} in Θ_{k,J_n}. For arbitrary points $\beta_{k,J+1}$,

$\beta'_{k,J+1}$ in $\Theta_{k,J+1}$, by the triangle inequality

$$\begin{aligned}
&\left\|\pi_n(\beta_{k,J+1}) - \pi_n(\beta_{k,J_n}) - \left[\pi_n(\beta'_{k,J+1}) - \pi_n(\beta'_{k,J_n})\right]\right\| \\
&= \left\|\sum_{j=J_n}^{J} \left[\pi_n(\beta_{k,j+1}) - \pi_n(\beta_{k,j})\right] - \sum_{j=J_n}^{J} \left[\pi_n(\beta'_{k,j+1}) - \pi_n(\beta'_{k,j})\right]\right\| \\
&\leq 2 \sum_{j=J_n}^{J} \max \|\pi_n(\beta_{k,j+1}) - \pi_n(\beta_{k,j})\| \qquad \text{(SC.295)}
\end{aligned}$$

where for fixed j the maximum is taken over all links $(\beta_{k,j+1}, \beta_{k,j})$ from $\Theta_{k,j+1}$ to $\Theta_{k,j}$. Thus the jth maximum is taken over at most $C2^{j+1}$ many links. By Assumption SC2(vi), (SC.294), (SC.295), the triangle inequality and the finite maximum inequality,

$$\begin{aligned}
&\left\|\max \left\|\pi_n(\beta_{k,J+1}) - \pi_n(\beta_{k,J_n}) - \left[\pi_n(\beta'_{k,J+1}) - \pi_n(\beta'_{k,J_n})\right]\right\|\right\|_2 \\
&\leq 2 \sum_{j=J_n}^{J} \|\max \|\pi_n(\beta_{k,j+1}) - \pi_n(\beta_{k,j})\|\|_2 \\
&\leq C \sum_{j=J_n}^{J} 2^{j/2} \max \|\|\pi_n(\beta_{k,j+1}) - \pi_n(\beta_{k,j})\|\|_2 \leq C m_2^2 \sum_{j=J_n}^{\infty} 2^{-j/2} \leq C m_2 \qquad \text{(SC.296)}
\end{aligned}$$

where β_{k,J_n} and β'_{k,J_n} are the endpoints of the chains starting at $\beta_{k,J+1}$ and $\beta'_{k,J+1}$ respectively. Since the set Θ_{k,J_n} has at most Cm_2 many elements, by the finite maximum inequality, the triangle inequality, (SC.289) and Assumption SC2(iv)

$$\left\|\max \left\|\pi_n(\beta_{k,J_n}) - \pi_n(\beta'_{k,J_n})\right\|\right\|_2 \leq C m_2^{1/2} \max \|\|\pi_n(\beta_{k,J_n})\|\|_2 \leq C m_2. \qquad \text{(SC.297)}$$

Therefore, by the triangle inequality, (SC.296) and (SC.297),

$$\begin{aligned}
&\left\|\max \left\|\pi_n(\beta_{k,J+1}) - \pi_n(\beta'_{k,J+1})\right\|\right\|_2 \\
&\leq \left\|\max \left\|\pi_n(\beta_{k,J+1}) - \pi_n(\beta_{k,J_n}) - \left[\pi_n(\beta'_{k,J+1}) - \pi_n(\beta'_{k,J_n})\right]\right\|\right\|_2 \\
&\qquad + \left\|\max \left\|\pi_n(\beta_{k,J_n}) - \pi_n(\beta'_{k,J_n})\right\|\right\|_2 \leq C m_2. \qquad \text{(SC.298)}
\end{aligned}$$

Let J go to infinity, by (SC.298) we deduce that

$$\left\|\sup_{\beta_k, \beta'_k \in \Theta_k} \left\|\pi_n(\beta_k) - \pi_n(\beta'_k)\right\|\right\|_2 \leq C m_2. \qquad \text{(SC.299)}$$

By (SC.297), (SC.299) and the triangle inequality,

$$\left\| \sup_{\beta_k \in \Theta_k} \|\pi_n(\beta_k)\| \right\|_2 \leq \left\| \sup_{\beta_k \in \Theta_k} \|\pi_n(\beta_k) - \pi_n(\beta_{k,0})\| \right\|_2 + \| \|\pi_n(\beta_{k,0})\| \|_2 \leq C m_2 \tag{SC.300}$$

which finishes the proof. *Q.E.D.*

Lemma SC29. *Under Assumptions SC1 and SC2(ii, iii, v, vi), we have*

$$\sup_{\beta_k \in \Theta_k} \left\| n^{-1} \sum_{i=1}^{n} u_{2,i} k_{1,i} \partial^1 \tilde{P}_2(\omega_{1,i}(\beta_k), \beta_k) \right\| = O_p(m_2^{5/2} n^{-1/2}).$$

PROOF OF LEMMA SC29. For ease of notations, we define $\pi_n(\beta_k) = n^{-1/2} \sum_{i=1}^{n} u_{2,i} k_{1,i} \partial^1 \tilde{P}_2(\omega_{1,i}(\beta_k), \beta_k)$ for any $\beta_k \in \Theta_k$. By Assumptions SC1(i) and (SC.80), $\mathbb{E}\left[u_{2,i}^2 k_{1,i}^2 | x_{1,i}\right] \leq C$. Therefore for any $\beta_{k,1}$ and $\beta_{k,2}$, we can use similar arguments in showing (SC.292) to obtain

$$\| \|\pi_n(\beta_{k,1}) - \pi_n(\beta_{k,2})\| \|_2 \leq C \xi_{2,m_2} \left| \beta_{k,1} - \beta_{k,2} \right|. \tag{SC.301}$$

Construct nested sets $\Theta_{k,1} \subset \Theta_{k,2} \subset \cdots \subset \Theta_k$ such that $\Theta_{k,j}$ is a maximal set of points in the sense that for every $\beta_{k,j}$, $\beta'_{k,j} \in \Theta_{k,j}$ there is $\left| \beta_{k,j} - \beta'_{k,j} \right| > 2^{-j}$. Since Θ_k is a compact set, the number of the points in $\Theta_{k,j}$ is less than $C2^j$. Link every point $\beta_{k,j+1} \in \Theta_{k,j+1}$ to a unique $\beta_{k,j} \in \Theta_{k,j}$ such that $|\beta_{k,j+1} - \beta_{k,j}| \leq 2^{-j}$. Let $J_n = \min\{j : 2^{-j} \leq C m_2^{-1}\}$. Consider any positive integer $J > J_n$. Obtain for every $\beta_{k,J+1}$ a chain $\beta_{k,J+1}, \ldots, \beta_{k,J_n}$ that connects it to a point β_{k,J_n} in Θ_{k,J_n}. For arbitrary points $\beta_{k,J+1}$, $\beta'_{k,J+1}$ in $\Theta_{k,J+1}$, by the triangle inequality and (SC.301)

$$\begin{aligned}
&\left\| \max \left\| \pi_n(\beta_{k,J+1}) - \pi_n(\beta_{k,J_n}) - \left[\pi_n(\beta'_{k,J+1}) - \pi_n(\beta'_{k,J_n}) \right] \right\| \right\|_2 \\
&\leq 2 \sum_{j=J_n}^{J} \| \max \|\pi_n(\beta_{k,j+1}) - \pi_n(\beta_{k,j})\| \|_2 \\
&\leq C \sum_{j=J_n}^{J} 2^{j/2} \max \| \|\pi_n(\beta_{k,j+1}) - \pi_n(\beta_{k,j})\| \|_2 \\
&\leq \xi_{2,m_2} \sum_{j=J_n}^{\infty} 2^{-j/2} \leq C \xi_{2,m_2} m_2^{-1}
\end{aligned} \tag{SC.302}$$

where β_{k,J_n} and β'_{k,J_n} are the endpoints of the chains starting at $\beta_{k,J+1}$ and $\beta'_{k,J+1}$ respectively. Since the set Θ_{k,J_n} has at most $C m_2$ many elements, by the finite maximum inequality, the triangle inequality, (SC.289) and Assumption SC2(iii)

$$\left\| \max \left\| \pi_n(\beta_{k,J_n}) - \pi_n(\beta'_{k,J_n}) \right\| \right\|_2 \leq C m_2^{1/2} \max \| \|\pi_n(\beta_{k,J_n})\| \|_2 \leq C m_2^{5/2}. \tag{SC.303}$$

Then the claim of the lemma follows by applying the chaining arguments in the proof of Lemma SC28. *Q.E.D.*

Lemma SC30. *Under Assumptions SC1 and SC2(ii, iii, v, vi), we have*

$$\sup_{\beta_k \in \Theta_k} \left\| n^{-1} \sum_{i=1}^{n} u_{2,i} k_{1,i} \partial^1 \tilde{P}_{2,i}(\beta_k) P_1(x_{1,i})' \right\| = O_p(m_2^{5/2} m_1^{1/2} n^{-1/2}).$$

Proof of Lemma SC30. For ease of notations, we define $n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}\partial^1 \tilde{P}_{2,i}(\beta_k)P_1(x_{1,i})'$ for any $\beta_k \in \Theta_k$. By Assumptions SC1(i) and (SC.80), $\mathbb{E}\left[u_{2,i}^2 k_{1,i}^2 | x_{1,i}\right] \leq C$. Therefore for any $\beta_{k,1}$ and $\beta_{k,2}$, we can use similar arguments in showing (SC.292) to obtain

$$\left\| \left\| \pi_n(\beta_{k,1}) - \pi_n(\beta_{k,2}) \right\| \right\|_2 \leq C m_1^{1/2} \xi_{2,m_2} \left| \beta_{k,1} - \beta_{k,2} \right|. \quad \text{(SC.304)}$$

Construct nested sets $\Theta_{k,1} \subset \Theta_{k,2} \subset \cdots \subset \Theta_k$ such that $\Theta_{k,j}$ is a maximal set of points in the sense that for every $\beta_{k,j}$, $\beta'_{k,j} \in \Theta_{k,j}$ there is $\left|\beta_{k,j} - \beta'_{k,j}\right| > 2^{-j}$. Since Θ_k is a compact set, the number of the points in $\Theta_{k,j}$ is less than $C2^j$. Link every point $\beta_{k,j+1} \in \Theta_{k,j+1}$ to a unique $\beta_{k,j} \in \Theta_{k,j}$ such that $|\beta_{k,j+1} - \beta_{k,j}| \leq 2^{-j}$. Let $J_n = \min\{j : 2^{-j} \leq Cm_2^{-1}\}$. Consider any positive integer $J > J_n$. Obtain for every $\beta_{k,J+1}$ a chain $\beta_{k,J+1}, \ldots, \beta_{k,J_n}$ that connects it to a point β_{k,J_n} in Θ_{k,J_n}. For arbitrary points $\beta_{k,J+1}$, $\beta'_{k,J+1}$ in $\Theta_{k,J+1}$, by the triangle inequality and (SC.301)

$$\left\| \max \left\| \pi_n(\beta_{k,J+1}) - \pi_n(\beta_{k,J_n}) - \left[\pi_n(\beta'_{k,J+1}) - \pi_n(\beta'_{k,J_n})\right] \right\| \right\|_2 \leq C\xi_{2,m_2} m_1^{1/2} m_2^{-1}. \quad \text{(SC.305)}$$

Since the set Θ_{k,J_n} has at most Cm_2 many elements, by the finite maximum inequality, the triangle inequality, (SC.289) and Assumption SC2(iii)

$$\left\| \max \left\| \pi_n(\beta_{k,J_n}) - \pi_n(\beta'_{k,J_n}) \right\| \right\|_2 \leq C m_2^{1/2} \max \left\| \left\| \pi_n(\beta_{k,J_n}) \right\| \right\|_2 \leq C\xi_{1,m_2} m_1^{1/2} m_2^{1/2}. \quad \text{(SC.306)}$$

Then the claim of the lemma follows by applying the chaining arguments in the proof of Lemma SC28. *Q.E.D.*

References

Ackerberg, D. A., K. Caves, and G. Frazer (2015): "Identification properties of recent production function estimators," *Econometrica*, 83(6), 2411–2451.

Andrews, D. W. (1991a): "Asymptotic optimality of generalized CL, cross-validation, and generalized cross-validation in regression with heteroskedastic errors," *Journal of Econometrics*, 47(2-3), 359–377.

Andrews, D. W. K. (1991b): "Asymptotic Normality of Series Estimators for Nonparametric and Semiparametric Regression Models," *Econometrica*, 59(2), 307–345.

Belloni, A., V. Chernozhukov, D. Chetverikov, and K. Kato (2015): "Some new asymptotic theory for least squares series: Pointwise and uniform results," *Journal of Econometrics*, 186(2), 345 – 366, High Dimensional Problems in Econometrics.

Chen, X. (2007): "Large Sample Sieve Estimation of Semi-Nonparametric Models," in *Handbook of Econometrics*, ed. by J. Heckman, and E. Leamer, vol. 6B, chap. 76. Elsevier, 1 edn.

Li, K.-C. (1987): "Asymptotic optimality for Cp, CL, cross-validation and generalized cross-validation: discrete index set," *The Annals of Statistics*, pp. 958–975.

Newey, W. K. (1997): "Convergence rates and asymptotic normality for series estimators," *Journal of Econometrics*, 79(1), 147 – 168.

Olley, G. S., and A. Pakes (1996): "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica: Journal of the Econometric Society*, pp. 1263–1297.

Tropp, J. A. (2012): "User-Friendly Tail Bounds for Sums of Random Matrices," *Foundations of Computational Mathematics*, 12(4), 389–434.

van der Vaart, A., and J. Wellner (1996): *Weak Convergence and Empirical Processes*. Springer-Verlag.