Supplemental Appendix to:
Identification and the Influence Function of Olley and Pakes’
(1996) Production Function Estimator

Jinyong Hahn
Department of Economics, UCLA

Zhipeng Liao
Department of Economics, UCLA

Geert Ridder
Department of Economics, USC

March 6, 2021

Abstract
This supplemental appendix contains additional technical results of Hahn, Liao, and Ridder (2021). Section SA provides detailed description of the multi-step estimator of $\beta_{k,0}$ mentioned in Hahn, Liao, and Ridder (2021). Section SB derives the asymptotic properties of the multi-step estimator and provides consistent estimation of its asymptotic variance. Section SC contains assumptions and auxiliary lemmas used in Section SB.
SA  The multi-step Series Estimator

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

$$y_{1,i} = l_{1,i} \beta_{t,0} + \phi(i_{1,i}, k_{1,i}) + \eta_{1,i}, \quad (SA.1)$$
$$\hat{y}_{2,i}^* = k_{2,i} \beta_{k,0} + g(\nu_{1,i}) + u_{2,i}, \quad (SA.2)$$

where $u_{2,i} \equiv \xi_{2,i} + \eta_{2,i}$, $y_{2,i}^* \equiv y_{2,i} - l_{2,i} \beta_{t,0}$ and $\nu_{1,i} \equiv \phi(i_{1,i}, k_{1,i}) - k_{1,i} \beta_{k,0}$, and $\xi_{2,i}$ is defined in equation (1) of the paper. The following restrictions are maintained throughout this appendix

$$E[\eta_{1,i} | i_{1,i}, k_{1,i}] = 0 \quad \text{and} \quad E[u_{2,i} | i_{1,i}, k_{1,i}] = 0. \quad (SA.3)$$

For any $\beta_k$, let

$$\nu_{1,i}(\beta_k) \equiv \phi(i_{1,i}, k_{1,i}) - k_{1,i} \beta_k \quad \text{and} \quad g(\nu_{1,i}(\beta_k); \beta_k) \equiv E[y_{2,i}^* - \beta_k k_{2,i} | \nu_{1,i}(\beta_k)]. \quad (SA.4)$$

Then by definition

$$\nu_{1,i} = \nu_{1,i}(\beta_{k,0}) \quad \text{and} \quad g(\nu_{1,i}) = g(\nu_{1,i}(\beta_{k,0}); \beta_{k,0}). \quad (SA.5)$$

The unknown parameters are $\beta_{t,0}$, $\beta_{k,0}$, $\phi(\cdot)$ and $g(\cdot; \beta_k)$ for any $\beta_k$ in $\Theta_k$, where $\Theta_k$ is a compact subset of $\mathbb{R}$ which contains $\beta_{k,0}$ as an interior point.

Suppose that we have data $\{(y_{t,i}, i_{t,i}, k_{t,i}, l_{t,i})_{t=1}^n\}_{i=1}^n$ and a preliminary estimator $\hat{\beta}_t$ of $\beta_{t,0}$. The asymptotic theory established here allows for a generic estimator of $\beta_{t,0}$, as long as certain regularity conditions (i.e., Assumptions [SC1(iii) and [SC4(i)] in Section SC) hold. For example, $\hat{\beta}_t$ 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 multi-step estimation procedure described in the paper.

**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 $\hat{y}_{1,i} \equiv y_{1,i} - l_{1,i} \hat{\beta}_t$. Then the unknown function $\phi(\cdot)$ is estimated by

$$\hat{\phi}(\cdot) \equiv P_1(\cdot)' (P_1' P_1)^{-1} (P_1' \hat{\mathbf{Y}}_1) \quad (SA.6)$$

where $P_1 \equiv (P_1(x_{1,1}), \ldots, P_1(x_{1,n}))'$ 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 $\hat{\beta}_t$ and $\hat{\phi}(\cdot)$ obtained in the first step, one can estimate $y_{2,i}^*$ by $\hat{y}_{2,i}^* \equiv y_{2,i} - \hat{\beta}_t l_{2,i}$ and estimate $\nu_{1,i}(\beta_k)$ by $\hat{\nu}_{1,i}(\beta_k) \equiv \hat{\phi}(x_{1,i}) - \beta_k k_{1,i}$. Let
\[ P_2(\nu) \equiv (p_{2,1}(\nu), \ldots, p_{2,m_2}(\nu))' \] be an \( m_2 \)-dimensional approximating functions. Then \( g(\cdot; \beta_k) \) is estimated by

\[ \hat{g}(\cdot; \beta_k) \equiv P_2(\cdot)\hat{\beta}_g(\beta_k), \quad \text{where} \quad \hat{\beta}_g(\beta_k) \equiv (P_2(\beta_k)'P_2(\beta_k))^{-1}P_2(\beta_k)'Y_2^*(\beta_k) \quad (SA.7) \]

where \( \hat{P}_2(\beta_k) \equiv (P_2(\hat{\nu}_{1,1}(\beta_k), \ldots, P_2(\hat{\nu}_{1,n}(\beta_k)))' \) and \( Y_2^*(\beta_k) \equiv (\hat{y}_{2,1}^* - \beta_k k_{2,1}, \ldots, \hat{y}_{2,n}^* - \beta_k k_{2,n})' \).

**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

\[ \hat{\beta}_k \equiv \arg \min_{\beta_k \in \Theta_k} n^{-1} \sum_{i=1}^n \hat{\ell}_i(\beta_k)^2, \quad \text{where} \quad \hat{\ell}_i(\beta_k) \equiv \hat{y}_{2,i}^* - k_{2,i}\beta_k - \hat{g}(\hat{\nu}_{1,i}(\beta_k); \beta_k). \quad (SA.8) \]

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 \( n^{-1} \sum_{i=1}^n \hat{\ell}_i(\beta_k)^2 \) to the population criterion function uniformly over \( \beta_k \in \Theta_k \). We impose the identification uniqueness condition of \( \beta_{k,0} \) in condition \( \text{(SB.9)} \) 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 \( \ell_i(\beta_k) \equiv y_{2,i} - l_{2,i}\beta_{k,0} - \beta_k k_{2,i} - g(\nu_{1,i}(\beta_k); \beta_k) \) for any \( \beta_k \in \Theta_k \). Suppose that for any \( \varepsilon > 0 \), there exists a constant \( \delta_\varepsilon > 0 \) such that

\[ \inf_{\{\beta_k \in \Theta_k : |\beta_k - \beta_{k,0}| \geq \varepsilon\}} \mathbb{E} \left[ \ell_i(\beta_k)^2 - \ell_i(\beta_{k,0})^2 \right] > \delta_\varepsilon. \quad \text{(SB.9)} \]

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:

$$n^{-1} \sum_{i=1}^{n} \hat{\ell}_i(\hat{\beta}_k) \left( k_{2,i} + \frac{\partial \hat{g}(\hat{\nu}_{1,i}(\hat{\beta}_k); \hat{\beta}_k)}{\partial \beta_k} \right) = 0$$  \hspace{1cm} (SB.10)

where for any $\beta_k \in \Theta_k$

$$\frac{\partial \hat{g}(\hat{\nu}_{1,i}(\hat{\beta}_k); \hat{\beta}_k)}{\partial \beta_k} = \hat{g}(\hat{\beta}_k) \frac{\partial P_2(\hat{\nu}_{1,i}(\hat{\beta}_k))}{\partial \nu} + P_2(\hat{\nu}_{1,i}(\hat{\beta}_k)) \frac{\partial \hat{g}(\hat{\beta}_k)}{\partial \beta_k}.  \hspace{1cm} (SB.11)$$

By the definition of $\hat{g}(\hat{\nu}_{1,i}(\hat{\beta}_k); \hat{\beta}_k)$ in (SA.7), we can write

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

which implies that

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

Therefore, the first-order condition (SB.10) can be reduced to

$$n^{-1} \sum_{i=1}^{n} \hat{\ell}_i(\hat{\beta}_k) \left( k_{2,i} + k_{1,i} \hat{g}(\hat{\beta}_k) \frac{\partial P_2(\hat{\nu}_{1,i}(\hat{\beta}_k))}{\partial \nu} \right) = 0  \hspace{1cm} (SB.12)$$

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

**Theorem SB1.** Let $\zeta_{1,i} \equiv k_{2,i} - E[k_{2,i}|\nu_{1,i}] - g_1(\nu_{1,i})(k_{1,i} - E[k_{1,i}|\nu_{1,i}])$ where

$$g_1(\nu) \equiv \frac{\partial g(\nu)}{\partial \nu}.  \hspace{1cm} (SB.13)$$

Suppose that

$$\Upsilon \equiv E[\zeta_{1,i}^2] > 0, \text{ where } \zeta_{1,i} \equiv k_{2,i} - E[k_{2,i}|\nu_{1,i}] - g_1(\nu_{1,i})(k_{1,i} - E[k_{1,i}|\nu_{1,i}]).  \hspace{1cm} (SB.14)$$

Then under (SB.9) in Lemma SB1 and Assumptions SC1, SC2 and SC3 in Section SC

$$\hat{\beta}_k - \beta_{k,0} = \Upsilon^{-1} n^{-1} \sum_{i=1}^{n} u_{2,i} \zeta_{1,i} - \Upsilon^{-1} n^{-1} \sum_{i=1}^{n} \eta_{1,i} g_1(\nu_{1,i}) (\zeta_{1,i} - \zeta_{2,i}) - \Upsilon^{-1} \Gamma(\hat{\beta}_l - \beta_{l,0}) + o_p(n^{-1/2}),  \hspace{1cm} (SB.15)$$
where \( \Gamma \equiv \mathbb{E}[l_{2,i} - l_{1,i}g_1(\nu_{1,i})] \varsigma_{1,i} + l_{1,i}g_1(\nu_{1,i}) \varsigma_{2,i} \) and \( \varsigma_{2,i} \equiv k_{2,i} - \mathbb{E}[k_{2,i}|x_{1,i}] \). Moreover

\[
n^{1/2}(\hat{\beta}_k - \beta_{k,0}) \to_d N(0, \Upsilon^{-1}\Omega\Upsilon^{-1}) \tag{SB.16}
\]

where \( \Omega \equiv \mathbb{E}\left[\left(\left(u_{2,i} - \eta_{1,i}g_1(\nu_{1,i})\right) \varsigma_{1,i} - \Gamma \varepsilon_{1,i} + \eta_{1,i}g_1(\nu_{1,i}) \varsigma_{2,i}\right)^2\right] \).

**Remark 1.** The local identification condition of \( \beta_{k,0} \) is imposed in (SB.14) which is important to ensure the root-n consistency of \( \hat{\beta}_k \). This condition is verified in Proposition 2 of the paper. □

**Remark 2.** The random variable \( \varepsilon_{1,i} \) in the definition of \( \Omega \) is from the linear representation of the estimator error in \( \hat{\beta}_l \), i.e.,

\[
\hat{\beta}_l - \beta_{l,0} = n^{-1} \sum_{i=1}^n \varepsilon_{1,i} + o_p(n^{-1/2}) \tag{SB.17}
\]

which is maintained in Assumption SC1(iii) in Section SC. The explicit form of \( \varepsilon_{1,i} \) depends on the estimation procedure of \( \beta_{l,0} \). For example, when \( \beta_{l,0} \) is estimated by the partially linear regression proposed in Olley and Pakes (1996),

\[
\varepsilon_{1,i} = \frac{l_{1,i} - \mathbb{E}[l_{1,i}|x_{1,i}]}{\mathbb{E}[l_{1,i} - \mathbb{E}[l_{1,i}|x_{1,i}]]} \eta_{1,i}.
\]

On the other hand, \( \varepsilon_{1,i} \) may take different forms in different estimation procedures (under different identification condition on \( \beta_{l,0} \)), such as the GMM procedure in Ackerberg, Caves, and Frazer (2015).

□

**Remark 3.** Since \( \mathbb{E}[s_{1,i}|\nu_{1,i}] = 0 \) for \( j = 1, 2 \),

\[
\mathbb{E}[l_{2,i}s_{1,i}] = \mathbb{E}[(l_{2,i} - \mathbb{E}[l_{2,i}|\nu_{1,i}]) \varsigma_{1,i}]
\]

by the law of iterated expectation. Similarly

\[
\mathbb{E}[l_{1,i}g_1(\nu_{1,i})(\varsigma_{1,i} - \varsigma_{2,i})] = \mathbb{E}[\mathbb{E}[l_{1,i}|x_{1,i}]g_1(\nu_{1,i})(\varsigma_{1,i} - \varsigma_{2,i})]
\]

Therefore we can write

\[
\Gamma = \mathbb{E}[l_{2,i}s_{1,i} - l_{1,i}g_1(\nu_{1,i})(\varsigma_{1,i} - \varsigma_{2,i})] = \mathbb{E}[(l_{2,i} - \mathbb{E}[l_{2,i}|\nu_{1,i}]) \varsigma_{1,i} - \mathbb{E}[l_{1,i}|x_{1,i}]g_1(\nu_{1,i})(\varsigma_{1,i} - \varsigma_{2,i})]. \tag{SB.18}
\]

\(^1\)See Assumption SC5 and Lemma SC26 in Subsection SC.5 for the regularity conditions and the derivation for (SB.17).
When the perpetual inventory method (PIM) i.e., \( k_{2,i} = (1 - \delta) k_{1,i} + i_{1,i} \) holds, we have \( \varsigma_{2,i} = 0 \) for any \( i = 1, \ldots, n \). Therefore, we deduce that
\[
\Gamma = \mathbb{E} \left[ (l_{2,i} - \mathbb{E}[l_{2,i}|\nu_{1,i}]) - g_{1}(\nu_{1,i})\mathbb{E}[l_{1,i}|x_{1,i}] \right] \varsigma_{1,i} \quad \text{(SB.19)}
\]
which appears in Proposition[3] of the paper. Moreover
\[
\Omega = \mathbb{E} \left[ ((u_{2,i} - \nu_{1,i} g_{1}(\nu_{1,i})) \varsigma_{1,i} - \Gamma \epsilon_{1,i})^{2} \right] \quad \text{(SB.20)}
\]
under PIM.

\[\text{□}\]

**Remark 4.** From the asymptotic expansion in (SB.15), we see that the asymptotic variance of \( \hat{\beta}_k \) is determined by three components. The first component, \( n^{-1} \sum_{i=1}^{n} u_{2,i} \varsigma_{1,i} \) comes from the third-step estimation with known \( \nu_{1,i} \). The second and the third components are from the first-step estimation. Specifically, the second one, \( n^{-1} \sum_{i=1}^{n} \eta_{1,i} g_{1}(\nu_{1,i}) (\varsigma_{1,i} - \varsigma_{2,i}) \) is from estimating \( \phi(\cdot) \) in the first step, while the third component \( \Gamma(\hat{\beta}_l - \beta_{l,0}) \) is due to the estimation error in \( \hat{\beta}_l \). \[\text{□}\]

**SB.2 Consistent variance estimation**

The asymptotic variance of \( \hat{\beta}_k \) can be estimated using the estimators of \( \varsigma_{1,i}, \varsigma_{2,i}, \epsilon_{1,i}, \eta_{1,i}, u_{2,i} \) and \( g_{1}(\nu_{1,i}) \). First, it is clear that \( g_{1}(\nu_{1,i}) \) can be estimated by \( \hat{g}_{1}(\hat{\nu}_{1,i}(\hat{\beta}_k); \hat{\beta}_k) \) where
\[
\hat{g}_{1}(\hat{\nu}_{1,i}(\hat{\beta}_k); \hat{\beta}_k) \equiv \hat{\beta}_g(\hat{\beta}_k)^{\prime} \frac{\partial P_{2}(\hat{\nu}_{1,i}(\hat{\beta}_k))}{\partial \nu} \quad \text{for any } \beta_k \in \Theta_k. \quad \text{(SB.21)}
\]
Second, the residual \( \varsigma_{1,i} \) can be estimated by
\[
\hat{\varsigma}_{1,i} \equiv \Delta \hat{k}_{2,i} - P_{2}(\hat{\nu}_{1,i}(\hat{\beta}_k))^{\prime}(\hat{P}_{2}(\hat{\beta}_k)^{\prime})^{-1} \sum_{i=1}^{n} P_{2}(\hat{\nu}_{1,i}(\hat{\beta}_k)) \Delta \hat{k}_{2,i} \quad \text{(SB.22)}
\]
where \( \Delta \hat{k}_{2,i} \equiv k_{2,i} - k_{1,i} \hat{g}_{1}(\hat{\nu}_{1,i}(\hat{\beta}_k); \hat{\beta}_k) \). Third, the residual \( \varsigma_{2,i} \) can be estimated by
\[
\hat{\varsigma}_{2,i} \equiv k_{2,i} - P_{1}(x_{1,i})^{\prime}(\hat{P}_{1}(x_{1,i})^{-1} \sum_{i=1}^{n} P_{1}(x_{1,i}) k_{2,i}. \quad \text{(SB.23)}
\]
Given the estimated residual \( \hat{\varsigma}_{1,i} \), the Hessian term \( \Upsilon \) can be estimated by
\[
\hat{\Upsilon}_n \equiv n^{-1} \sum_{i=1}^{n} \hat{\varsigma}_{1,i}^{2}. \quad \text{(SB.24)}
\]
Moreover the Jacobian term $\Gamma$ can be estimated by

$$
\hat{\Gamma}_n \equiv n^{-1} \sum_{i=1}^{n} \left[ (l_{2,i} - l_{1,i}\hat{\nu}_1(i\hat{\beta}_k))\hat{\varsigma}_{1,i} + l_{1,i}\hat{\nu}_1(i\hat{\beta}_k)\hat{\varsigma}_{2,i} \right].
$$

(SB.25)

Define $\hat{u}_{2,i} \equiv \hat{y}_{2,i} - l_{2,i}\hat{\beta}_l - k_{2,i}\hat{\beta}_k - \hat{g}(\hat{\nu}_1(i\hat{\beta}_k))$ and $\hat{\eta}_{1,i} \equiv y_{1,i} - l_{1,i}\hat{\beta}_l - \hat{\varphi}(x_{1,i})$. Then $\Omega$ is estimated by

$$
\hat{\Omega}_n \equiv n^{-1} \sum_{i=1}^{n} \left( (\hat{u}_{2,i} - \hat{\eta}_{1,i}\hat{g}_1(\hat{\nu}_1(i\hat{\beta}_k)))\hat{\varsigma}_{1,i} - \hat{\Gamma}_n\hat{\epsilon}_{1,i} + \hat{\eta}_{1,i}\hat{g}_1(\hat{\nu}_1(i\hat{\beta}_k))\hat{\varsigma}_{2,i} \right)^2
$$

(SB.26)

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

**Theorem SB2.** Suppose that the conditions in Theorem SB1 hold. Then under Assumption SC4 in Section SC, we have

$$
\hat{\Upsilon}_n = \Upsilon + o_p(1) \quad \text{and} \quad \hat{\Omega}_n = \Omega + o_p(1)
$$

(SB.27)

and moreover

$$
\frac{n^{1/2}(\hat{\beta}_k - \beta_k,0)}{(\hat{\Upsilon}_n^{-1} \hat{\Omega}_n \hat{\Upsilon}_n^{-1})^{1/2}} \rightarrow_d N(0,1).
$$

(SB.28)

**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.80) in the proof of Lemma SC8 and Assumption SC2(i)

$$
\sup_{\beta_k \in \Theta_k} \mathbb{E} \left[ \ell_i(\beta_k)^2 \right] \leq C
$$

(SB.29)

which together with Lemma SC8 implies that

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

(SB.30)

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

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

(SB.31)
By the definition of \( \hat{\ell}_i(\beta_k) \) and \( \ell_i(\beta_k) \), we can write

\[
\begin{align*}
\frac{1}{n} \sum_{i=1}^{n} \hat{\ell}_i(\beta_k)^2 &- \mathbb{E} \left[ \ell_i(\beta_k)^2 \right] \\
= \frac{1}{n} \sum_{i=1}^{n} (\ell_i(\beta_k)^2 - \mathbb{E} \left[ \ell_i(\beta_k)^2 \right]) + 2 \frac{1}{n} \sum_{i=1}^{n} \ell_i(\beta_k)(\hat{y}_{2,i}^* - y_{2,i}^*) \\
&- 2 \frac{1}{n} \sum_{i=1}^{n} \ell_i(\beta_k) (\hat{g}(\hat{\nu}_{1,i}(\beta_k); \beta_k) - g(\nu_{1,i}(\beta_k); \beta_k)) \\
&- 2 \frac{1}{n} \sum_{i=1}^{n} (\hat{y}_{2,i}^* - y_{2,i}^*) (\hat{g}(\hat{\nu}_{1,i}(\beta_k); \beta_k) - g(\nu_{1,i}(\beta_k); \beta_k)) \\
&+ \frac{1}{n} \sum_{i=1}^{n} (\hat{y}_{2,i}^* - y_{2,i}^*)^2 + \frac{1}{n} \sum_{i=1}^{n} \left( \hat{g}(\hat{\nu}_{1,i}(\beta_k); \beta_k) - g(\nu_{1,i}(\beta_k); \beta_k) \right)^2,
\end{align*}
\]

which together with Assumption SC2(vi), Lemma SC7, Lemma SC8, (SB.30), (SB.31) and the Cauchy-Schwarz inequality implies that

\[
\sup_{\beta_k \in \Theta_k} \left| \frac{1}{n} \sum_{i=1}^{n} \hat{\ell}_i(\beta_k)^2 - \mathbb{E} \left[ \ell_i(\beta_k)^2 \right] \right| = o_p(1). \tag{SB.32}
\]

The consistency of \( \hat{\beta}_k \) follows from its definition in (SA.8), (SB.32), the identification uniqueness condition of \( \beta_{k,0} \) assumed in (SB.9) and the standard arguments of showing the consistency of the extremum estimator. \( Q.E.D. \)

**Lemma SB2.** Let \( g_{1,i} \equiv g_1(\nu_{1,i}) \) and \( \hat{J}_i(\beta_k) \equiv \hat{\ell}_i(\beta_k) (k_{2,i} - k_{1,i}) \hat{g}_1(\hat{\nu}_{1,i}(\beta_k); \beta_k) \) for any \( \beta_k \in \Theta_k \), where \( \hat{g}_1(\hat{\nu}_{1,i}(\beta_k); \beta_k) \) is defined in (SB.21). Then under Assumptions SC1, SC2, and SC3 we have

\[
\frac{1}{n} \sum_{i=1}^{n} \hat{J}_i(\beta_{k,0}) = \frac{1}{n} \sum_{i=1}^{n} \left( u_{2,i} \varsigma_{1,i} - \eta_{1,i} g_1(\nu_{1,i}) (\varsigma_{1,i} - \varsigma_{2,i}) \right) - \Gamma(\hat{\beta}_l - \beta_{l,0}) + o_p(n^{-1/2}). \tag{SB.33}
\]

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

\[
\begin{align*}
\frac{1}{n} \sum_{i=1}^{n} \hat{\ell}_i(\beta_{k,0}) (k_{2,i} - k_{1,i}) \hat{g}_1(\hat{\nu}_{1,i}(\beta_{k,0}); \beta_{k,0})) \\
= \frac{1}{n} \sum_{i=1}^{n} (\hat{y}_{2,i}^* (\beta_{k,0}) - g(\nu_{1,i})) (k_{2,i} - k_{1,i}) g_1(\nu_{1,i}) \\
&- \frac{1}{n} \sum_{i=1}^{n} (\hat{g}(\hat{\nu}_{1,i}(\beta_{k,0}); \beta_{k,0}) - g(\nu_{1,i})) (k_{2,i} - k_{1,i}) g_1(\nu_{1,i}) + o_p(n^{-1/2}) \tag{SB.34}
\end{align*}
\]
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]

\[
\begin{align*}
  n^{-1} \sum_{i=1}^{n} \left( \hat{g}(\hat{\nu}_{1,i}(\beta_{k,0}); \beta_{k,0}) - g(\nu_{1,i}) \right)(k_{2,i} - k_{1,i}g_{1,i}) \\
  = n^{-1} \sum_{i=1}^{n} u_{2,i}(\nu_{1,i}) - \mathbb{E}[l_{2,i}(\nu_{1,i})]\left[\hat{\beta}_l - \beta_{l,0}\right] \\
  + n^{-1} \sum_{i=1}^{n} g_{1,i}(\hat{\phi}(x_{1,i}) - \phi(x_{1,i}))s_{1,i} + o_p(n^{-1/2}),
\end{align*}
\]

(SB.35)

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

\[
\begin{align*}
  n^{-1} \sum_{i=1}^{n} \left( \hat{y}_{2,i}^*(\beta_{k,0}) - g(\nu_{1,i}) \right)(k_{2,i} - k_{1,i}g_{1,i}) \\
  = n^{-1} \sum_{i=1}^{n} u_{2,i}(k_{2,i} - k_{1,i}g_{1,i}) - (\hat{\beta}_l - \beta_{l,0})n^{-1} \sum_{i=1}^{n} l_{2,i}(k_{2,i} - k_{1,i}g_{1,i}) \\
  = n^{-1} \sum_{i=1}^{n} u_{2,i}(k_{2,i} - k_{1,i}g_{1,i}) - (\hat{\beta}_l - \beta_{l,0})\mathbb{E}[l_{2,i}(k_{2,i} - k_{1,i}g_{1,i})] + o_p(n^{-1/2}) \quad \text{(SB.36)}
\end{align*}
\]

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

\[
\begin{align*}
  n^{-1} \sum_{i=1}^{n} l_{2,i}(k_{2,i} - k_{1,i}g_{1,i}) = \mathbb{E}[l_{2,i}(k_{2,i} - k_{1,i}g_{1,i})] + o_p(n^{-1/2})
\end{align*}
\]

which holds by the Markov inequality, Assumptions [SC1(i)] and [SC2(i, ii)]. Therefore by (SB.34), (SB.35) and (SB.36), we obtain

\[
\begin{align*}
  n^{-1} \sum_{i=1}^{n} \hat{\ell}_{i}(\beta_{k,0}) (k_{2,i} - k_{1,i}\hat{g}_{1}(\hat{\nu}_{1,i}(\beta_{k,0}); \beta_{k,0})) \\
  = n^{-1} \sum_{i=1}^{n} u_{2,i}s_{1,i} - (\hat{\beta}_l - \beta_{l,0})\mathbb{E}[l_{2,i}s_{1,i}] - n^{-1} \sum_{i=1}^{n} g_{1,i}(\hat{\phi}(x_{1,i}) - \phi(x_{1,i}))s_{1,i} + o_p(n^{-1/2}) \quad \text{(SB.37)}
\end{align*}
\]

The claim of the lemma follows from (SB.37) and Lemma [SC13] \( Q.E.D. \)

**Lemma SB3.** Under Assumptions [SC1], [SC2] and [SC5], we have

\[
\begin{align*}
  n^{-1} \sum_{i=1}^{n} (\hat{J}_i(\beta_k) - \hat{J}_i(\beta_{k,0})) = -(\hat{\beta}_k - \beta_{k,0}) \left( n^{-1} \sum_{i=1}^{n} g_{1,i}^2 + o_p(n^{-1/2}) \right). 
\end{align*}
\]
Proof of Lemma SB3. First note that by the definition of $\hat{J}_i(\beta_k)$ and $\hat{\ell}_i(\beta_k)$, we can write

$$n^{-1} \sum_{i=1}^{n} (\hat{J}_i(\beta_k) - \hat{J}_i(\beta_k,0))$$

$$= -(\hat{\beta}_k - \beta_k,0)n^{-1} \sum_{i=1}^{n} k_{2,i}(k_{2,i} - k_{1,i}\hat{g}_1(\hat{\nu}_{1,i}(\beta_k_k); \beta_k))$$

$$-n^{-1} \sum_{i=1}^{n} (\hat{g}(\hat{\nu}_{1,i}(\beta_k); \beta_k) - \hat{g}(\hat{\nu}_{1,i}(\beta_k,0); \beta_k))(k_{2,i} - k_{1,i}\hat{g}_1(\hat{\nu}_{1,i}(\beta_k,0); \beta_k,0))$$

$$-n^{-1} \sum_{i=1}^{n} u_{2,i}k_{1,i}(\hat{g}_1(\hat{\nu}_{1,i}(\beta_k); \beta_k) - \hat{g}_1(\hat{\nu}_{1,i}(\beta_k,0); \beta_k,0))$$

$$+(\hat{\beta}_l - \beta_{l,0})n^{-1} \sum_{i=1}^{n} l_{2,i}k_{1,i}(\hat{g}_1(\hat{\nu}_{1,i}(\beta_k); \beta_k) - \hat{g}_1(\hat{\nu}_{1,i}(\beta_k,0); \beta_k,0))$$

which together with Assumption SC1(iii), Lemma SC17, Lemma SC19 and Lemma SC23 implies that

$$n^{-1} \sum_{i=1}^{n} (\hat{J}_i(\beta_k) - \hat{J}_i(\beta_k,0)) = -(\hat{\beta}_k - \beta_k,0)E[k_{2,i}(k_{2,i} - k_{1,i}\varphi_{1,i})]$$

$$+ (\hat{\beta}_k - \beta_k,0) [E[k_{1,i}\varphi_{1,i}] + E[k_{2,i}\varphi_{1,i}]]$$

$$+ (\hat{\beta}_k - \beta_k,0) o_p(1) + o_p(n^{-1/2})$$

$$= -(\hat{\beta}_k - \beta_k,0) (E[\varphi_i^2] + o_p(1)) + o_p(n^{-1/2})$$

which finishes the proof. Q.E.D.

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

$$|\Gamma| \leq E \bigg[ |(l_{2,i} - l_{1,i}\varphi_{1,i})(\varphi_{1,i} + l_{1,i}\varphi_{1,i})| \bigg] \leq C$$

(SB.39)

and

$$\Omega = E \bigg[ (u_{2,i}^2 - \eta_{1,i}\varphi_{1,i})(\varphi_{1,i} + \eta_{1,i}\varphi_{1,i})^2 \bigg]$$

$$\leq C E[u_{2,i}^4 + \eta_{1,i}^4 + k_{1,i}^4 + k_{2,i}^4 + \varphi_{1,i}^2] \leq C.$$  

(SB.40)

By Assumption SC1(i), (SB.40) and the Lindeberg–Lévy central limit theorem,

$$n^{-1/2} \sum_{i=1}^{n} \bigg( (u_{2,i} - \eta_{1,i}\varphi_{1,i})(\varphi_{1,i} - \varphi_{1,i} + \eta_{1,i}\varphi_{1,i}) \bigg) \rightarrow_d N(0, \Omega).$$

(SB.41)
By (SB.12), Assumption SC1(iii), Lemma SB2 and Lemma SB3, we can write

\[
0 = n^{-1} \sum_{i=1}^{n} \hat{J}_i(\beta_{k,0}) + n^{-1} \sum_{i=1}^{n} (\hat{J}_i(\hat{\beta}_k) - \hat{J}_i(\beta_{k,0})) \\
= n^{-1} \sum_{i=1}^{n} (u_{2,i} \varsigma_{1,i} - \eta_{1,i} g_1(\nu_{1,i}) (\varsigma_{1,i} - \varsigma_{2,i})) - \Gamma n^{1/2} (\hat{\beta}_l - \beta_{l,0}) \\
- (\hat{\beta}_k - \beta_{k,0}) (E[\varsigma_{1,i}^2] + o_p(1)) + o_p(n^{-1/2})
\]  
(SB.42)

which together with (SB.14) and (SB.41) implies that

\[
\hat{\beta}_k - \beta_{k,0} = \Upsilon^{-1} n^{-1} \sum_{i=1}^{n} (u_{2,i} \varsigma_{1,i} - \eta_{1,i} g_1(\nu_{1,i}) (\varsigma_{1,i} - \varsigma_{2,i})) - \Upsilon^{-1} \Gamma (\hat{\beta}_l - \beta_{l,0}) + o_p(n^{-1/2}).  
\]  
(SB.43)

This proves (SB.15). The claim in (SB.16) follows from Assumption SC1(iii), (SB.41) and (SB.43). Q.E.D.

**Proof of Theorem SB2** The results in (SB.27) are proved in Lemma SC25(i, iii), which together with Theorem SB1, Assumption SC4(iii) and the Slutsky Theorem proves the claim in (SB.28). 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 conditions (SB.9) and (SB.14) are assumed throughout this section. 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 \(Q_{m_1} \equiv E[P_1(x_{1,i})P_1(x_{1,i})]\). 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) \(E[\eta_{1,i}|x_{1,i}] = 0\) and
\( \mathbb{E}[P_{i,1}^2 + \eta_{1,i}^2 | x_{1,i}] \leq C; \) (iii) there exist i.i.d. random variables \( \varepsilon_{1,i} \) with \( \mathbb{E}[\varepsilon_{1,i}^4] \leq C \) such that

\[
\hat{\beta}_1 - \beta_{t,0} = n^{-1} \sum_{i=1}^{n} \varepsilon_{1,i} + o_p(n^{-1/2});
\]

(iv) there exist \( r_\phi > 0 \) and \( \phi_{\phi,m} \in \mathbb{R}^m \) 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) \( C^{-1} \leq \min(Q_{m_1}) \leq \max(Q_{m_1}) \leq C \) uniformly over \( m_1 \); (vi) \( m_1^2n^{-1} + n^{1/2}m_1^{-r_\phi} = o(1) \) and \( \log(m_1)\xi_{0,m_1}^2m_1^{-1} = o(1) \) where \( \xi_{0,m_1} \) is a nondecreasing sequence such that \( \sup_{x \in \mathcal{X}} \| P_1(x) \| \leq \xi_{0,m_1} \).

Assumption SCI(iii) assumes that there exists a root-\( n \) consistent estimator \( \hat{\beta}_1 \) of \( \beta_{t,0} \). Different estimation procedures of \( \hat{\beta}_1 \) may give different forms for \( \varepsilon_{1,i} \). For example, \( \hat{\beta}_1 \) 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 SCI are fairly standard in series estimation; see, for example, Andrews (1991), Newey (1997) and Chen (2007). 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 SCI, we have

\[
n^{-1} \sum_{i=1}^{n} |\hat{\phi}(x_{1,i}) - \phi(x_{1,i})|^2 = O_p(m_1n^{-1}) \quad (\text{SC.44})
\]

and moreover

\[
\sup_{x_1 \in \mathcal{X}} |\hat{\phi}(x_1) - \phi(x_1)| = O_p(\xi_{0,m_1}m_1^{1/2}n^{-1/2}). \quad (\text{SC.45})
\]

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

\[
\|n^{-1}P_1^\prime P_1 - Q_{m_1}\|_S = O_p((\log m_1)^{1/2}\xi_{0,m_1}n^{-1/2}) = o_p(1) \quad (\text{SC.46})
\]

See SCI.270 in Subsection SCI.5 for the form of \( \varepsilon_{1,i} \) when \( \beta_{t,0} \) is estimated by the partially linear regression proposed in Olley and Pakes (1996).

For some approximating functions such as power series, Assumptions SCI (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.
which together with Assumption $\text{SC1}(v)$ implies that

$$C^{-1} \leq \lambda_{\min}(n^{-1}P_1'P_1) \leq \lambda_{\max}(n^{-1}P_1'P_1) \leq C$$  \hspace{1cm} (SC.47)

uniformly over $m_1$ with probability approaching 1 (wpa1). Since $\hat{y}_{1,i} = y_{1,i} - l_{1,i}\hat{\beta}_l = \phi(x_{1,i}) + \eta_{1,i} - l_{1,i}(\hat{\beta}_l - \beta_{l,0})$, we can write

$$\hat{\beta}_\phi - \beta_{\phi,m_1} = (P_1'P_1)^{-1} \sum_{i=1}^{n} P_1(x_{1,i})\eta_{1,i} + (P_1'P_1)^{-1} \sum_{i=1}^{n} P_1(x_{1,i})(\phi(x_{1,i}) - \phi_{m_1}(x_{1,i})) - (\hat{\beta}_l - \beta_{l,0})(P_1'P_1)^{-1} \sum_{i=1}^{n} P_1(x_{1,i})l_{1,i}. \hspace{1cm} (SC.48)$$

By Assumption $\text{SC1}(i, ii, v)$ and the Markov inequality

$$n^{-1} \sum_{i=1}^{n} P_1(x_{1,i})\eta_{1,i} = O_p(m_1^{1/2}n^{-1/2})) \hspace{1cm} (SC.49)$$

which together with Assumption $\text{SC1}(vi)$, (SC.46) and (SC.47) implies that

$$[(n^{-1}P_1'P_1)^{-1} - Q_{m_1}^{-1}]n^{-1} \sum_{i=1}^{n} P_1(x_{1,i})\eta_{1,i} = O_p((\log m_1)^{1/2}\xi_{0,m_1}m_1^{1/2}n^{-1}) = o_p(n^{-1/2}). \hspace{1cm} (SC.50)$$

By Assumption $\text{SC1}(iv, vi)$ and (SC.47)

$$(P_1'P_1)^{-1} \sum_{i=1}^{n} P_1(x_{1,i})(\phi(x_{1,i}) - \phi_{m_1}(x_{1,i})) = O_p(m^{-r_\phi}) = O_p(n^{-1/2}). \hspace{1cm} (SC.51)$$

Under Assumption $\text{SC1}(i, ii, v, vi)$, we can use similar arguments in showing (SC.49) to get

$$n^{-1} \sum_{i=1}^{n} P_1(x_{1,i})l_{1,i} - E[P_1(x_{1,i})l_{1,i}] = O_p(m_1^{1/2}n^{-1/2}) = o_p(1). \hspace{1cm} (SC.52)$$

By Assumption $\text{SC1}(i, ii, v)$,

$$\|E[l_{1,i}P_1(x_{1,i})]\|^2 \leq \lambda_{\max}(Q_{m_1})E[l_{1,i}P_1(x_{1,i})'^*Q_{m_1}]E[P_1(x_{1,i})l_{1,i}] \leq CE[l_{1,i}^2] \leq C \hspace{1cm} (SC.53)$$
which combined with (SC.52) implies that
\[ n^{-1} \sum_{i=1}^{n} P_1(x_{1,i})l_{1,i} = O_p(1). \] (SC.54)

By Assumption SCI (iii, v, vi), (SC.46), (SC.52), (SC.53) and (SC.54),
\[ (\hat{\beta}_t - \beta_{t,0})(P_1' P_1)^{-1} \sum_{i=1}^{n} P_1(x_{1,i})l_{1,i} = Q_{m_1}^{-1} E[P_1(x_{1,i})l_{1,i}](\hat{\beta}_t - \beta_{t,0}) + O_p(n^{-1/2}) \]
which combined with Assumption SCI (vi), (SC.48), (SC.50) and (SC.51) shows that
\[ \hat{\beta}_\phi - \beta_{\phi,m_1} = Q_{m_1}^{-1} \left( \sum_{i=1}^{n} P_1(x_{1,i})\eta_{1,i} - E[P_1(x_{1,i})l_{1,i}](\hat{\beta}_t - \beta_{t,0}) \right) + O_p(n^{-1/2}) = O_p(m_1^{1/2} n^{-1/2}) \] (SC.55)
where the second equality follows from Assumptions SCI (iii, v), (SC.49) and (SC.53). By the Cauchy-Schwarz inequality
\[
\begin{align*}
n^{-1} \sum_{i=1}^{n} |\hat{\phi}(x_{1,i}) - \phi(x_{1,i})|^2 & \leq 2n^{-1} \sum_{i=1}^{n} |\hat{\phi}(x_{1,i}) - \phi_{m_1}(x_{1,i})|^2 + 2n^{-1} \sum_{i=1}^{n} |\phi_{m_1}(x_{1,i}) - \phi(x_{1,i})|^2 \\
& \leq 2\lambda_{\text{max}}(n^{-1} P_1' P_1) \left\| \hat{\beta}_\phi - \beta_{\phi,m_1} \right\| ^2 + 2 \sup_{x \in \mathcal{X}} |\phi_{m_1}(x) - \phi(x)| = O_p(m_1^{1/2} n^{-1/2})
\end{align*}
\] (SC.56)
where the equality is by Assumptions SCI (iv, vi), (SC.47) and (SC.55), which proves (SC.44). By the triangle inequality, the Cauchy-Schwarz inequality, Assumption SCI (iv, vi) and (SC.55)
\[
\sup_{x_1 \in \mathcal{X}} |\hat{\phi}(x_1) - \phi(x_1)| \leq \sup_{x_1 \in \mathcal{X}} |\hat{\phi}(x_1) - \phi_{m_1}(x_1)| + \sup_{x_1 \in \mathcal{X}} |\phi_{m_1}(x_1) - \phi(x_1)| \\
\leq \xi_{0,m_1} \left\| \hat{\beta}_\phi - \beta_{\phi,m_1} \right\| + O(m_1^{-r_\phi}) = O_p(\xi_{0,m_1} m_1^{1/2} n^{-1/2}) \] (SC.57)
which proves the claim in (SC.44). \( Q.E.D. \)

\textbf{SC.2 Auxiliary results for the consistency of \( \hat{\beta}_k \)}

Recall that \( \nu_{1,i}(\beta_k) \equiv \phi(x_{1,i}) - \beta_k k_{1,i} \) and \( g(\nu; \beta_k) \equiv E[y_{2,i}^* - \beta_k k_{2,i}|\nu_{1,i}(\beta_k) = \nu] \). For any \( \beta_k \in \Theta_k \), let \( \Omega(\beta_k) \equiv [a_{\beta_k}, b_{\beta_k}] \) denote the support of \( \nu_{1,i}(\beta_k) \) with \( c_\nu < a_{\beta_k} < b_{\beta_k} < C_\nu \), where \( c_\nu \) and \( C_\nu \) 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 \( |g(\beta_k)|_d = \max_{0 \leq j \leq d} \sup_{\nu \in \Omega(\beta_k)} |\partial^j g(\nu; \beta_k)/\partial \nu^j| \).

\textbf{Assumption SC2.} (i) \( E[(y_{2,i}^*)^4 + I_{2,i}^4 + k_{2,i}^4|x_{1,i}] \leq C \); (ii) \( g(\nu; \beta_k) \) is twice 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 \( g_{m_2}(\nu; \beta_k) \equiv P_2(\nu)' \beta_{g,m_2}(\beta_k) ; \)

\[ 14 \]
(iv) for any $\beta_k \in \Theta_k$ there exists a nonsingular matrix $B(\beta_k)$ such that for $\tilde{P}_2(\nu_1(\beta_k); \beta_k) \equiv B(\beta_k)P_2(\nu_1(\beta_k))$,

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

uniformly over $\beta_k \in \Theta_k$, where $Q_{m_2}(\beta_k) \equiv E[\tilde{P}_2(\nu_1(\beta_k); \beta_k) \tilde{P}_2(\nu_1(\beta_k); \beta_k)^\prime]$; (v) for $j = 0, 1, 2, 3$, there exist sequences $\xi_{j,m_2}$ such that $\sup_{\beta_k \in \Theta_k} \sup_{\nu \in \Omega_c(\beta_k)} \| \partial^j \tilde{P}_2(\nu; \beta_k) / \partial \nu^j \| \leq \xi_{j,m_2}$ where $j_1 \leq j$ and $\varepsilon = m_2^{-2}$; (vi) $\xi_{j,m_2} \leq C m_2^{j+1}$ and $\xi_{0,m_1}(m_1^{1/2} m_2^{3/2} + (\log(n))^{1/2}) n^{-1/2} + n^{1/2} m_2^{-\tau_g} = o(1)$.

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

Let $\tilde{P}_2(\nu_1(\beta_k,1); \beta_k,2) \equiv B(\beta_k,2)P_2(\nu_1(\beta_k,1))$ for any $\beta_k,1, \beta_k,2 \in \Theta_k$. Since the series estimator $\hat{g}(\nu_1(\beta_k); \beta_k) = P_2(\nu_1(\beta_k))' \hat{\phi}(\beta_k)$ is invariant to any non-singular transformation on $P_2(\nu)$, throughout the rest of the Appendix we let

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

where $\tilde{P}_{2,i}(\beta_k) \equiv \tilde{P}_2(\nu_1(\beta_k); \beta_k)$. Define

$$\partial^j \tilde{P}_2(\nu; \beta_k) \equiv \frac{\partial^j \tilde{P}_2(\nu; \beta_k)}{\partial \nu^j} \quad \text{and} \quad \partial^j \tilde{P}_{2,i}(\beta_k) \equiv \partial^j \tilde{P}_2(\nu_1(\beta_k); \beta_k)$$

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

**Lemma SC5.** Under Assumptions $\text{SC1}$ and $\text{SC2}$, we have

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

**Proof of Lemma SC5.** Since $\hat{\nu}_1,i(\beta_k) = \hat{\phi}(x_{1,i}) - \beta_k k_{1,i}$, by Lemma SC4

$$\sup_{\beta_k \in \Theta_k} \max_{1 \leq n} \left| \hat{\nu}_1,i(\beta_k) - \nu_1,i(\beta_k) \right| = \max_{1 \leq n} \left| \hat{\phi}(x_{1,i}) - \phi(x_{1,i}) \right| = O_p(\xi_{0,m_1} m_1^{1/2} n^{-1/2}) = o_p(1) \quad (SC.58)$$

\footnote{Note that we define $\tilde{P}_{2,i}(\beta_k) \equiv P_2(\hat{\omega}_1,i(\beta_k))$ in Section SA. We change its definition here since the asymptotic properties of the series estimator $\hat{g}(\hat{\omega}_1,i(\beta_k); \beta_k) = P_2(\hat{\omega}_1,i(\beta_k))' \hat{\phi}(\beta_k)$ shall be investigated under the new definition $\tilde{P}_{2,i}(\beta_k) \equiv B(\beta_k)P_2(\hat{\omega}_1,i(\beta_k))$.}
which together with Assumption SC2(vi) implies that
\[ \hat{v}_{1,i}(\beta_k) \in \Omega_\varepsilon(\beta_k) \text{ wpa1} \tag{SC.59} \]
for any \( i \leq n \) and uniformly over \( \beta_k \in \Theta_k \). By the mean value expansion, we have for any \( b \in \mathbb{R}^{m_2} \)
\[ \left| b'(\hat{P}_{2,i}(\beta_k) - \hat{P}_{2,i}(\beta_k)) \right| = \left| b' \partial_1 \hat{P}_2 \left( \hat{v}_{1,i}(\beta_k); \beta_k \right) (\hat{v}_{1,i}(\beta_k) - v_{1,i}(\beta_k)) \right| \tag{SC.60} \]
where \( \hat{v}_{1,i}(\beta_k) \) lies between \( v_{1,i}(\beta_k) \) and \( \hat{v}_{1,i}(\beta_k) \). Since \( v_{1,i}(\beta_k) \) and \( \hat{v}_{1,i}(\beta_k) \) 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 \( \hat{v}_{1,i}(\beta_k) \). By the Cauchy-Schwarz inequality, Assumption SC2(v) and SC2(vi)
\[ \left| b'(\hat{P}_{2,i}(\beta_k) - \hat{P}_{2,i}(\beta_k)) \right| \leq \| b \| \xi_{1,m_2} |\hat{\phi}(x_{1,i}) - \phi(x_{1,i})| \text{ wpa1.} \]
Therefore,
\[ b'(\hat{P}_2(\beta_k) - \hat{P}_2(\beta_k))'(\hat{P}_2(\beta_k) - \hat{P}_2(\beta_k))b \]
\[ = \sum_{i=1}^{n} (b'(\hat{P}_{2,i}(\beta_k) - \hat{P}_{2,i}(\beta_k)))^2 \leq \| b \|^2 \xi_{1,m_2}^2 \sum_{i=1}^{n} |\hat{\phi}(x_{1,i}) - \phi(x_{1,i})|^2 \text{ wpa1,} \]
which together with Lemma SC4 implies that
\[ \sup_{\beta_k \in \Theta_k} \| \hat{P}_2(\beta_k) - \hat{P}_2(\beta_k) \|_S = O_p(\xi_{1,m_2}m_1^{1/2}). \tag{SC.61} \]
By Lemma SC32 and Assumption SC2(iv, vi), we have uniformly over \( \beta_k \in \Theta_k \)
\[ C^{-1} \leq \lambda_{\min}(n^{-1}\hat{P}_2(\beta_k)'\hat{P}_2(\beta_k)) \leq \lambda_{\max}(n^{-1}\hat{P}_2(\beta_k)'\hat{P}_2(\beta_k)) \leq C \text{ wpa1.} \tag{SC.62} \]
By the triangle inequality, Assumption SC2(vi), (SC.61) and (SC.62), we get
\[ \sup_{\beta_k \in \Theta_k} \left\| n^{-1}\hat{P}_2(\beta_k)'\hat{P}_2(\beta_k) - n^{-1}\hat{P}_2(\beta_k)'\hat{P}_2(\beta_k) \right\|_S \]
\[ \leq \sup_{\beta_k \in \Theta_k} n^{-1} \left\| (\hat{P}_2(\beta_k) - \hat{P}_2(\beta_k))'(\hat{P}_2(\beta_k) - \hat{P}_2(\beta_k)) \right\|_S \]
\[ \quad + \sup_{\beta_k \in \Theta_k} n^{-1} \left\| (\hat{P}_2(\beta_k) - \hat{P}_2(\beta_k))'(\hat{P}_2(\beta_k)) \right\|_S \]
\[ \quad + \sup_{\beta_k \in \Theta_k} n^{-1} \left\| \hat{P}_2(\beta_k)'(\hat{P}_2(\beta_k) - \hat{P}_2(\beta_k)) \right\|_S = O_p(\xi_{1,m_2}m_1^{1/2}n^{-1/2}) \]
which proves the claim of the lemma. \textit{Q.E.D.}
Lemma SC6. Under Assumptions SC1 and SC2, we have

$$\sup_{\beta_k \in \Theta_k} n^{-1} \sum_{i=1}^{n} \left| \tilde{P}_{2,i}(\beta_k)' \hat{\beta}_g(\beta_k) - g(\nu_{1,i}(\beta_k); \beta_k) \right|^2 = O_p((m^5_2/2 + \xi^2_{1,m_2}m_1)n^{-1}) = o_p(1)$$

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

**Proof of Lemma SC6** Let \( \tilde{\beta}_{g,m_2}(\beta_k) \equiv (B(\beta_k))^{-1} \beta_{g,m_2}(\beta_k) \) and \( \beta_{g,m_2}(\beta_k) \) is defined in Assumption SC2(iii). By the Cauchy-Schwarz inequality and Assumption SC2(iii)

$$n^{-1} \sum_{i=1}^{n} \left| \hat{P}_{2,i}(\beta_k)' \hat{\beta}_g(\beta_k) - g(\nu_{1,i}(\beta_k); \beta_k) \right|^2 \leq 2n^{-1} \sum_{i=1}^{n} \left| \hat{P}_{2,i}(\beta_k)' \hat{\beta}_g(\beta_k) - g_{m_2}(\nu_{1,i}(\beta_k); \beta_k) \right|^2$$

$$+ 2n^{-1} \sum_{i=1}^{n} \left| g_{m_2}(\nu_{1,i}(\beta_k); \beta_k) - g(\nu_{1,i}(\beta_k); \beta_k) \right|^2 \leq 2\lambda_{\max}(n^{-1} \hat{P}_2(\beta_k)' \hat{P}_2(\beta_k)) ||\hat{\beta}_g(\beta_k) - \tilde{\beta}_{g,m_2}(\beta_k)||^2 + Cm_2^{-2r_g} \quad (SC.63)$$

uniformly over \( \beta_k \in \Theta_k \), where \( g_{m_2}(\nu_{1,i}(\beta_k); \beta_k) \equiv \hat{P}_{2,i}(\beta_k)' \hat{\beta}_{g,m_2}(\beta_k) \) for any \( \beta_k \in \Theta_k \). We next show that

$$\sup_{\beta_k \in \Theta_k} \left| \hat{\beta}_g(\beta_k) - \tilde{\beta}_{g,m_2}(\beta_k) \right|^2 = O_p((m^5_2/2 + \xi^2_{1,m_2}m_1)n^{-1}) = o_p(1) \quad (SC.64)$$

which together with (SC.62) and (SC.63) proves the claim of the lemma.

Let \( u_{2,i}(\beta_k) \equiv y^*_2,i - k_{2,i}\beta_k - g(\nu_{1,i}(\beta_k), \beta_k) \). Then we can write

$$\hat{\beta}_g(\beta_k) - \tilde{\beta}_{g,m_2}(\beta_k) = (\hat{P}_2(\beta_k)' \hat{P}_2(\beta_k))^{-1} \hat{P}_2(\beta_k)'(\hat{Y}_2^*(\beta_k) - \hat{P}_2(\beta_k)' \tilde{\beta}_{g,m_2}(\beta_k))$$

$$= (\hat{P}_2(\beta_k)' \hat{P}_2(\beta_k))^{-1} \sum_{i=1}^{n} \hat{P}_{2,i}(\beta_k)(g(\nu_{1,i}(\beta_k), \beta_k) - g_{m_2}(\nu_{1,i}(\beta_k), \beta_k))$$

$$-(\hat{\beta}_1 - \beta_{1,0})(\hat{P}_2(\beta_k)' \hat{P}_2(\beta_k))^{-1} \sum_{i=1}^{n} \hat{P}_{2,i}(\beta_k)l_{2,i}$$

$$+(\hat{P}_2(\beta_k)' \hat{P}_2(\beta_k))^{-1} \sum_{i=1}^{n} \hat{P}_{2,i}(\beta_k)u_{2,i}(\beta_k) \quad (SC.65)$$

where \( g_{m_2}(\nu_{1,i}(\beta_k), \beta_k) \equiv \hat{P}_{2,i}(\beta_k)' \tilde{\beta}_{g,m_2}(\beta_k) \). By Assumption SC2(vi), Lemma SC5 and (SC.62), we have uniformly over \( \beta_k \in \Theta_k \)

$$C^{-1} \leq \lambda_{\min}(n^{-1} \hat{P}_2(\beta_k)' \hat{P}_2(\beta_k)) \leq \lambda_{\max}(n^{-1} \hat{P}_2(\beta_k)' \hat{P}_2(\beta_k)) \leq C \quad \text{wpa}1 \quad (SC.66)$$

which implies that \( \hat{P}_2(\beta_k)' \hat{P}_2(\beta_k)^{-1} \hat{P}_2(\beta_k)' \) is an idempotent matrix uniformly over \( \beta_k \in \Theta_k \).
Therefore,

$$\left\| (\hat{P}_2(\beta_k)' \hat{P}_2(\beta_k))^{-1} \sum_{i=1}^{n} \hat{P}_{2,i}(\beta_k)(g(\nu_{1,i}(\beta_k), \beta_k) - g_m(\hat{\nu}_{1,i}(\beta_k), \beta_k)) \right\|^2 \leq O_p(1) \sum_{i=1}^{n} (g(\nu_{1,i}(\beta_k), \beta_k) - g_m(\hat{\nu}_{1,i}(\beta_k), \beta_k))^2.$$  \hspace{1cm} (SC.67)

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

$$\sup_{\beta_k \in \Theta_k} |g(\nu_{1,i}(\beta_k); \beta_k)| \leq C.$$ \hspace{1cm} (SC.68)

Therefore,

$$\sup_{\beta_k \in \Theta_k} \left\| \beta_{g,m_2}(\beta_k) \right\|^2 \leq \sup_{\beta_k \in \Theta_k} (\lambda_{\min}(Q_{m_2}(\beta_k)))^{-1} \left\| \hat{P}_{2,i}(\beta_k)' \beta_{g,m_2}(\beta_k) \right\|^2_2 \leq C \sup_{\beta_k \in \Theta_k} \left\| g(\nu_{1,i}(\beta_k); \beta_k) - g_m(\nu_{1,i}(\beta_k); \beta_k) \right\|^2_2 + C \sup_{\beta_k \in \Theta_k} \left\| g(\nu_{1,i}(\beta_k); \beta_k) \right\|^2 \leq C.$$ \hspace{1cm} (SC.69)

By the second order expansion, Assumption SC2(iii, v, vi), Lemma SC4, (SC.68) and (SC.69), we have uniformly over $\beta_k \in \Theta_k$,

$$n^{-1} \sum_{i=1}^{n} (g_{m_2}(\nu_{1,i}(\beta_k), \beta_k) - g_{m_2}(\hat{\nu}_{1,i}(\beta_k), \beta_k))^2 \leq 2n^{-1} \sum_{i=1}^{n} (\partial_1 \hat{P}_{2,i}(\beta_k)' \beta_{g,m_2}(\beta_k)(\hat{\phi}(x_{1,i}) - \phi(x_{1,i})))^2$$

$$+ 2n^{-1} \sum_{i=1}^{n} (\partial_2 \hat{P}_2(\hat{\nu}_{1,i}(\beta_k); \beta_k)' \beta_{g,m_2}(\beta_k)(\hat{\phi}(x_{1,i}) - \phi(x_{1,i}))^2)^2$$

$$= O_p(m_n^{-1}) + O_p(\varepsilon_{g,m_2,m}^2\varepsilon_{m,1}^2n^{-2}) = O_p(m_n^{-1})$$

where $\hat{\nu}_{1,i}(\beta_k)$ is between $\nu_{1,i}(\beta_k)$ and $\hat{\nu}_{1,i}(\beta_k)$ and it lies in $\Omega_{\varepsilon}(\beta_k)$ uniformly over $\beta_k \in \Theta_k$ wpal
by (SC.59), which together with Assumption SC2(iii, vi) implies that
\[
    n^{-1} \sum_{i=1}^{n} (g(\nu_{1,i}(\beta_k), \beta_k) - g_{m2}(\hat{\nu}_{1,i}(\beta_k), \beta_k))^2 \\
    \leq C n^{-1} \sum_{i=1}^{n} (g(\nu_{1,i}(\beta_k), \beta_k) - g_{m2}(\hat{\nu}_{1,i}(\beta_k), \beta_k))^2 \\
    + C n^{-1} \sum_{i=1}^{n} (g_{m2}(\nu_{1,i}(\beta_k), \beta_k) - g_{m2}(\hat{\nu}_{1,i}(\beta_k), \beta_k))^2 \\
    = O_p(m_1n^{-1} + m_2^{-2r_g^2}) = O_p(m_1n^{-1}). \quad (SC.70)
\]

From (SC.67) and (SC.70), we get uniformly over \( \beta_k \in \Theta_k \)
\[
    (\hat{P}_2(\beta_k)'\hat{P}_2(\beta_k))^{-1} \sum_{i=1}^{n} \hat{P}_{2,i}(\beta_k)(g(\nu_{1,i}(\beta_k), \beta_k) - g_{m2}(\hat{\nu}_{1,i}(\beta_k), \beta_k)) = O_p(m_1^{1/2}n^{-1/2}). \quad (SC.71)
\]

By Assumptions SC1(i) and SC2(i), and the Markov inequality,
\[
    n^{-1} \sum_{i=1}^{n} l_{2,i}^2 = O_p(1) \quad (SC.72)
\]
which together with Assumption SC1(iii) and (SC.66) implies that
\[
    (\hat{\beta}_l - \beta_{l,0})(\hat{P}_2(\beta_k)'\hat{P}_2(\beta_k))^{-1} \sum_{i=1}^{n} \hat{P}_{2,i}(\beta_k) l_{2,i}^2 = O_p(n^{-1/2}) \quad (SC.73)
\]
uniformly over \( \beta_k \in \Theta_k \). By the mean value expansion, the Cauchy-Schwarz inequality and the triangle inequality, we have for any \( b \in \mathbb{R}^{m_2} \)
\[
    \left| n^{-1} \sum_{i=1}^{n} b'(\hat{P}_{2,i}(\beta_k) - \hat{P}_{2,i}(\beta_k))u_{2,i}(\beta_k) \right| \\
    = \left| n^{-1} \sum_{i=1}^{n} b'\partial^1 \hat{P}_2(\hat{\nu}_{1,i}(\beta_k); \beta_k)(\hat{\nu}_{1,i}(\beta_k) - \nu_{1,i}(\beta_k))u_{2,i}(\beta_k) \right| \\
    \leq \| b \| \xi_{1,m_2} n^{-1} \sum_{i=1}^{n} \left| (\hat{\phi}(x_{1,i}) - \phi(x_{1,i}))u_{2,i}(\beta_k) \right|. \quad (SC.74)
\]

By the definition of \( u_{2,i}(\beta_k) \), we can use Assumptions SC1(i) and SC2(i), (SC.68) and the Markov inequality to deduce
\[
    \sup_{\beta_k \in \Theta_k} n^{-1} \sum_{i=1}^{n} (u_{2,i}(\beta_k))^2 = O_p(1). \quad (SC.75)
\]
Thus by the Cauchy-Schwarz inequality, Lemma SC4 and (SC.75),

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

which together with (SC.66) and (SC.74) implies that

$$(\hat{P}_2(\beta_k)')\hat{P}_2(\beta_k))^{-1} \sum_{i=1}^{n} (\hat{P}_{2,i}(\beta) - \hat{P}_{2,i}(\beta_k)) u_{2,i}(\beta_k) = O_p(\xi_{1,m_2} m_1^{1/2} n^{-1/2})$$  \hspace{1cm} (SC.76)$$

uniformly over $\beta_k \in \Theta_k$. Applying Lemma SC33 and (SC.66) yields

$$(\hat{P}_2(\beta_k)')\hat{P}_2(\beta_k))^{-1} \sum_{i=1}^{n} \hat{P}_{2,i}(\beta_k) u_{2,i}(\beta_k) = O_p(m_2^{5/4} n^{-1/2})$$  \hspace{1cm} (SC.77)$$

uniformly over $\beta_k \in \Theta_k$. The claim in (SC.64) then follows from Assumption SC2(vi), (SC.65), (SC.71), (SC.73), (SC.76) and (SC.77).

Q.E.D.

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

$$
\sup_{\beta_k \in \Theta_k} n^{-1} \sum_{i=1}^{n} \left| \hat{\phi}(\hat{\nu}_{1,i}(\beta_k); \beta_k) - g(\nu_{1,i}(\beta_k); \beta_k) \right|^2 = O_p((m_2^{5/2} + \xi_{1,m_2}^2 m_1 n^{-1})) = o_p(1).
$$

**Proof of Lemma SC7.** By the triangle inequality, (SC.64) and (SC.69)

$$
\sup_{\beta_k \in \Theta_k} \left\| \hat{\beta}_g(\beta_k) \right\| \leq \sup_{\beta_k \in \Theta_k} \left\| \hat{\beta}_{g,m_2}(\beta_k) \right\| + \sup_{\beta_k \in \Theta_k} \left\| \hat{\beta}_g(\beta_k) - \hat{\beta}_{g,m_2}(\beta_k) \right\| = O_p(1).$$  \hspace{1cm} (SC.78)$$

By the mean value expansion, the Cauchy-Schwarz inequality, Assumption SC2(v, vi), Lemma SC4 and (SC.78),

$$
\sup_{\beta_k \in \Theta_k} n^{-1} \sum_{i=1}^{n} \left| (\hat{P}_{2,i}(\beta_k) - \hat{P}_{2,i}(\beta_k))' \hat{\beta}_g(\beta_k) \right|^2
$$

$$
= \sup_{\beta_k \in \Theta_k} n^{-1} \sum_{i=1}^{n} \left| \partial^1 \hat{P}_2(\hat{\nu}_{1,i}(\beta_k); \beta_k)' \hat{\beta}_g(\beta_k)(\hat{\nu}_{1,i}(\beta_k) - \nu_{1,i}(\beta_k)) \right|^2
$$

$$
\leq \xi_{1,m_2}^2 n^{-1} \sum_{i=1}^{n} \left( \hat{\phi}(x_{1,i}) - \phi(x_{1,i}) \right)^2 \sup_{\beta_k \in \Theta_k} \left\| \hat{\beta}_g(\beta_k) \right\| = O_p(\xi_{1,m_2}^2 m_1 n^{-1}) = o_p(1)$$  \hspace{1cm} (SC.79)$$

where $\hat{\nu}_{1,i}(\beta_k)$ is between $\hat{\nu}_{1,i}(\beta_k)$ and $\nu_{1,i}(\beta_k)$ and hence by (SC.59) it lies in $\Omega_\epsilon(\beta_k)$ 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.79).

Q.E.D.
Lemma SC8. Under Assumptions SC1 and SC2, we have

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

Proof of Lemma SC8. For any \( \beta_k \in \Theta_k \), by the Cauchy-Schwarz inequality and (SC.68),

\[
\ell_i(\beta_k)^2 \leq C \left[ (y_{2,i}^*)^2 + k_{2,i}^2,\beta_k^2 + g(\nu_{1,i}(\beta_k); \beta_k) \right] \leq C(1 + (y_{2,i}^*)^2 + k_{2,i}^2). \tag{SC.80}
\]

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

\[
|\ell_i(\beta_{k,1}) - \ell_i(\beta_{k,2})| \leq (C + k_{2,i}) |\beta_{k,1} - \beta_{k,2}|. \tag{SC.81}
\]

By Assumption SC2(ii), (SC.80) and (SC.81), we get

\[
\mathbb{E} \left[ |\ell_i(\beta_{k,1})^2 - \ell_i(\beta_{k,2})^2|^2 \right] \leq C(\beta_{k,1} - \beta_{k,2})^2 \tag{SC.82}
\]

for any \( \beta_k \in \Theta_k \), which implies that

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

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

\[
\|\ell_i(\beta_{k,1})^2 - \ell_i(\beta_{k,2})^2\|^2 \leq C |\beta_{k,1} - \beta_{k,2}|. \tag{SC.83}
\]

By Assumptions SC1(i) and SC2(i), and (SC.68),

\[
\mathbb{E} \left[ \left( n^{-1/2} \sum_{i=1}^{n} (\ell_i(\beta_k)^2 - \mathbb{E} [\ell_i(\beta_k)^2]) \right)^2 \right] = \mathbb{E} [\ell_i(\beta_k)^4] - (\mathbb{E} [\ell_i(\beta_k)^2])^2 \leq C \left( \mathbb{E} [(y_{2,i}^*)^4 + k_{2,i}^4,\beta_k^4 + g(\nu_{1,i}(\beta_k); \beta_k)] \right) \leq C
\]

for any \( \beta_k \in \Theta_k \), which implies that

\[
n^{-1/2} \sum_{i=1}^{n} (\ell_i(\beta_k)^2 - \mathbb{E} [\ell_i(\beta_k)^2]) = O_p(1) \tag{SC.84}
\]
for any \( \beta_k \in \Theta_k \). Moreover, by Assumption SC(i) and (SC.83)

\[
\mathbb{E} \left[ n^{-1/2} \sum_{i=1}^{n} \left( \ell_i(\beta_{k,1})^2 - \ell_i(\beta_{k,2})^2 - \mathbb{E} \left[ \ell_i(\beta_{k,1})^2 - \ell_i(\beta_{k,1})^2 \right] \right) \right]^2
\]

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

Collecting the results in (SC.84) and (SC.85), we can invoke Theorem 2.2.4 in van der Vaart and Wellner (1996) to deduce that

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

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 \( \varphi(\nu) \equiv \gamma_2(\nu) - \gamma_1(\nu)g_1(\nu) \) where \( g_1(\nu) \equiv \partial g(\nu)/\partial \nu \) and \( \gamma_j(\nu) \equiv \mathbb{E}[k_j,i|\nu_1,i = \nu] \) for \( j = 1, 2 \). For any \( \beta_k \in \Theta_k \) and \( i = 1, \ldots, n \), let

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

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

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

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

where \( \varphi_{m_2}(\nu) \equiv P_2(\nu)\beta_{\varphi,m_2} \); (iii) there exists \( \beta_{c,m_1} \in \mathbb{R}^{m_2} \) such that

\[
\left\| (\varsigma_{1,i} - \varsigma_{2,i})g_1(\nu_{1,i}) - P_1 (x_{1,i})' \beta_{c,m_1} \right\|_2 \to 0 \text{ as } m_1 \to \infty;
\]

(iv) \( n^{1/2}m_2^{-r_\varphi} + m_1m_2^2n^{-1/2} = o(1) \).

Assumptions SC3(i, ii) require that the function \( \varphi(\nu) \) is smooth and can be well approximated by the approximating functions \( P_2(\nu) \). By the definition of \( \varsigma_{1,i} \) and \( \varsigma_{2,i} \), we can write

\[
\varsigma_{1,i} - \varsigma_{2,i} = \mathbb{E}[k_{2,i}|x_{1,i}] - \mathbb{E}[k_{2,i}|\nu_{1,i}] - (k_{1,i} - \mathbb{E}[k_{1,i}|\nu_{1,i}])g_1(\nu_{1,i})
\]

which combined with Assumption SC2(i, ii) implies that \( (\varsigma_{1,i} - \varsigma_{2,i})g_1(\nu_{1,i}) \) is a function of \( x_{1,i} \) with finite \( L_2 \)-norm. Assumption SC3(iii) requires that \( (\varsigma_{1,i} - \varsigma_{2,i})g_1(\nu_{1,i}) \) can be approximated by the
approximating functions $P_i(x_{1,i})$. Assumption SC3(iv) restricts the numbers of the approximating functions and the smoothness of $\varphi(\nu)$.

**Lemma SC9.** Under Assumptions SC1, SC2 and SC3(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.65) and the results in (SC.71) and (SC.73) to get

$$
\hat{\beta}_g(\beta_{k,0}) - \tilde{\beta}_{g,m_2}(\beta_{k,0}) = (\hat{P}_2(\beta_{k,0})'\hat{P}_2(\beta_{k,0}))^{-1} \sum_{i=1}^n \hat{P}_{2,i}(\beta_{k,0})u_{2,i} + O_p(m_1^{1/2}n^{-1/2}). \tag{SC.86}
$$

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

$$
n^{-1} \sum_{i=1}^n b'(\hat{P}_{2,i}(\beta_{k,0}) - \tilde{P}_{2,i}(\beta_{k,0}))u_{2,i} = n^{-1} \sum_{i=1}^n b'\partial^1 \tilde{P}_{2,i}(\beta_{k,0})(\hat{\phi}_i - \phi_i)u_{2,i} \tag{SC.87}
$$

$$
+ n^{-1} \sum_{i=1}^n b'\partial^2 \tilde{P}_2(\nu_{1,i};\beta_{k,0})(\hat{\phi}_i - \phi_i)^2u_{2,i}
$$

where $\nu_{1,i}$ is between $\hat{\nu}_{1,i}(\beta_{k,0})$ and $\nu_{1,i}(\beta_{k,0})$. By (SC.59), $\tilde{\nu}_{1,i} \in \Omega_{\epsilon}(\beta_{k,0})$ for any $i = 1, \ldots, n$ wpa1. By Assumption SC2(i) and (SC.68),

$$
\mathbb{E} \left[ u_{2,i}^2 \right] \leq C. \tag{SC.88}
$$

By Assumption SC1(i, v, vi), (SC.88) and the Markov inequality

$$
\left\| n^{-1} \sum_{i=1}^n |u_{2,i}|P_1(x_{1,i})P_1(x_{1,i})' - \mathbb{E} \left[ |u_{2,i}|P_1(x_{1,i})P_1(x_{1,i})' \right] \right\| = o_p(1). \tag{SC.89}
$$

Since $\lambda_{\max}(\mathbb{E} \left[ |u_{2,i}|P_1(x_{1,i})P_1(x_{1,i})' \right]) \leq C$ by Assumption SC1(v) and (SC.88), from (SC.89) we deduce that

$$
\lambda_{\max} \left( n^{-1} \sum_{i=1}^n |u_{2,i}|P_1(x_{1,i})P_1(x_{1,i})' \right) \leq C \text{ wpa1.} \tag{SC.90}
$$

By (SC.55) and (SC.90), we get

$$
n^{-1} \sum_{i=1}^n |u_{2,i}(\hat{\phi}_i - \phi_{m_1,i})^2| = O_p(m_1n^{-1}) \tag{SC.91}
$$

23
where $\phi_{m_1,i} \equiv \phi_{m_1}(x_{1,i})$. By Assumption $\text{SCI}(i, \text{iv})$ and (SC.88), and the Markov inequality
\[
n^{-1} \sum_{i=1}^{n} |u_{2,i}(\phi_{m_1,i} - \phi_i)|^2 = O_p(m_1^{-2r\phi})
\]
which together with (SC.91) and Assumption $\text{SCI}(\text{vi})$ implies that
\[
n^{-1} \sum_{i=1}^{n} |u_{2,i}| (\hat{\phi}_i - \phi_i)^2 = O_p(m_1 n^{-1}). \tag{SC.92}
\]
By the Cauchy-Schwarz inequality and the triangle inequality, Assumption $\text{SC2}(\text{v})$ and (SC.92)
\[
\left| n^{-1} \sum_{i=1}^{n} b' \partial^2 \tilde{P}_2(\tilde{\nu}_{1,i}; \beta_{k,0}) (\hat{\phi}_i - \phi_i)^2 u_{2,i} \right| \leq \|b\| O_p(\xi_2 m_2 m_1 n^{-1}). \tag{SC.93}
\]
By Assumptions $\text{SCI}(i, \text{v})$ and $\text{SC2}(\text{v})$, and (SC.88),
\[
\mathbb{E} \left[ \left| n^{-1} \sum_{i=1}^{n} u_{2,i} \partial \tilde{P}_2,i (\beta_{k,0}) P_1(x_{1,i})' \right|^2 \right] \leq C \xi_1^2 m_1 m_2 n^{-1}
\]
which together with the Cauchy-Schwarz inequality, the Markov inequality and (SC.55) implies that
\[
\left| n^{-1} \sum_{i=1}^{n} u_{2,i} b' \partial \tilde{P}_2,i (\beta_{k,0}) (\phi_{m_2,i} - \phi_i) u_{2,i} \right| \leq \|b\| \|\hat{\beta}_\phi - \beta_{\phi,m_1}\| \left| n^{-1} \sum_{i=1}^{n} u_{2,i} \partial \tilde{P}_2,i (\beta_{k,0}) P_1(x_{1,i})' \right| = \|b\| \xi_1 m_2 m_1 n^{-1}. \tag{SC.94}
\]
By Assumptions $\text{SCI}(i, \text{iv, vi})$ and $\text{SC2}(\text{v})$, and (SC.88),
\[
\mathbb{E} \left[ \left| n^{-1} \sum_{i=1}^{n} \partial \tilde{P}_2,i (\phi_{m_2,i} - \phi_i) u_{2,i} \right|^2 \right] \leq C \xi_1^2 m_2 n^{-2}
\]
which together with the Cauchy-Schwarz inequality and the Markov inequality implies that
\[
\left| n^{-1} \sum_{i=1}^{n} b' \partial \tilde{P}_2,i (\beta_{k,0}) (\phi_{m_2,i} - \phi_i) u_{2,i} \right| \leq \|b\| \xi_1 m_2 n^{-1}. \tag{SC.95}
\]
Collecting the results in \((\text{SC.94})\) and \((\text{SC.95})\) obtains
\[
\left| n^{-1} \sum_{i=1}^{n} b' \partial_1 \tilde{P}_{2,i}(\beta_{k,0}) \left( \hat{\phi}_i - \phi_i \right) u_{2,i} \right| \leq \|b\| O_p(\xi_{1,m} m_1 n^{-1}). \tag{SC.96}
\]

Therefore, from Assumptions \(\text{SC2(vi)}\) and \(\text{SC3(iv)}\), \((\text{SC.66})\), \((\text{SC.87})\), \((\text{SC.93})\) and \((\text{SC.96})\) we can deduce
\[
(\hat{\mathbf{P}}_2(\beta_{k,0})' \hat{\mathbf{P}}_2(\beta_{k,0}))^{-1} \sum_{i=1}^{n} (\tilde{P}_{2,i}(\beta_{k,0}) - \hat{P}_{2,i}(\beta_{k,0})) u_{2,i} = O_p(m_1^{1/2} n^{-1/2}). \tag{SC.97}
\]

By Assumptions \(\text{SC1(i)}\) and \(\text{SC2(v)}\), and \((\text{SC.88})\),
\[
n^{-1} \sum_{i=1}^{n} \tilde{P}_{2,i}(\beta_{k,0}) u_{2,i} = O_p(m_2^{1/2} n^{-1/2})
\]

which together with \((\text{SC.66})\) implies that
\[
(\hat{\mathbf{P}}_2(\beta_{k,0})' \hat{\mathbf{P}}_2(\beta_{k,0}))^{-1} \sum_{i=1}^{n} \hat{P}_{2,i}(\beta_{k,0}) u_{2,i} = O_p(m_2^{1/2} n^{-1/2}). \tag{SC.98}
\]

The claim of the lemma follows from \((\text{SC.86})\), \((\text{SC.97})\) and \((\text{SC.98})\). \(Q.E.D.\)

**Lemma SC10.** Under Assumptions \(\text{SC1}, \text{SC2} \text{ and } \text{SC3} \), we have:
\[
n^{-1} \sum_{i=1}^{n} \hat{\ell}_i(\beta_{k,0}) k_{1,i}(\hat{g}_{1,i}(\beta_{k,0}) - g_1(\nu_{1,i})) = o_p(n^{-1/2}).
\]

**Proof of Lemma SC10** By the definition of \(\hat{\ell}_i(\beta_{k,0})\), we can write
\[
n^{-1} \sum_{i=1}^{n} \hat{\ell}_i(\beta_{k,0}) k_{1,i}(\hat{g}_{1,i}(\beta_{k,0}) - g_1(\nu_{1,i})) = n^{-1} \sum_{i=1}^{n} (g(\nu_{1,i}) - \hat{g}_i(\beta_{k,0})) k_{1,i}(\hat{g}_{1,i}(\beta_{k,0}) - g_1(\nu_{1,i})) + n^{-1} \sum_{i=1}^{n} (\hat{g}_{2,i}(\beta_{k,0}) - g(\nu_{1,i})) k_{1,i}(\hat{g}_{1,i}(\beta_{k,0}) - g_1(\nu_{1,i})). \tag{SC.99}
\]

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, Lemma SC9 (SC.66) and (SC.70)

\[
n^{-1} \sum_{i=1}^{n} (\hat{g}_i(\beta_{k,0}) - g(\nu_{1,i}))^2 \leq C n^{-1} \sum_{i=1}^{n} \left(\hat{P}_{2,i}(\beta_{k,0})' (\hat{\beta}_g(\beta_{k,0}) - \tilde{\beta}_{g,m_2}(\beta_{k,0}))\right)^2 \\
+ C n^{-1} \sum_{i=1}^{n} \left(\hat{P}_{2,i}(\beta_{k,0})' \tilde{\beta}_{g,m_2}(\beta_{k,0}) - g(\nu_{1,i}))^2 \leq C \left\| \hat{\beta}_g(\beta_{k,0}) - \tilde{\beta}_{g,m_2}(\beta_{k,0}) \right\|^2 \lambda_{\text{max}}(n^{-1} \hat{P}_2(\beta_{k,0})' \hat{P}_2(\beta_{k,0})) + O_p(m_1 n^{-1}) \\
= O_p((m_1 + m_2)n^{-1}) \tag{SC.100}
\]

Similarly, we can show that

\[
n^{-1} \sum_{i=1}^{n} (\hat{g}_{1,i}(\beta_{k,0}) - g_1(\nu_{1,i}))^2 \leq C n^{-1} \sum_{i=1}^{n} \left(\partial^1 \hat{P}_{2,i}(\beta_{k,0})' (\hat{\beta}_g(\beta_{k,0}) - \tilde{\beta}_{g,m_2}(\beta_{k,0}))\right)^2 \\
+ C n^{-1} \sum_{i=1}^{n} \left((\partial^1 \hat{P}_{2,i}(\beta_{k,0}) - \partial^1 \hat{P}_{2,i}(\beta_{k,0}))' \tilde{\beta}_{g,m_2}(\beta_{k,0})\right)^2 \\
+ C n^{-1} \sum_{i=1}^{n} \left(\partial^1 \hat{P}_{2,i}(\beta_{k,0})' \tilde{\beta}_{g,m_2}(\beta_{k,0}) - g_1(\nu_{1,i})\right)^2 \leq C \xi_{1,m_2}^2 \left\| \hat{\beta}_g(\beta_{k,0}) - \tilde{\beta}_{g,m_2}(\beta_{k,0}) \right\|^2 + O_p(\xi_2^2 m_1 n^{-1}) \\
= O_p(\xi_1^2 m_1 + m_2)n^{-1} + \xi_{2,m_2}^2 m_1 n^{-1}). \tag{SC.101}
\]

Therefore, by the Cauchy-Schwarz inequality, Assumption SC3(iv), (SC.100), and (SC.101),

\[
n^{-1} \sum_{i=1}^{n} (\hat{g}_i(\beta_{k,0}) - g(\nu_{1,i}))k_{1,i}(\hat{g}_{1,i}(\beta_{k,0}) - g_1(\nu_{1,i})) = o_p(n^{-1/2}). \tag{SC.102}
\]

Since \( \hat{y}_{2,i}^*(\beta_{k,0}) - g(\nu_{1,i}) = u_{2,i} - l_{1,i}(\hat{\beta}_1 - \beta_{1,0}) \), we can write

\[
n^{-1} \sum_{i=1}^{n} (\hat{y}_{2,i}^*(\beta_{k,0}) - g(\nu_{1,i}))k_{1,i}(\hat{g}_{1,i}(\beta_{k,0}) - g_1(\nu_{1,i})) \\
= n^{-1} \sum_{i=1}^{n} u_{2,i}k_{1,i}(\hat{g}_{1,i}(\beta_{k,0}) - g_1(\nu_{1,i})) \\
- (\hat{\beta}_1 - \beta_{1,0})n^{-1} \sum_{i=1}^{n} l_{1,i}k_{1,i}(\hat{g}_{1,i}(\beta_{k,0}) - g_1(\nu_{1,i})). \tag{SC.103}
\]

Since \( k_{1,i} \) has bounded support, by Assumptions SC1(i, ii, iii), SC2(vi) and SC3(iv), (SC.101) and
the Markov inequality,

\[(\hat{\beta}_l - \beta_{l,0})n^{-1} \sum_{i=1}^{n} l_{1,i} k_{1,i}(\hat{g}_{1,i}(\beta_{k,0}) - g_1(\nu_{1,i})) = o_p(n^{-1/2}). \]  \hspace{2cm} (SC.104)

Let

\[\partial^1 \tilde{P}_{2,i}(\beta_k) \equiv \partial^1 \tilde{P}_{2}(\nu_{1,i}(\beta_k); \beta_k) \text{ for any } \beta_k \in \Theta_k. \]

Then we can write

\[n^{-1} \sum_{i=1}^{n} u_{2,i} k_{1,i}(\hat{g}_{1,i}(\beta_{k,0}) - g_1(\nu_{1,i})) = \text{O}_p(\frac{n^{-1/2}}{\xi_{1,m_2}}). \]  \hspace{2cm} (SC.105)

By Assumptions SC1(i) and SC2(iii), (SC.88) and the Markov inequality, we have

\[n^{-1} \sum_{i=1}^{n} u_{2,i} k_{1,i}(\partial^1 \tilde{P}_{2,i}(\beta_{k,0}) - \partial^1 \tilde{P}_{2,i}(\beta_{k,0}))' \hat{\beta}_g(\beta_{k,0}) \]

\[+n^{-1} \sum_{i=1}^{n} u_{2,i} k_{1,i} \partial^1 \tilde{P}_{2,i}(\beta_{k,0})' (\hat{\beta}_g(\beta_{k,0}) - \tilde{\beta}_{g,m_2}(\beta_{k,0})) \]

\[+n^{-1} \sum_{i=1}^{n} u_{2,i} k_{1,i} \left(\partial^1 \tilde{P}_{2,i}(\beta_{k,0})' \tilde{\beta}_{g,m_2}(\beta_{k,0}) - g_1(\nu_{1,i})\right). \]  \hspace{2cm} (SC.106)

Similarly,

\[n^{-1} \sum_{i=1}^{n} u_{2,i} k_{1,i} \partial^1 \tilde{P}_{2,i}(\beta_{k,0}) = \text{O}_p(\frac{n^{-1/2}}{\xi_{1,m_2}}) \]

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

\[n^{-1} \sum_{i=1}^{n} u_{2,i} k_{1,i} \partial^1 \tilde{P}_{2,i}(\beta_{k,0})' (\hat{\beta}_g(\beta_{k,0}) - \tilde{\beta}_{g,m_2}(\beta_{k,0})) = o_p(n^{-1/2}). \]  \hspace{2cm} (SC.107)

By Assumption SC1(i), (SC.88) and the Markov inequality

\[n^{-1} \sum_{i=1}^{n} u_{2,i}^2 k_{1,i}^2 = \text{O}_p(1). \]  \hspace{2cm} (SC.108)
Let $\hat{\phi}_i \equiv \hat{\phi}(x_{1,i})$ and $\phi_i \equiv \phi(x_{1,i})$. By the second order expansion,

$$
 n^{-1} \sum_{i=1}^{n} u_{2,i}k_{1,i}(\partial^1 \hat{P}_{2,i}(\beta_{k,0}) - \partial^1 \tilde{P}_{2,i}(\beta_{k,0}))' \hat{\beta}_g(\beta_{k,0}) \\
 = n^{-1} \sum_{i=1}^{n} u_{2,i}k_{1,i}(\hat{\phi}_i - \phi_i)\partial^2 \tilde{P}_{2,i}(\beta_{k,0})' \hat{\beta}_g(\beta_{k,0}) \\
 + n^{-1} \sum_{i=1}^{n} u_{2,i}k_{1,i}(\hat{\phi}_i - \phi_i)^2 \partial^3 \tilde{P}_2(\tilde{\nu}_{1,i}(\beta_{k,0}); \beta_{k,0})' \hat{\beta}_g(\beta_{k,0}) \tag{SC.109}
$$

where $\tilde{\nu}_{1,i}(\beta_{k,0})$ is between $\tilde{\nu}_{1,i}(\beta_{k,0})$ and $\nu_{1,i}(\beta_{k,0})$. Using similar arguments for proving $\text{(SC.92)}$, we can show that

$$
 n^{-1} \sum_{i=1}^{n} |u_{2,i}k_{1,i}| (\hat{\phi}_i - \phi_i)^2 = O_p(m_1n^{-1}). \tag{SC.110}
$$

By the Cauchy-Schwarz inequality, Assumption $\text{SC2(v)}$, Lemma $\text{SC4}$, $\text{(SC.78)}$ and $\text{(SC.110)}$

$$
 n^{-1} \sum_{i=1}^{n} u_{2,i}k_{1,i}(\hat{\phi}_i - \phi_i)^2 \partial^3 \tilde{P}_2(\tilde{\nu}_{1,i}(\beta_{k,0}); \beta_{k,0})' \hat{\beta}_g(\beta_{k,0}) = O_p(\xi_3 m_2 n^{-1}) = o_p(n^{-1/2}) \tag{SC.111}
$$

where the second equality is by Assumptions $\text{SC2(vi)}$ and $\text{SC3(iv)}$. By Assumptions $\text{SC1(i, v)}$ and $\text{SC2(v)}$, and $\text{(SC.88)}$

$$
 n^{-1} \sum_{i=1}^{n} u_{2,i}k_{1,i} P_1(x_{1,i}) \partial^2 \tilde{P}_{2,i}(\beta_{k,0})' = O_p(\xi_2 m_2 n^{-1/2})
$$

which together with Lemma $\text{SC4}$ and $\text{(SC.78)}$ implies that

$$
 n^{-1} \sum_{i=1}^{n} u_{2,i}k_{1,i}(\hat{\phi}_i - \phi_i)\partial^2 \tilde{P}_{2,i}(\beta_{k,0})' \hat{\beta}_g(\beta_{k,0}) = O_p(\xi_2 m_2 n^{-1}) = o_p(n^{-1/2}) \tag{SC.112}
$$

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

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

which together with $\text{(SC.112)}$ implies that

$$
 n^{-1} \sum_{i=1}^{n} u_{2,i}k_{1,i}(\hat{\phi}_i - \phi_i)\partial^2 \tilde{P}_{2,i}(\beta_{k,0})' \hat{\beta}_g(\beta_{k,0}) = o_p(n^{-1/2}) \tag{SC.113}
$$
Collecting the results in (SC.109), (SC.111) and (SC.113) we get

$$n^{-1} \sum_{i=1}^{n} u_{2,i} k_{1,i} (\partial^{1} \hat{P}_{2,i}(\beta_{k,0}) - \partial^{1} \tilde{P}_{2,i}(\beta_{k,0})) \tilde{g}(k_{1,i} (\beta_{k,0})) = o_p(n^{-1/2}).$$  \(\text{(SC.114)}\)

By (SC.103), (SC.104), (SC.105), (SC.106), (SC.107) and (SC.114),

$$n^{-1} \sum_{i=1}^{n} (\hat{y}_{2,i}^{*}(\beta_{k,0}) - g(\nu_{1,i})) k_{1,i} g_{1}(k_{1,i} (\beta_{k,0}) - g_{1}(\nu_{1,i})) = o_p(n^{-1/2}).$$  \(\text{(SC.115)}\)

The claim of the lemma follows from (SC.99), (SC.102) and (SC.115).  \(Q.E.D.\)

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

$$n^{-1} \sum_{i=1}^{n} (\hat{P}_{2,i}(\beta_{k,0}) - \tilde{P}_{2,i}(\beta_{k,0})) \tilde{g}(k_{2,i} - k_{1,i} g_{1}(\nu_{1,i})) = o_p(n^{-1/2}).$$  \(\text{(SC.116)}\)

**Proof of Lemma SC11.** First we write

$$n^{-1} \sum_{i=1}^{n} (\hat{P}_{2,i}(\beta_{k,0}) - \tilde{P}_{2,i}(\beta_{k,0})) \tilde{g}(k_{2,i} - k_{1,i} g_{1}(\nu_{1,i}))$$

$$= n^{-1} \sum_{i=1}^{n} g_{1}(\nu_{1,i}) (\hat{g}(x_{1,i}) - g(x_{1,i})) (k_{2,i} - k_{1,i} g_{1}(\nu_{1,i})) + o_p(n^{-1/2}).$$

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

$$n^{-1} \sum_{i=1}^{n} (k_{2,i} - k_{1,i} g_{1}(\nu_{1,i}))^2 = O_p(1).$$  \(\text{(SC.117)}\)

Therefore by Assumption SC2(iii, vi) and (SC.117), we have

$$n^{-1} \sum_{i=1}^{n} (\hat{P}_{2,i}(\beta_{k,0}) - g(\nu_{1,i})) (k_{2,i} - k_{1,i} g_{1}(\nu_{1,i})) = o_p(n^{-1/2}).$$  \(\text{(SC.118)}\)
Recall that $\hat{\phi}_i \equiv \hat{\phi}(x_{1,i})$ and $\phi_i \equiv \phi(x_{1,i})$. By the second order expansion,

\[
\begin{align*}
&n^{-1} \sum_{i=1}^{n} \left( \hat{P}_{2,i}(\beta_{k,0}) - \bar{P}_{2,i}(\beta_{k,0}) \right) \beta_{g,m_2}(\beta_{k,0})(k_{2,i} - k_{1,i}g_1(\nu_{1,i})) \\
= & \ n^{-1} \sum_{i=1}^{n} \partial^1 \hat{P}_{2,i}(\beta_{k,0}) \beta_{g,m_2}(\beta_{k,0})(\hat{\phi}_i - \phi_i)(k_{2,i} - k_{1,i}g_1(\nu_{1,i})) \\
&+ n^{-1} \sum_{i=1}^{n} \partial^2 \hat{P}_{2}(\nu_{1,i}; \beta_{k,0}) \beta_{g,m_2}(\beta_{k,0})(\hat{\phi}_i - \phi_i)^2(k_{2,i} - k_{1,i}g_1(\nu_{1,i})). \quad (SC.119)
\end{align*}
\]

By the Cauchy-Schwarz inequality and the triangle inequality, Assumption $SC2(v)$, $SC.59$ and $SC.69$

\[
\begin{align*}
&\left| n^{-1} \sum_{i=1}^{n} \partial^2 \hat{P}_{2}(\nu_{1,i}; \beta_{k,0}) \beta_{g,m_2}(\beta_{k,0})(\hat{\phi}_i - \phi_i)^2(k_{2,i} - k_{1,i}g_1(\nu_{1,i})) \right| \\
\leq & \ \ O_p(\xi_{2,m_2}) n^{-1} \sum_{i=1}^{n} |k_{2,i} - k_{1,i}g_1(\nu_{1,i})| (\hat{\phi}_i - \phi_i)^2. \quad (SC.120)
\end{align*}
\]

Since $\mathbb{E}[|k_{2,i} - k_{1,i}g_1(\nu_{1,i})|^2 |x_{1,i}] \leq C$ by Assumption $SC2(i, ii)$, we can use the similar arguments for showing $SC.92$ to get

\[
\begin{align*}
&n^{-1} \sum_{i=1}^{n} |k_{2,i} - k_{1,i}g_1(\nu_{1,i})| (\hat{\phi}_i - \phi_i)^2 = O_p(m_1n^{-1})
\end{align*}
\]

which combined with Assumption $SC3(iv)$ and $SC.120$ implies that

\[
\begin{align*}
&n^{-1} \sum_{i=1}^{n} \partial^2 \hat{P}_{2}(\nu_{1,i}; \beta_{k,0}) \beta_{g,m_2}(\beta_{k,0})(\hat{\phi}_i - \phi_i)^2(k_{2,i} - k_{1,i}g_1(\nu_{1,i})) = o_p(n^{-1/2}). \quad (SC.121)
\end{align*}
\]

By the Cauchy-Schwarz inequality, Assumption $SC2(ii, iii, vi)$, Lemma $SC4$ and $SC.117$

\[
\begin{align*}
&n^{-1} \sum_{i=1}^{n} \partial^1 \hat{P}_{2,i}(\beta_{k,0}) \beta_{g,m_2}(\beta_{k,0})(\hat{\phi}_i - \phi_i)(k_{2,i} - k_{1,i}g_1(\nu_{1,i})) \\
= & \ n^{-1} \sum_{i=1}^{n} \partial^1 \hat{P}_{2,i}(\beta_{k,0}) \beta_{g,m_2}(\beta_{k,0})(\hat{\phi}_i - \phi_i)(k_{2,i} - k_{1,i}g_1(\nu_{1,i})) + o_p(n^{-1/2})
\end{align*}
\]
which together with (SC.119) and (SC.121) shows that

$$n^{-1} \sum_{i=1}^{n} (\hat{P}_{2,i}(\beta_{k,0}) - \tilde{P}_{2,i}(\beta_{k,0}))' \tilde{\beta}_{g,m_2}(\beta_{k,0})(k_{2,i} - k_{1,i}g_{1}(\nu_{1,i}))$$

$$= n^{-1} \sum_{i=1}^{n} g_{1}(\nu_{1,i})(\hat{\phi}_{i} - \phi_{i})(k_{2,i} - k_{1,i}g_{1}(\nu_{1,i})) + o_{p}(n^{-1/2}). \quad (SC.122)$$

The claim of the lemma follows from (SC.116), (SC.118) and (SC.122). Q.E.D.

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

$$n^{-1} \sum_{i=1}^{n} (\hat{g}_{i}(\beta_{k,0}) - g(\nu_{1,i}))(k_{2,i} - k_{1,i}g_{1}(\nu_{1,i}))$$

$$= n^{-1} \sum_{i=1}^{n} u_{2,i} \varphi(\nu_{1,i}) - \mathbb{E}[l_{2,i} \varphi(\nu_{1,i})](\hat{\beta}_{l} - \beta_{l,0})$$

$$+ n^{-1} \sum_{i=1}^{n} g_{1}(\nu_{1,i})(\hat{\phi}(x_{1,i}) - \phi(x_{1,i}))s_{1,i} + o_{p}(n^{-1/2})$$

where $\varphi(\nu_{1,i}) \equiv \mathbb{E}[k_{2,i} - k_{1,i}g_{1}(\nu_{1,i})|\nu_{1,i}]$ and $s_{1,i}$ is defined in (SB.14).

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

$$n^{-1} \sum_{i=1}^{n} (\hat{g}_{i}(\beta_{k,0}) - g(\nu_{1,i}))(k_{2,i} - k_{1,i}g_{1}(\nu_{1,i}))$$

$$= (\hat{\beta}_{g}(\beta_{k,0}) - \tilde{\beta}_{g,m_2}(\beta_{k,0}))'n^{-1} \sum_{i=1}^{n} \hat{P}_{2,i}(\beta_{k,0})(k_{2,i} - k_{1,i}g_{1}(\nu_{1,i}))$$

$$+ n^{-1} \sum_{i=1}^{n} (\hat{P}_{2,i}(\beta_{k,0})' \tilde{\beta}_{g,m_2}(\beta_{k,0}) - g(\nu_{1,i}))(k_{2,i} - k_{1,i}g_{1}(\nu_{1,i})). \quad (SC.123)$$

In view of Lemma SC11 and (SC.123), the claim of the lemma follows if

$$(\hat{\beta}_{g}(\beta_{k,0}) - \tilde{\beta}_{g,m_2}(\beta_{k,0}))'n^{-1} \sum_{i=1}^{n} \hat{P}_{2,i}(\beta_{k,0})(k_{2,i} - k_{1,i}g_{1}(\nu_{1,i}))$$

$$= n^{-1} \sum_{i=1}^{n} u_{2,i} \varphi(\nu_{1,i}) - \mathbb{E}[l_{2,i} \varphi(\nu_{1,i})](\hat{\beta}_{l} - \beta_{l,0})$$

$$- n^{-1} \sum_{i=1}^{n} g_{1}(\nu_{1,i})(\hat{\phi}(x_{1,i}) - \phi(x_{1,i}))\varphi(\nu_{1,i}) + o_{p}(n^{-1/2}). \quad (SC.124)$$

We next prove (SC.124).
Let \( \hat{\beta}_\varphi(\beta_{k,0}) \equiv (\hat{P}_2(\beta_{k,0})'\hat{P}_2(\beta_{k,0}))^{-1} \sum_{i=1}^{n} \hat{P}_{2,i}(\beta_{k,0})(k_{2,i} - k_{1,i}g_1(\nu_{1,i})) \). Then we can use the decomposition in (SC.65) to write

\[
(\hat{\beta}_g(\beta_{k,0}) - \hat{\beta}_{g,m_2}(\beta_{k,0}))'n^{-1} \sum_{i=1}^{n} \hat{P}_{2,i}(\beta_{k,0})(k_{2,i} - k_{1,i}g_1(\nu_{1,i})) = (\hat{\beta}_g(\beta_{k,0}) - \hat{\beta}_{g,m_2}(\beta_{k,0}))'(n^{-1}\hat{P}_2(\beta_{k,0})'\hat{P}_2(\beta_{k,0}))\hat{\beta}_\varphi(\beta_{k,0})
\]

\[
= n^{-1}\sum_{i=1}^{n} \hat{\beta}_\varphi(\beta_{k,0})'\hat{P}_{2,i}(\beta_{k,0})(g(\nu_{1,i}(\beta_{k,0}), \beta_{k,0}) - g_{m_2}(\nu_{1,i}(\beta_{k,0}), \beta_{k,0}))
\]

\[
- (\hat{\beta}_l - \beta_{l,0})n^{-1}\sum_{i=1}^{n} \hat{\beta}_\varphi(\beta_{k,0})'\hat{P}_{2,i}(\beta_{k,0})l_{2,i} + n^{-1}\sum_{i=1}^{n} \hat{\beta}_\varphi(\beta_{k,0})'\hat{P}_{2,i}(\beta_{k,0})u_{2,i} \quad (\text{SC.125})
\]

where \( g_{m_2}(\nu_{1,i}(\beta_{k,0}), \beta_{k,0}) \equiv \hat{P}_{2,i}(\beta_{k,0})'\hat{\beta}_{g,m_2}(\beta_{k,0}) \). Under Assumptions \( \text{SC1}, \text{SC2} \) and \( \text{SC3} \), we can use the same arguments for proving Lemma \( \text{SC9} \) to show that

\[
\hat{\beta}_\varphi(\beta_{k,0}) - \hat{\beta}_{\varphi,m_2}(\beta_{k,0}) = O_p((m_1^{1/2} + m_2^{1/2})n^{-1/2}) = o_p(1) \quad (\text{SC.126})
\]

where \( \hat{\beta}_{\varphi,m_2}(\beta_{k,0}) \equiv (B(\beta_{k,0})')^{-1}\beta_{\varphi,m_2} \) and \( \beta_{\varphi,m_2} \) is defined in Assumption \( \text{SC3}(\text{ii}) \). By Assumptions \( \text{SC1}(\text{i}, \text{v}) \) and \( \text{SC3}(\text{ii}, \text{iv}) \), and \( \text{(SC.126)} \), we can use similar arguments for showing \( \text{(SC.69)} \) and \( \text{(SC.78)} \) to deduce

\[
\left\| \hat{\beta}_{\varphi,m_2}(\beta_{k,0}) \right\| = O(1) \quad \text{and} \quad \left\| \hat{\beta}_\varphi(\beta_{k,0}) \right\| = O_p(1) \quad (\text{SC.127})
\]

Moreover, we can use similar arguments for proving \( \text{(SC.100)} \) to show that

\[
n^{-1}\sum_{i=1}^{n} (\hat{\beta}_\varphi(\beta_{k,0})'\hat{P}_{2,i}(\beta_{k,0}) - \varphi(\nu_{1,i}))^2 = O_p((m_1 + m_2)n^{-1}) \quad (\text{SC.128})
\]

The rest of the proof is divided into 3 steps. The claim in \( \text{(SC.124)} \) follows from \( \text{(SC.125)}, \text{(SC.129)}, \text{(SC.131)} \) and \( \text{(SC.133)} \) below.

**Step 1.** In this step, we show that

\[
n^{-1}\sum_{i=1}^{n} (\hat{\beta}_\varphi(\beta_{k,0})'\hat{P}_{2,i}(\beta_{k,0})(g(\nu_{1,i}(\beta_{k,0}), \beta_{k,0}) - g_{m_2}(\nu_{1,i}(\beta_{k,0}), \beta_{k,0}))
\]

\[
- n^{-1}\sum_{i=1}^{n} g_1(\nu_{1,i})(\hat{\phi}(x_{1,i}) - \phi(x_{1,i}))\varphi(\nu_{1,i}) + o_p(n^{-1/2}). \quad (\text{SC.129})
\]

Recall that \( \hat{\phi}_i \equiv \hat{\phi}(x_{1,i}) \) and \( \phi_i \equiv \phi(x_{1,i}) \). By Assumptions \( \text{SC2}(\text{iii}, \text{vi}) \) and \( \text{SC3}(\text{i}, \text{iv}) \), \( \text{SC.70} \) and
we get
\[
\begin{align*}
\sum_{i=1}^{n} \beta_{i} (\beta_{k,0})' \hat{P}_{2,i} (\beta_{k,0}) (g_{m2} (\nu_{1,i} (\beta_{k,0}) , \beta_{k,0}) & - g (\nu_{1,i} )) \\
= \sum_{i=1}^{n} \varphi (\nu_{1,i}) (g_{m2} (\nu_{1,i} (\beta_{k,0}) , \beta_{k,0}) & - g (\nu_{1,i} )) + O_{p}(n^{-1/2}) \\
= \sum_{i=1}^{n} \varphi (\nu_{1,i}) (\hat{P}_{2,i} (\beta_{k,0}) & - \hat{P}_{2,i} (\beta_{k,0}) )' \beta_{g,m2} (\beta_{k,0}) + O_{p}(n^{-1/2}).
\end{align*}
\] (SC.130)

By the second order expansion, Assumptions SC2 ii, iii, v, vi) and SC3 i, iv), Lemma SC4, SC59 and (SC.69)
\[
\begin{align*}
\sum_{i=1}^{n} \varphi (\nu_{1,i}) (\hat{P}_{2,i} (\beta_{k,0}) & - \hat{P}_{2,i} (\beta_{k,0}) )' \beta_{g,m2} (\beta_{k,0}) \\
= \sum_{i=1}^{n} \varphi (\nu_{1,i}) (\hat{\phi}_{i} - \phi_{i}) \partial \hat{P}_{2,i} (\beta_{k,0})' \beta_{g,m2} (\beta_{k,0}) \\
+ \sum_{i=1}^{n} \varphi (\nu_{1,i}) (\hat{\phi}_{i} - \phi_{i})^{2} \partial^{2} \hat{P}_{2,i} (\nu_{1,i} (\beta_{k,0}) ; \beta_{k,0})' \beta_{g,m2} (\beta_{k,0}) + O_{p}(n^{-1/2}) \\
= \sum_{i=1}^{n} \varphi (\nu_{1,i}) (\hat{\phi}_{i} - \phi_{i}) g_{1} (\nu_{1,i}) + O_{p}(n^{-1/2})
\end{align*}
\] which together with (SC.130) proves (SC.129).

**Step 2.** In this step, we show that
\[
(\hat{\beta}_{1} - \beta_{1,0}) n^{-1} \sum_{i=1}^{n} \beta_{\varphi} (\beta_{k,0})' \hat{P}_{2,i} (\beta_{k,0}) l_{2,i} = \mathbb{E} [l_{2,i} \varphi (\nu_{1,i}) ] (\hat{\beta}_{1} - \beta_{1,0}) + O_{p}(n^{-1/2}).
\] (SC.131)

By the Cauchy-Schwarz inequality, (SC.72) and (SC.128)
\[
\begin{align*}
\sum_{i=1}^{n} \beta_{\varphi} (\beta_{k,0})' \hat{P}_{2,i} (\beta_{k,0}) l_{2,i} &= \sum_{i=1}^{n} \varphi (\nu_{1,i}) l_{2,i} + O_{p}(m_{1}^{1/2} + m_{2}^{1/2}) n^{-1/2} \\
&= \mathbb{E} [l_{2,i} \varphi (\nu_{1,i}) ] + O_{p}(m_{1}^{1/2} + m_{2}^{1/2}) n^{-1/2}
\end{align*}
\] (SC.132)

where the second equality is by the Markov inequality, Assumptions SC1 i, SC2 i and SC3 i. The claim in (SC.131) follows by Assumptions SC1 iii, SC2 i and SC3 ii, vi), and (SC.132).

**Step 3.** In this step, we show that
\[
\sum_{i=1}^{n} \beta_{\varphi} (\beta_{k,0})' \hat{P}_{2,i} (\beta_{k,0}) u_{2,i} = n^{-1} \sum_{i=1}^{n} u_{2,i} \varphi (\nu_{1,i}) + O_{p}(n^{-1/2}).
\] (SC.133)
By the second order expansion,

\[ n^{-1} \sum_{i=1}^{n} \beta_\varphi(\beta_{k,0})' \hat{P}_{2,i}(\beta_{k,0}) u_{2,i} = n^{-1} \sum_{i=1}^{n} \beta_\varphi(\beta_{k,0})' \hat{P}_{2,i}(\beta_{k,0}) u_{2,i} \]
\[ + n^{-1} \sum_{i=1}^{n} \beta_\varphi(\beta_{k,0})' \partial^1 \hat{P}_{2,i}(\beta_{k,0})(\hat{\phi}_i - \phi_i) u_{2,i} \quad \text{(SC.134)} \]
\[ + n^{-1} \sum_{i=1}^{n} \beta_\varphi(\beta_{k,0})' \partial^2 \hat{P}_{2}(\varphi_1; \beta_{k,0})((\hat{\phi}_i - \phi_i)^2 u_{2,i} \]

which together with Assumption SC3(vi), SC393, SC396 and SC3127 implies that

\[ n^{-1} \sum_{i=1}^{n} \beta_\varphi(\beta_{k,0})' \hat{P}_{2,i}(\beta_{k,0}) u_{2,i} = n^{-1} \sum_{i=1}^{n} \beta_\varphi(\beta_{k,0})' \hat{P}_{2,i}(\beta_{k,0}) u_{2,i} + o_p(n^{-1/2}). \quad \text{(SC.135)} \]

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

\[ n^{-1} \sum_{i=1}^{n} \hat{P}_{2,i}(\beta_{k,0}) u_{2,i} = O_p(m_1^{1/2} n^{-1/2}) \quad \text{(SC.136)} \]

we deduce that

\[ n^{-1} \sum_{i=1}^{n} \beta_\varphi(\beta_{k,0})' \hat{P}_{2,i}(\beta_{k,0}) u_{2,i} = n^{-1} \sum_{i=1}^{n} \beta_\varphi(\beta_{k,0})' \hat{P}_{2,i}(\beta_{k,0}) u_{2,i} + o_p(n^{-1/2}) \]
\[ = n^{-1} \sum_{i=1}^{n} \varphi(\nu_{1,i}) u_{2,i} + o_p(n^{-1/2}) \]

where the first equality is by SC126, SC136 and Assumption SC3(vi), the second equality is by Assumptions SC1(i) and SC3(vi), SC88 and the Markov inequality. \textit{Q.E.D.}

**Lemma SC13.** \textit{Under Assumptions SC1, SC2 and SC3, we have}

\[ n^{-1} \sum_{i=1}^{n} g_1(\nu_{1,i})(\hat{\phi}(x_{1,i}) - \phi(x_{1,i})) \varsigma_{1,i} \]
\[ = n^{-1} \sum_{i=1}^{n} \eta_{1,i} g_1(\nu_{1,i})(\varsigma_{1,i} - \varsigma_{2,i}) - \mathbb{E}[l_{1,i} g_1(\nu_{1,i})(\varsigma_{1,i} - \varsigma_{2,i})](\hat{\beta}_t - \beta_{t,0}) + o_p(n^{-1/2}) \]

where \( \varsigma_{2,i} \equiv k_{2,i} - \mathbb{E}[k_{2,i}|x_{1,i}] \).

**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

\[ n^{-1} \sum_{i=1}^{n} g_{1,i}(\hat{\phi}(x_{1,i}) - \phi(x_{1,i}))\varsigma_{1,i} \]

\[ = (\hat{\beta}_\phi - \beta_{0,m_1})' n^{-1} \sum_{i=1}^{n} P_1(x_{1,i})g_{1,i}\varsigma_{1,i} + n^{-1} \sum_{i=1}^{n} g_{1,i}(\phi_{m_1}(x_{1,i}) - \phi(x_{1,i}))\varsigma_{1,i} \quad \text{(SC.137)} \]

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

\[ n^{-1} \sum_{i=1}^{n} g_{1,i}(\phi_{m_1}(x_{1,i}) - \phi(x_{1,i}))\varsigma_{1,i} = o_p(n^{-1/2}). \quad \text{(SC.138)} \]

By the definition of \( \varsigma_{1,i} \) and \( \varsigma_{2,i} \), we can write

\[ \varsigma_{1,i} = E[\varsigma_{1,i}|x_{1,i}] + \varsigma_{2,i}. \quad \text{(SC.139)} \]

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

\[ n^{-1} \sum_{i=1}^{n} P_1(x_{1,i})g_{1,i}\varsigma_{1,i} - E[P_1(x_{1,i})g_{1,i}\varsigma_{1,i}] = O_p(m_1^{1/2}n^{-1/2}) \]

which together with Assumption SC3(iv), SC55 and SC139 implies that

\[ (\hat{\beta}_\phi - \beta_{0,m_1})' n^{-1} \sum_{i=1}^{n} P_1(x_{1,i})g_{1,i}\varsigma_{1,i} \]

\[ = n^{-1} \sum_{i=1}^{n} \eta_{1,i}P_1(x_{1,i})'Q_{m_1}^{-1}E[P_1(x_{1,i})g_{1,i}(\varsigma_{1,i} - \varsigma_{2,i})] \]

\[ - (\hat{\beta}_l - \beta_{0,0})E[l_{1,i}P_1(x_{1,i})'Q_{m_1}^{-1}E[P_1(x_{1,i})g_{1,i}(\varsigma_{1,i} - \varsigma_{2,i})] + o_p(n^{-1/2}). \quad \text{(SC.140)} \]

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

\[ E \left[ n^{-1} \sum_{i=1}^{n} \eta_{1,i} \left[ P_1(x_{1,i})'Q_{m_1}^{-1}E[P_1(x_{1,i})g_{1,i}(\varsigma_{1,i} - \varsigma_{2,i})] - g_{1,i}(\varsigma_{1,i} - \varsigma_{2,i}) \right]^2 \right] \]

\[ \leq Cn^{-1}E \left[ P_1(x_{1,i})'Q_{1,m_1}^{-1}E[P_1(x_{1,i})g_{1,i}(\varsigma_{1,i} - \varsigma_{2,i})] - g_{1,i}(\varsigma_{1,i} - \varsigma_{2,i}) \right]^2 = o(n^{-1}) \]

which together with the Markov inequality implies that

\[ n^{-1} \sum_{i=1}^{n} \eta_{1,i}P_1(x_{1,i})'Q_{m_1}^{-1}E[P_1g_{1,i}(\varsigma_{1,i} - \varsigma_{2,i})] = n^{-1} \sum_{i=1}^{n} \eta_{1,i}g_{1,i}(\varsigma_{1,i} - \varsigma_{2,i}) + o_p(n^{-1/2}). \quad \text{(SC.141)} \]
By Hölder’s inequality, Assumptions [SC1](ii, v), [SC2](ii) and [SC3](iii)

\[
\begin{align*}
|E \left[ l_{1,i} P_1(x_{1,i})' \right] Q_{m_1}^{-1} E \left[ P_1(x_{1,i}) g_1,i(s_{1,i} - s_{2,i}) \right] &- E[l_{1,i} g_1,i(s_{1,i} - s_{2,i})]|^2 \\
= |E \left[ l_{1,i} (P_1(x_{1,i})' Q_{m_1}^{-1} E \left[ P_1(x_{1,i}) g_1,i(s_{1,i} - s_{2,i}) \right] - g_1,i(s_{1,i} - s_{2,i}) \right] |^2 \\
\leq E \left[ l_{1,i}^2 \right] E \left[ (P_1(x_{1,i})' Q_{m_1}^{-1} E \left[ P_1(x_{1,i}) g_1,i(s_{1,i} - s_{2,i}) \right] - g_1,i(s_{1,i} - s_{2,i}) \right]^2 \right] = o(1)
\end{align*}
\]

which combined with Assumption [SC1](iii) implies that

\[
(\hat{\beta}_l - \beta_{l,0}) E[l_{1,i} P_1(x_{1,i})' Q_{m_1}^{-1} E \left[ P_1(x_{1,i}) g_1,i(s_{1,i} - s_{2,i}) \right]] = (\hat{\beta}_l - \beta_{l,0}) E[l_{1,i} g_1,i(s_{1,i} - s_{2,i})] + o_p(n^{-1/2})
\]

(SC.142)

The claim of the lemma follows from (SC.137), (SC.138), (SC.140), (SC.141) and (SC.142). Q.E.D.

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

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

**Proof of Lemma SC14.** Using the decomposition in (SC.65), and applying the results in (SC.71), (SC.73) and (SC.77), we have

\[
\begin{align*}
\hat{\beta}_g(\hat{\beta}_k) - \tilde{\beta}_{g,m_2}(\hat{\beta}_k) &= (P_2(\hat{\beta}_k)' \hat{P}_2(\hat{\beta}_k))^{-1} \sum_{i=1}^n (\hat{P}_{2,i}(\hat{\beta}_k) - \tilde{P}_{2,i}(\tilde{\beta}_k)) u_{2,i}(\hat{\beta}_k) + O_p((m_1^{1/2} + m_2) n^{-1/2}).
\end{align*}
\]

(SC.143)

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

\[
\begin{align*}
&n^{-1} \sum_{i=1}^n b'(\hat{P}_{2,i}(\hat{\beta}_k) - \tilde{P}_{2,i}(\tilde{\beta}_k)) u_{2,i}(\hat{\beta}_k) \\
&= n^{-1} \sum_{i=1}^n b' \partial^1 \hat{P}_{2,i}(\hat{\beta}_k)(\hat{\phi}_i - \phi_i) u_{2,i}(\hat{\beta}_k) + n^{-1} \sum_{i=1}^n b' \partial^2 \hat{P}_2(\hat{\nu}_{1,i}(\hat{\beta}_k); \hat{\beta}_k)(\hat{\phi}_i - \phi_i)^2 u_{2,i}(\hat{\beta}_k)
\end{align*}
\]

(SC.144)

where $\hat{\nu}_{1,i}(\hat{\beta}_k)$ lies between $\nu_{1,i}(\hat{\beta}_k)$ and $\nu_{1,i}(\hat{\beta}_k)$. By (SC.68) and the compactness of $\Theta_k$,

\[
\sup_{\beta_k \in \Theta_k} |u_{2,i}(\beta_k)| \leq C + |y_{2,i}^*| + |k_{2,i}|
\]

(SC.145)

Using similar arguments in showing (SC.92), we have

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

36
which together with the Cauchy-Schwarz inequality, the triangle inequality, Assumptions \textbf{SC2} (vi) and \textbf{SC3} (iv), and (SC.145) implies that

\[
\left| n^{-1} \sum_{i=1}^{n} b' \partial^2 \hat{P}_{2,i}(\beta_k; \hat{\beta}_k)(\hat{\phi}_i - \phi_i)^2 u_{2,i}(\hat{\beta}_k) \right| \\
\leq \|b\| \|\xi_{2,m_2}n^{-1} \sum_{i=1}^{n} (\hat{\phi}_i - \phi_i)^2 (C + |y_{2,i}^2|) = \|b\| o_p(m_1^{-1/2}n^{-1/2}). \quad (SC.146)
\]

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

\[
n^{-1} \sum_{i=1}^{n} b' \partial^1 \tilde{P}_{2,i}(\hat{\beta}_k)(\hat{\phi}_i - \phi_i) u_{2,i} = n^{-1} \sum_{i=1}^{n} b' \partial^1 \tilde{P}_{2,i}(\hat{\beta}_k)(\hat{\phi}_i - \phi_i) u_{2,i} \\
- (\hat{\beta}_k - \beta_{k,0})n^{-1} \sum_{i=1}^{n} b' \partial^1 \tilde{P}_{2,i}(\hat{\beta}_k) k_{2,i}(\hat{\phi}_i - \phi_i) \\
- n^{-1} \sum_{i=1}^{n} b' \partial^1 \tilde{P}_{2,i}(\hat{\beta}_k)(\hat{\phi}_i - \phi_i)(g(\nu_{1,i}(\hat{\beta}_k), \hat{\beta}_k) - g(\nu_{1,i})). \quad (SC.147)
\]

By the Cauchy-Schwarz inequality, the triangle inequality, Assumption \textbf{SC2} (i, v) and Lemma \textbf{SC4}

\[
\left| n^{-1} \sum_{i=1}^{n} b' \partial^1 \tilde{P}_{2,i}(\hat{\beta}_k) k_{2,i}(\hat{\phi}_i - \phi_i) \right| \leq \|b\| o_p(\xi_{1,m_2}m_1^{-1/2}n^{-1/2}). \quad (SC.148)
\]

Similarly we can show that

\[
\left| n^{-1} \sum_{i=1}^{n} b' \partial^1 \tilde{P}_{2,i}(\hat{\beta}_k)(\hat{\phi}_i - \phi_i)(g(\nu_{1,i}(\hat{\beta}_k), \hat{\beta}_k) - g(\nu_{1,i})) \right| \\
\leq \|b\| |\hat{\beta}_k - \beta_{k,0}| o_p(\xi_{1,m_2}m_1^{-1/2}n^{-1/2}). \quad (SC.149)
\]

By the Cauchy-Schwarz inequality, the triangle inequality, Assumption \textbf{SC3} (iv), Lemma \textbf{SC35} Lemma \textbf{SC36} and \textbf{SC55},

\[
\left| n^{-1} \sum_{i=1}^{n} b' \partial^1 \tilde{P}_{2,i}(\hat{\beta}_k)(\hat{\phi}_i - \phi_i) u_{2,i} \right| \leq \|b\| o_p(m_2^{-1/2}m_1^{-1}n^{-1}) \leq \|b\| o_p((m_1^{-1/2} + m_2)n^{-1/2}) \quad (SC.150)
\]
Collecting the results in (SC.144), (SC.146), (SC.147), (SC.148), (SC.149) and (SC.150), we have

\[
\frac{1}{n}\sum_{i=1}^{n} (\hat{P}_{2,i}(\hat{\beta}_k) - \hat{P}_{2,i}(\hat{\beta}_k))u_{2,i}(\hat{\beta}_k) = (\hat{\beta}_k - \beta_k, 0)O_p(\xi_1, m_2 m_1^{1/2} n^{-1/2}) + O_p((m_1^{1/2} + m_2) n^{-1/2})
\]

which together with (SC.143) proves the claim of the lemma. \(Q.E.D.\)

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

\[
\frac{1}{n}\sum_{i=1}^{n} \left| \hat{g}_i(\hat{\beta}_k) - g(\nu_{1,i}, \hat{\beta}_k) \right|^2 = (\hat{\beta}_k - \beta_k, 0)^2 O_p(\xi_1, m_2 m_1 n^{-1}) + O_p((m_1 + m_2) n^{-1})
\]

where \(\hat{g}_i(\hat{\beta}_k) \equiv \hat{g}(\nu_{1,i}, \hat{\beta}_k), \hat{\beta}_k).\)

**Proof of Lemma SC15.** First note that by (SC.70),

\[
\frac{1}{n}\sum_{i=1}^{n} \left| g_{m_2}(\nu_{1,i}, \hat{\beta}_k) - g(\nu_{1,i}, \hat{\beta}_k) \right|^2 = O_p(m_1 n^{-1}) \quad \text{(SC.151)}
\]

where \(g_{m_2}(\nu_{1,i}, \hat{\beta}_k) \equiv \hat{P}_{2,i}(\hat{\beta}_k)' \hat{\beta}_{g,m_2}(\hat{\beta}_k).\) By Lemma SC14 and (SC.66)

\[
\frac{1}{n}\sum_{i=1}^{n} \left| \hat{P}_{2,i}(\hat{\beta}_k)' \hat{\beta}_g(\hat{\beta}_k) - g_{m_2}(\nu_{1,i}, \hat{\beta}_k) \right|^2 \\
\leq \lambda_{\max} \left( \frac{1}{n} \hat{P}_{2}(\hat{\beta}_k)' \hat{P}_{2}(\hat{\beta}_k) \right) \left| \hat{\beta}_g(\hat{\beta}_k) - \hat{\beta}_{g,m_2}(\hat{\beta}_k) \right|^2 \\
= (\hat{\beta}_k - \beta_k, 0)^2 O_p(\xi_1, m_2 m_1 n^{-1}) + O_p((m_1 + m_2) n^{-1})
\]

which together with (SC.151) finishes the proof. \(Q.E.D.\)

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

\[
\frac{1}{n}\sum_{i=1}^{n} \left| \hat{g}_{1,i}(\hat{\beta}_k) - g(\nu_{1,i}, \hat{\beta}_k) \right|^2 = o_p(1)
\]

where \(\hat{g}_{1,i}(\hat{\beta}_k) \equiv \partial \hat{P}_{2,i}(\hat{\beta}_k)' \hat{\beta}_g(\hat{\beta}_k).\)

**Proof of Lemma SC16.** First, we can use similar arguments for showing (SC.70) to get

\[
\frac{1}{n}\sum_{i=1}^{n} \left| \partial \hat{P}_{2,i}(\hat{\beta}_k)' \hat{\beta}_{g,m_2}(\hat{\beta}_k) - g(\nu_{1,i}, \hat{\beta}_k) \right|^2 = O_p(\xi_2, m_2 n^{-1}) \quad \text{(SC.152)}
\]
By Assumption SC2(v), Lemma SC14 and the consistency of \( \hat{\beta}_k \)

\[
\begin{align*}
n^{-1} \sum_{i=1}^{n} \left| \partial^{1} \hat{P}_{2,i}(\hat{\beta}_k)^{'} \hat{\beta}_g(\hat{\beta}_k) - \partial^{1} \hat{P}_{2,i}(\hat{\beta}_k)^{'} \hat{\beta}_{g,m_2}(\hat{\beta}_k) \right|^2 \\
\leq \xi_{1,m_2}^2 \left| \hat{\beta}_g(\hat{\beta}_k) - \hat{\beta}_{g,m_2}(\hat{\beta}_k) \right|^2 \\
= o_p(\xi_{1,m_2}^2 m_1 n^{-1}) + O_p\left(\xi_{1,m_2}^2 (m_1 + m_2) n^{-1}\right) \\
\text{(SC.153)}
\end{align*}
\]

which together with Assumption SC3(iv) and (SC.152) proves the claim of the lemma. \( Q.E.D. \)

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

\[
\begin{align*}
n^{-1} \sum_{i=1}^{n} k_{2,i}(k_{2,i} - k_{1,i} \hat{g}_{1,i}(\hat{\beta}_k)) &= \mathbb{E}[k_{2,i}(k_{2,i} - k_{1,i} g_1(\nu_{1,i}))] + o_p(1) \\
\text{(SC.154)}
\end{align*}
\]

and

\[
\begin{align*}
n^{-1} \sum_{i=1}^{n} l_{2,i} k_{1,i}(\hat{g}_{1,i}(\hat{\beta}_k) - \hat{g}_{1,i}(\beta_k,0)) &= o_p(1). \\
\text{(SC.155)}
\end{align*}
\]

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

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

where the third equality is by the Markov inequality. This proves the claim in (SC.154). Similarly, by Assumption SC2(ii), Lemma SC16 and the consistency of \( \hat{\beta}_k \), we have

\[
\begin{align*}
&n^{-1} \sum_{i=1}^{n} \left| \hat{g}_{1,i}(\hat{\beta}_k) - \hat{g}_{1,i}(\beta_k,0) \right|^2 \\
&\leq 2n^{-1} \sum_{i=1}^{n} \left| g_1(\nu_{1,i}; \hat{\beta}_k) - g_1(\nu_{1,i}; \beta_k,0) \right|^2 + o_p(1) \\
&\leq C(\hat{\beta}_k - \beta_k,0)^2 + o_p(1) = o_p(1). \\
\text{(SC.156)}
\end{align*}
\]

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.156) proves the claim in (SC.155). \( Q.E.D. \)
Lemma SC18. Under Assumptions [SC1, SC2 and SC3], we have

\[ n^{-1} \sum_{i=1}^{n} (\hat{g}_i(\hat{\beta}_k) - \hat{g}_i(\beta_{k,0}))(k_{2,i} - k_{1,i} \hat{g}_1,i(\beta_{k,0})) = -(\hat{\beta}_k - \beta_{k,0}) (E[(\gamma_{2,i} + \epsilon_{1,i}g_{1,i})(k_{2,i} - k_{1,i}g_{1,i})] + o_p(1)) + O_p((m_2 + m_1^{1/2})n^{-1/2}) \]

where \( g_{1,i} \equiv g_1(\nu_{1,i}), \epsilon_{1,i} \equiv k_{1,i} - E[k_{1,i}|\nu_{1,i}], \gamma_{2,i} \equiv E[k_{2,i}|\nu_{1,i}], \nu_{1,i} \) and \( g_1(\cdot) \) are defined in (SA.5) and (SB.13) respectively.

Proof of Lemma SC18. First note that

\[ n^{-1} \sum_{i=1}^{n} (\hat{g}_i(\hat{\beta}_k) - \hat{g}_i(\beta_{k,0}))(k_{2,i} - k_{1,i} \hat{g}_1,i(\beta_{k,0})) = -n^{-1} \sum_{i=1}^{n} k_{1,i}(\hat{g}_i(\hat{\beta}_k) - \hat{g}_i(\beta_{k,0}))(\hat{g}_1,i(\beta_{k,0}) - g_{1,i}) \]

\[ + n^{-1} \sum_{i=1}^{n} (\hat{g}_i(\hat{\beta}_k) - g(\nu_{1,i} - \hat{\nu}_{1,i}))(\hat{g}_1,i(\beta_{k,0}) - g_{1,i}) \]

\[ + n^{-1} \sum_{i=1}^{n} (g(\nu_{1,i} - \hat{\nu}_{1,i}))(k_{2,i} - k_{1,i}g_{1,i}). \quad (SC.157) \]

By the Cauchy-Schwarz inequality, Assumption SC3 iv), Lemma SC15 and (SC.101),

\[ n^{-1} \sum_{i=1}^{n} k_{1,i}(\hat{g}_i(\hat{\beta}_k) - \hat{g}_i(\beta_{k,0}))(\hat{g}_1,i(\beta_{k,0}) - g_{1,i}) = (\hat{\beta}_k - \beta_{k,0})o_p(1) + o_p(n^{-1/2}). \quad (SC.158) \]

Similarly, we can use the Cauchy-Schwarz inequality, Lemma SC15 and (SC.117) to get

\[ n^{-1} \sum_{i=1}^{n} (\hat{g}_i(\hat{\beta}_k) - g(\nu_{1,i} - \hat{\nu}_{1,i}))(\hat{g}_1,i(\beta_{k,0}) - g_{1,i}) \]

\[ = (\hat{\beta}_k - \beta_{k,0})o_p(1) + O_p((m_2 + m_1^{1/2})n^{-1/2}). \quad (SC.159) \]

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

\[ n^{-1} \sum_{i=1}^{n} (g(\nu_{1,i} - \hat{\nu}_{1,i})(\hat{\beta}_{k,0}))(k_{2,i} - k_{1,i}g_{1,i}) \]

\[ = (\hat{\beta}_k - \beta_{k,0})n^{-1} \sum_{i=1}^{n} \frac{\partial g(\nu_{1,i} - \hat{\nu}_{1,i})(\beta_{k,0})}{\partial \beta_k}(k_{2,i} - k_{1,i}g_{1,i}) + (\hat{\beta}_k - \beta_{k,0})o_p(1). \quad (SC.160) \]
Since
\[
\frac{\partial g(\nu_{1,i}(\beta_{k,0}); \beta_{k,0})}{\partial \beta_k} = \frac{\partial (\gamma_1(\nu_{1,i}(\beta_k)) - \beta_k \gamma_2(\nu_{1,i}(\beta_k)))}{\partial \beta_k} \bigg|_{\beta_{k}=\beta_{k,0}} = -\gamma_2(\nu_{1,i}) + \frac{\partial (\gamma_1(\nu_{1,i}(\beta_k)) - \beta_{k,0} \gamma_2(\nu_{1,i}(\beta_k)))}{\partial \beta_k} \bigg|_{\beta_{k}=\beta_{k,0}} = -\gamma_2(\nu_{1,i}) - g_{1,i}(k_{1,i} - \mathbb{E}[k_{1,i}|\nu_{1,i}])
\]

where the third equality is by the derivative formula in [Newey (1994) (Example 1 Continued, p.1358), by Assumptions SC1(i) and SC2(i, ii), and the Markov inequality, where the third equality is by the derivative formula in [Newey (1994) (Example 1 Continued, p.1358), by Assumptions SC1(i) and SC2(i, ii), and the Markov inequality, where
\[
\begin{align*}
n^{-1} \sum_{i=1}^{n} \frac{\partial g(\nu_{1,i}(\beta_{k,0}); \beta_{k,0})}{\partial \beta_k} (k_{2,i} - k_{1,i} g_{1,i}) \\
= -n^{-1} \sum_{i=1}^{n} (\gamma_{2,i} + \epsilon_{1,i} g_{1,i})(k_{2,i} - k_{1,i} g_{1,i}) = -\mathbb{E}[(\gamma_{2,i} + \epsilon_{1,i} g_{1,i})(k_{2,i} - k_{1,i} g_{1,i})] + o_p(n^{-1/2})
\end{align*}
\]

which together with (SC.160) implies that
\[
\begin{align*}
n^{-1} \sum_{i=1}^{n} (g(\nu_{1,i}(\beta_k); \beta_{k}) - g(\nu_{1,i}(\beta_{k,0}); \beta_{k,0}))(k_{2,i} - k_{1,i} g_{1,i}) \\
= - (\beta_k - \beta_{k,0}) (\mathbb{E}[(\gamma_{2,i} + \epsilon_{1,i} g_{1,i})(k_{2,i} - k_{1,i} g_{1,i})] + o_p(1)) + o_p(n^{-1/2}). \quad (SC.161)
\end{align*}
\]

The claim of the lemma follows from (SC.157), (SC.158), (SC.159) and (SC.161). \(Q.E.D.\)

**Lemma SC19.** Under Assumptions SC1, SC2 and SC3, we have
\[
\begin{align*}
n^{-1} \sum_{i=1}^{n} u_{2,i} k_{1,i} (\hat{g}_{1,i}(\hat{\beta}_k) - \hat{g}_{1,i}(\beta_{k,0})) &= (\hat{\beta}_k - \beta_{k,0}) o_p(1) + o_p(n^{-1/2}).
\end{align*}
\]

**Proof of Lemma SC19.** By the second order expansion,
\[
\begin{align*}
n^{-1} \sum_{i=1}^{n} u_{2,i} k_{1,i} \hat{g}_{1,i}(\hat{\beta}_k) &= n^{-1} \sum_{i=1}^{n} u_{2,i} k_{1,i} \partial^1 \hat{P}_{2,i}(\hat{\beta}_k)' \hat{\beta}_g(\hat{\beta}_k) \\
&= n^{-1} \sum_{i=1}^{n} u_{2,i} k_{1,i} \partial^1 \hat{P}_{2,i}(\hat{\beta}_k)' \hat{\beta}_g(\hat{\beta}_k) \\
&+ n^{-1} \sum_{i=1}^{n} u_{2,i} k_{1,i} (\hat{\phi}_i - \phi_i) \partial^2 \hat{P}_{2,i}(\hat{\beta}_k)' \hat{\beta}_g(\hat{\beta}_k) \\
&+ n^{-1} \sum_{i=1}^{n} u_{2,i} k_{1,i} (\hat{\phi}_i - \phi_i)^2 \partial^2 \hat{P}_{2,i}(\hat{\beta}_k)' \hat{\beta}_g(\hat{\beta}_k) \quad (SC.162)
\end{align*}
\]
where \( \tilde{\nu}_{1,i} \) is between \( \tilde{\nu}_{1,i}(\hat{\beta}_k) \) and \( \nu_{1,i}(\hat{\beta}_k) \). By (SC.59), \( \tilde{\nu}_{1,i} \in \Omega_\epsilon(\hat{\beta}_k) \) for any \( i = 1, \ldots, n \). By Assumption SC3(iv), Lemma SC14 and Lemma SC34

\[
 n^{-1} \sum_{i=1}^{n} u_{2,i}k_{1,i} \partial^1 \tilde{P}_{2,i}(\hat{\beta}_k)'(\hat{\beta}_g(\hat{\beta}_k) - \hat{\beta}_{g,m_2}(\hat{\beta}_k)) = (\hat{\beta}_k - \beta_{k,0})o_p(1) + o_p(n^{-1/2})
\]

which together with Assumption SC2(iii, vi) implies that

\[
 n^{-1} \sum_{i=1}^{n} u_{2,i}k_{1,i} \partial^1 \tilde{P}_{2,i}(\hat{\beta}_k)'\hat{\beta}_g(\hat{\beta}_k) = (\hat{\beta}_k - \beta_{k,0})o_p(1) + o_p(n^{-1/2}). \tag{SC.163}
\]

Using similar arguments for proving (SC.149) we can show that

\[
 \left| n^{-1} \sum_{i=1}^{n} u_{2,i}k_{1,i}(\hat{\phi}_i - \phi_i) \partial^2 \tilde{P}_{2,i}(\hat{\beta}_k)'b \right| \leq \|b\| O_p(m_2^2m_1n^{-1}) \tag{SC.164}
\]

for any \( b \in \mathbb{R}^{m_2} \). By the Cauchy-Schwarz inequality, Assumption SC3(iv), (SC.78) and (SC.164)

\[
 \left| n^{-1} \sum_{i=1}^{n} u_{2,i}k_{1,i}(\hat{\phi}_i - \phi_i) \partial^2 \tilde{P}_{2,i}(\hat{\beta}_k)'\hat{\beta}_g(\hat{\beta}_k) \right| \leq \|\hat{\beta}_g(\hat{\beta}_k)\| O_p(m_2^{7/2}m_1n^{-1}) = o_p(n^{-1/2}) \tag{SC.165}
\]

By the Cauchy-Schwarz inequality and the triangle inequality, Assumption SC2(v), (SC.59), (SC.78) and (SC.92)

\[
 \left| n^{-1} \sum_{i=1}^{n} u_{2,i}k_{1,i}(\hat{\phi}_i - \phi_i)^2 \partial^4 \tilde{P}_{2,i}(\tilde{\nu}_{1,i}; \hat{\beta}_k)'\hat{\beta}_g(\hat{\beta}_k) \right| \\
 \leq O_p(\xi_{3,m_2}) n^{-1} \sum_{i=1}^{n} |u_{2,i}|(\hat{\phi}_i - \phi_i)^2 = O_p(\xi_{3,m_2}m_1n^{-1}) = o_p(n^{-1/2}) \tag{SC.166}
\]

where the second equality is by Assumption SC3(iv). Combining the results in (SC.162), (SC.163), (SC.165) and (SC.166), we get

\[
 n^{-1} \sum_{i=1}^{n} u_{2,i}k_{1,i} \tilde{g}_{1,i}(\hat{\beta}_k) = n^{-1} \sum_{i=1}^{n} u_{2,i}k_{1,i}g_{1}(\nu_{1,i}(\hat{\beta}_k); \hat{\beta}_k) + (\hat{\beta}_k - \beta_{k,0})o_p(1) + o_p(n^{-1/2}). \tag{SC.167}
\]

Similarly, we can show that

\[
 n^{-1} \sum_{i=1}^{n} u_{2,i}k_{1,i} \tilde{g}_{1,i}(\beta_{k,0}) = n^{-1} \sum_{i=1}^{n} u_{2,i}k_{1,i}g_{1}(\nu_{1,i}(\beta_{k,0}); \beta_{k,0}) + o_p(n^{-1/2})
\]
which together with (SC.167) implies that

\[ n^{-1} \sum_{i=1}^{n} u_{2,i} k_{1,i}(\hat{g}_{1,i}(\hat{\beta}_k) - \hat{g}_{1,i}(\beta_{k,0})) \]

\[ = n^{-1} \sum_{i=1}^{n} u_{2,i} k_{1,i}(g_1(\nu_{1,i}(\hat{\beta}_k); \hat{\beta}_k) - g_1(\nu_{1,i}(\beta_{k,0}); \beta_{k,0})) \]

\[ + (\hat{\beta}_k - \beta_{k,0})o_p(1) + o_p(n^{-1/2}). \]  

(SC.168)

Therefore by Assumptions SC1(i), SC2(ii), (SC.108) and the consistency of \( \hat{\beta}_k \),

\[ n^{-1} \sum_{i=1}^{n} u_{2,i} k_{1,i}(g_1(\nu_{1,i}(\hat{\beta}_k); \hat{\beta}_k) - g_1(\nu_{1,i}(\beta_{k,0}); \beta_{k,0})) = (\hat{\beta}_k - \beta_{k,0})o_p(1) \]

which together with (SC.168) proves the claim of the lemma. Q.E.D.

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

\[ \hat{\beta}_k - \beta_{k,0} = O_p((m_1^{1/2} + m_2)n^{-1/2}). \]  

(SC.169)

Proof of Lemma SC20. Recall that \( \hat{J}_i(\beta_k) = \hat{\ell}_i(\beta_k) (k_{2,i} - k_{1,i} \hat{g}_1(\nu_{1,i}(\beta_k); \beta_k)) \) for any \( \beta_k \in \Theta_k \). The first order condition of \( \hat{\beta}_k \), i.e. (SB.12), can be written as

\[ n^{-1} \sum_{i=1}^{n} (\hat{J}_i(\beta_{k,0}) - \hat{J}_i(\hat{\beta}_k)) = n^{-1} \sum_{i=1}^{n} \hat{J}_i(\beta_{k,0}) \]  

(SC.170)

where by Lemma SB2 and (SB.41)

\[ n^{-1} \sum_{i=1}^{n} \hat{J}_i(\beta_{k,0}) = O_p(n^{-1/2}). \]  

(SC.171)

Using Assumption SC1(iii), Lemma SC17, Lemma SC18 and Lemma SC19 we can use the decomposition in (SB.38) to deduce that

\[ n^{-1} \sum_{i=1}^{n} (\hat{J}_i(\hat{\beta}_k) - \hat{J}_i(\beta_{k,0})) = -(\hat{\beta}_k - \beta_{k,0}) (E[\hat{\nu}_{1,i}^2] + o_p(1)) + O_p((m_1^{1/2} + m_2)n^{-1/2}). \]  

(SC.172)

The claim of the lemma follows from (SB.14), (SC.170), (SC.171) and (SC.172). Q.E.D.

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

\[ \hat{\beta}_{g,*} - \beta_{g,(\hat{\beta}_k,0)} = O_p(\xi_1 m_2 (m_1^{1/2})n^{-1/2}) = o_p(1) \]
where $\tilde{\beta}_{g,*} \equiv (B(\beta_{k,0}))^{-1}B(\tilde{\beta}_k)^\prime\tilde{\beta}_g(\tilde{\beta}_k)$.

**Proof of Lemma SC21** By the definition of $\tilde{\beta}_{g,*}$, we can write

$$\tilde{\beta}_{g,*} = (\hat{P}_2^\prime(\tilde{\beta}_k)^\prime\hat{P}_2(\tilde{\beta}_k))^{-1}\hat{P}_2^\prime(\tilde{\beta}_k)^\prime\hat{Y}_2(\tilde{\beta}_k)$$

where $\hat{P}_2^\prime(\beta_k) \equiv (\hat{P}_{2,1}(\beta_k), \ldots, \hat{P}_{2,n}(\beta_k))^\prime$ and $\hat{P}_{2,i}(\beta_k) \equiv B(\beta_{k,0})P_2(\nu_{1,i}(\beta_k))$. Therefore we have the following decomposition

$$\tilde{\beta}_{g,*} - \hat{\beta}_g(\beta_{k,0}) = \left[(\hat{P}_2^\prime(\tilde{\beta}_k)^\prime\hat{P}_2(\tilde{\beta}_k))^{-1} - (\hat{P}_2(\beta_{k,0}))^{-1}\hat{P}_2^\prime(\tilde{\beta}_k)^\prime\hat{Y}_2(\tilde{\beta}_k)\right] + (\hat{P}_2(\beta_{k,0}))^{-1}(\hat{P}_2^\prime(\tilde{\beta}_k) - \hat{P}_2(\beta_{k,0}))^\prime\hat{Y}_2^\prime(\beta_k)$$

SC.173

By the Markov inequality, Assumptions SC2(i) and SC2(i), and (SC.66),

$$\left\| (\hat{P}_2(\beta_{k,0}))^{-1}\hat{P}_2^\prime(\beta_{k,0})^\prime\hat{Y}_2(\beta_k) \right\|^2 \leq \frac{n^{-1}\sum_{i=1}^n k_{2,i}^2}{\lambda_{\min}(n^{-1}\hat{P}_2(\beta_{k,0})^\prime\hat{P}_2(\beta_{k,0}))} = O_p(1). \quad \text{SC.174}$$

Since $\hat{Y}_2^\prime(\tilde{\beta}_k) - \hat{Y}_2^\prime(\beta_{k,0}) = -\hat{\beta}_k - \beta_{k,0})K_2$, by Lemma SC20 and (SC.174) we get

$$\left((\hat{P}_2(\beta_{k,0}))^{-1}\hat{P}_2^\prime(\beta_{k,0})^\prime(\hat{Y}_2(\tilde{\beta}_k) - \hat{Y}_2(\beta_{k,0})) = O_p((m_2 + m_1^{1/2})n^{-1/2}). \quad \text{SC.175}$$

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

$$b'(\hat{P}_{2,i}(\tilde{\beta}_k) - \hat{P}_{2,i}(\beta_{k,0})) = -b'(1)\hat{P}_2(\nu_{1,i}(\tilde{\beta}_k); \beta_{k,0}) = k_{1,i}(\tilde{\beta}_k - \beta_{k,0}) \quad \text{SC.176}$$

where $\hat{\beta}_k$ lies between $\tilde{\beta}_k$ and $\beta_{k,0}$. By Assumption SC2(iv), Lemma SC4 and Lemma SC20, $\nu_{1,i}(\tilde{\beta}_k) \in \Omega_\varepsilon(\beta_{k,0})$ for any $i = 1, \ldots, n$ wpa1. Therefore by the Cauchy-Schwarz inequality, Assumption SC2(v) and (SC.176)

$$\left| b'(\hat{P}_{2,i}(\tilde{\beta}_k) - \hat{P}_{2,i}(\beta_{k,0})) \right| \leq \|b\| \varepsilon_{1,m_2} \left| k_{1,i}(\tilde{\beta}_k - \beta_{k,0}) \right|$$

wpa1. Therefore we have wpa1,

$$b'(\hat{P}_{2}(\tilde{\beta}_k) - \hat{P}_2(\beta_{k,0}))^\prime(\hat{P}_2(\tilde{\beta}_k) - \hat{P}_2(\beta_{k,0}))b = \sum_{i=1}^n (b'(\hat{P}_{2,i}(\tilde{\beta}_k) - \hat{P}_{2,i}(\beta_{k,0})))^2 \leq \|b\|^2 \varepsilon_{1,m_2}^2(\hat{\beta}_k - \beta_{k,0})^2 \sum_{i=1}^n k_{1,i}^2$$
which together with Lemma [SC20] implies that

\[ \| \hat{P}_2(\hat{\beta}_k) - \hat{P}_2(\beta_{k,0}) \|_S = |\hat{\beta}_k - \beta_{k,0}|O_p(\xi_{1,m_2}n^{1/2}) = O_p(\xi_{1,m_2}(m_2 + m_1^{1/2})n^{1-2/2}). \]  

(SC.177)

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

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

which together with the Markov inequality, Assumptions [SC1(i)] and [SC2(i)], and the compactness of \( \Theta_k \) implies that

\[ \sup_{\beta_k \in \Theta_k} n^{-1} \sum_{i=1}^n (y_{2,i}^*(\beta_k))^2 = O_p(1). \]  

(SC.178)

By the Cauchy-Schwarz inequality, (SC.66), (SC.177) and (SC.178),

\[ \left\| (\hat{P}_2(\beta_{k,0})'\hat{P}_2(\beta_{k,0}))^{-1}(\hat{P}_2(\beta_k) - \hat{P}_2(\beta_{k,0}))'\hat{Y}_2^*(\beta_k) \right\| \]

\[ \leq \frac{n^{-1}\|\hat{P}_2(\beta_k) - \hat{P}_2(\beta_{k,0})\|_S \left\| \hat{Y}_2^*(\beta_k) \right\|}{\lambda_{\min}(n^{-1}\hat{P}_2(\beta_{k,0})'\hat{P}_2(\beta_{k,0}))} = O_p(\xi_{1,m_2}(m_2 + m_1^{1/2})n^{1-2/2}). \]  

(SC.179)

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

\[ \left[ (\hat{P}_2(\beta_k)'\hat{P}_2(\hat{\beta}_k))^{-1} - (\hat{P}_2(\beta_{k,0})'\hat{P}_2(\beta_{k,0}))^{-1} \right] \hat{P}_2(\beta_k) \hat{Y}_2(\hat{\beta}_k) \]

\[ = (\hat{P}_2(\beta_{k,0})'\hat{P}_2(\beta_{k,0}))^{-1}(\hat{P}_2(\beta_k) - \hat{P}_2(\beta_{k,0}))'\hat{P}_2(\beta_{k,0})\hat{\beta}_g(\hat{\beta}_k) \]

\[ + (\hat{P}_2(\beta_{k,0})'\hat{P}_2(\beta_{k,0}))^{-1}\hat{P}_2(\beta_{k,0})'(\hat{P}_2(\beta_k) - \hat{P}_2(\beta_{k,0})) \hat{\beta}_{g,*}. \]  

(SC.180)

By the Cauchy-Schwarz inequality, (SC.66), (SC.78) and (SC.177),

\[ \left\| (\hat{P}_2(\beta_{k,0})'\hat{P}_2(\beta_{k,0}))^{-1}(\hat{P}_2(\beta_k) - \hat{P}_2(\beta_{k,0}))'\hat{P}_2(\beta_{k,0})\hat{\beta}_g(\hat{\beta}_k) \right\| \]

\[ \leq \frac{n^{-1}\|\hat{P}_2(\beta_k) - \hat{P}_2(\beta_{k,0})\|_S \left\| \hat{P}_2(\beta_{k,0})\hat{\beta}_g(\hat{\beta}_k) \right\|}{\lambda_{\min}(n^{-1}\hat{P}_2(\beta_{k,0})'\hat{P}_2(\beta_{k,0}))} = O_p(\xi_{1,m_2}(m_2 + m_1^{1/2})n^{1-2/2}). \]  

(SC.181)

By the definition of \( \hat{\beta}_{g,*} \), and the mean value expansion

\[ \left\| (\hat{P}_2(\beta_{k,0}) - \hat{P}_2^*(\beta_k))\hat{\beta}_{g,*} \right\|^2 = \sum_{i=1}^n (\hat{\beta}_g(\hat{\beta}_k))'(\hat{P}_2(\nu_{1,i}(\beta_k,0); \beta_k) - \hat{P}_2(\nu_{1,i}(\beta_k; \beta_k)))^2 \]

\[ = (\hat{\beta}_k - \beta_{k,0})^2 \sum_{i=1}^n k_{1,i}^2(\beta_k)'\partial^1 \hat{P}_2(\nu_{1,i}(\beta_k; \beta_k))^2 \]  

(SC.182)
where \( \tilde{\beta}_k \) lies between \( \hat{\beta}_k \) and \( \beta_{k,0} \). By Assumption [SC3 iv], Lemma [SC4] and Lemma [SC20] \( \hat{\nu}_{1,i}(\tilde{\beta}_k) \in \Omega_{\tilde{\xi}_n}(\tilde{\beta}_k) \) for any \( i = 1, \ldots, n \) w.p.a.1. By the Cauchy-Schwarz inequality, Assumption [SC2 v], Lemma [SC20] (SC.78) and (SC.182)

\[
n^{-1/2}(\dot{P}_2(\hat{\beta}_k) - \dot{P}^*_2(\hat{\beta}_k))\tilde{\beta}_{g,*} = (\hat{\beta}_k - \beta_{k,0})O_p(\xi_{1,m_2}) = O_p(\xi_{1,m_2}(m_2 + m_1^{1/2})n^{-1/2}) \quad \text{(SC.183)}
\]

which together with (SC.66) implies that

\[
(\dot{P}_2(\beta_{k,0})'\dot{P}_2(\beta_{k,0}))^{-1}\dot{P}_2(\beta_{k,0})'(\dot{P}_2(\beta_{k,0}) - \dot{P}^*_2(\beta_{k}))\tilde{\beta}_{g,*} = O_p(\xi_{1,m_2}(m_2 + m_1^{1/2})n^{-1/2}). \quad \text{(SC.184)}
\]

Combining the results in (SC.180), (SC.181) and (SC.184) we get

\[
[(\dot{P}^*_2(\hat{\beta}_k)\dot{P}^*_2(\hat{\beta}_k)')^{-1} - (\dot{P}_2(\beta_{k,0})'\dot{P}_2(\beta_{k,0}))^{-1}]\dot{P}^*_2(\hat{\beta}_k)'\dot{Y}_2(\hat{\beta}_k) = O_p(\xi_{1,m_2}(m_2 + m_1^{1/2})n^{-1/2})
\]

which together with Assumption [SC3 iv], (SC.173), (SC.175) and (SC.179) proves the lemma. \( Q.E.D. \)

**Lemma SC22.** Let \( U_2 = (u_{2,1}, \ldots, u_{2,n})' \), \( \hat{G}_n = (\hat{g}(\hat{\nu}_{1,1}(\hat{\beta}_k); \hat{\beta}_k), \ldots, \hat{g}(\hat{\nu}_{1,n}(\hat{\beta}_k); \hat{\beta}_k))' \) and \( G_n = (g(\nu_{1,1}), \ldots, g(\nu_{1,n}))' \). Then under Assumptions [SC7], [SC2] and [SC3] we have

(i) \( n^{-1}U_2'(\dot{P}^*_2(\hat{\beta}_k) - \dot{P}_2(\beta_{k,0}))\tilde{\beta}_{\varphi}(\beta_{k,0}) = (\hat{\beta}_k - \beta_{k,0})o_p(1) + o_p(n^{-1/2}) \)

(ii) \( n^{-1}L_2(\dot{P}^*_2(\hat{\beta}_k) - \dot{P}_2(\beta_{k,0}))\tilde{\beta}_{\varphi}(\beta_{k,0}) = o_p(1) \)

(iii) \( n^{-1}K_2(\dot{P}^*_2(\hat{\beta}_k) - \dot{P}_2(\beta_{k,0}))\tilde{\beta}_{\varphi}(\beta_{k,0}) = o_p(1) \)

(iv) \( n^{-1}(\hat{G}_n - G_n)'(\dot{P}^*_2(\hat{\beta}_k) - \dot{P}_2(\beta_{k,0}))\tilde{\beta}_{\varphi}(\beta_{k,0}) = (\hat{\beta}_k - \beta_{k,0})o_p(1) \)

**Proof of Lemma SC22** (i) First note that

\[
n^{-1}U_2'(\dot{P}^*_2(\hat{\beta}_k) - \dot{P}_2(\beta_{k,0}))\tilde{\beta}_{\varphi}(\beta_{k,0})
= n^{-1} \sum_{i=1}^{n} u_{2,i} (\dot{P}_2(\hat{\nu}_{1,i}(\hat{\beta}_k); \beta_{k,0}) - \dot{P}_2(\nu_{1,i}(\hat{\beta}_k); \beta_{k,0}))\tilde{\beta}_{\varphi}(\beta_{k,0})
- n^{-1} \sum_{i=1}^{n} u_{2,i} (\dot{P}_2(\hat{\nu}_{1,i}(\beta_{k,0}); \beta_{k,0}) - \dot{P}_2(\nu_{1,i}(\beta_{k,0}); \beta_{k,0}))\tilde{\beta}_{\varphi}(\beta_{k,0})
+ n^{-1} \sum_{i=1}^{n} u_{2,i} (\dot{P}_2(\nu_{1,i}(\beta_{k,0}); \beta_{k,0}) - \dot{P}_2(\nu_{1,i}(\hat{\beta}_k); \beta_{k,0}))\tilde{\beta}_{\varphi}(\beta_{k,0}). \quad \text{(SC.185)}
\]
By the second order expansion,

\[
\begin{align*}
&n^{-1} \sum_{i=1}^{n} u_{2,i} \left( \tilde{P}_2(\nu_{1,i}(\hat{\beta}_k); \beta_{k,0}) - \tilde{P}_2(\nu_{1,i}(\hat{\beta}_k); \beta_{k,0})' \hat{\beta}_\varphi(\beta_{k,0}) \right) \\
&= n^{-1} \sum_{i=1}^{n} u_{2,i} (\hat{\phi}_i - \phi_i) \partial^1 \tilde{P}_2(\nu_{1,i}(\hat{\beta}_k); \beta_{k,0})' \hat{\beta}_\varphi(\beta_{k,0}) \\
&\quad + n^{-1} \sum_{i=1}^{n} u_{2,i} (\hat{\phi}_i - \phi_i)^2 \partial^2 \tilde{P}_2(\nu_{1,i}(\hat{\beta}_k); \beta_{k,0})' \hat{\beta}_\varphi(\beta_{k,0})
\end{align*}
\]

(SC.186)

where \( \tilde{\nu}_{1,i}(\hat{\beta}_k) \) is between \( \nu_{1,i}(\hat{\beta}_k) \) and \( \nu_{1,i}(\hat{\beta}_k) \). By Assumption SC3(iv), Lemma SC4 and Lemma SC20, both \( \nu_{1,i}(\hat{\beta}_k) \) and \( \nu_{1,i}(\hat{\beta}_k) \) are in \( \Omega_\varepsilon(\beta_{k,0}) \) for any \( i = 1, \ldots, n \). Therefore \( \tilde{\nu}_{1,i}(\hat{\beta}_k) \in \Omega_\varepsilon(\beta_{k,0}) \) for any \( i = 1, \ldots, n \).

By the triangle inequality, the Cauchy-Schwarz inequality, Assumptions SC2(v) and SC3(iv), (SC.92) and (SC.127)

\[
\left| n^{-1} \sum_{i=1}^{n} u_{2,i} (\hat{\phi}_i - \phi_i)^2 \partial^2 \tilde{P}_2(\nu_{1,i}(\hat{\beta}_k); \beta_{k,0})' \hat{\beta}_\varphi(\beta_{k,0}) \right| \\
\leq O_p(\xi_{2,m_2}) n^{-1} \sum_{i=1}^{n} |u_{2,i}| (\hat{\phi}_i - \phi_i)^2 = O_p(\xi_{2,m_2} m_1 n^{-1}) = o_p(n^{-1/2}). \tag{SC.187}
\]

Using similar arguments for proving (SC.165), we can show that

\[
\begin{align*}
&n^{-1} \sum_{i=1}^{n} u_{2,i} (\hat{\phi}_i - \phi_i) \partial^1 \tilde{P}_2(\nu_{1,i}(\hat{\beta}_k); \beta_{k,0})' \hat{\beta}_\varphi(\beta_{k,0}) = o_p(n^{-1/2})
\end{align*}
\]

which together with (SC.186) and (SC.187) implies that

\[
\begin{align*}
&n^{-1} \sum_{i=1}^{n} u_{2,i} \left( \tilde{P}_2(\nu_{1,i}(\hat{\beta}_k); \beta_{k,0}) - \tilde{P}_2(\nu_{1,i}(\hat{\beta}_k); \beta_{k,0})' \hat{\beta}_\varphi(\beta_{k,0}) \right) \\
&\quad - \tilde{P}_2(\nu_{1,i}(\hat{\beta}_k); \beta_{k,0})' \hat{\beta}_\varphi(\beta_{k,0}) = o_p(n^{-1/2}). \tag{SC.188}
\end{align*}
\]

Similarly, we can show that

\[
\begin{align*}
\begin{align*}
&n^{-1} \sum_{i=1}^{n} u_{2,i} \left( \tilde{P}_2(\nu_{1,i}(\hat{\beta}_k); \beta_{k,0}) - \tilde{P}_2(\nu_{1,i}(\hat{\beta}_k); \beta_{k,0})' \hat{\beta}_\varphi(\beta_{k,0}) \right) \\
&\quad - \tilde{P}_2(\nu_{1,i}(\hat{\beta}_k); \beta_{k,0})' \hat{\beta}_\varphi(\beta_{k,0}) = o_p(n^{-1/2}). \tag{SC.189}
\end{align*}
\]
By the third order expansion,

\[
    n^{-1} \sum_{i=1}^{n} u_{2,i}(\tilde{P}_2(\nu_{1,i}(\hat{\beta}_k); \beta_{k,0}) - \tilde{P}_2(\nu_{1,i}(\beta_{k,0}); \beta_{k,0}))' \hat{\beta}_\varphi(\beta_{k,0})
\]

\[
    = (\hat{\beta}_k - \beta_{k,0}) n^{-1} \sum_{i=1}^{n} u_{2,i} \partial^1 \tilde{P}_2, i(\beta_{k,0})' \hat{\beta}_\varphi(\beta_{k,0})
\]

\[
    + (\hat{\beta}_k - \beta_{k,0})^2 n^{-1} \sum_{i=1}^{n} u_{2,i} \partial^2 \tilde{P}_2, i(\beta_{k,0})' \hat{\beta}_\varphi(\beta_{k,0})
\]

\[
    + (\hat{\beta}_k - \beta_{k,0})^3 n^{-1} \sum_{i=1}^{n} u_{2,i} \partial^3 \tilde{P}_2(\nu_{1,i}(\hat{\beta}_k); \beta_{k,0})' \hat{\beta}_\varphi(\beta_{k,0})
\]

(\text{SC.190})

where \( \hat{\beta}_k \) is between \( \hat{\beta}_k \) and \( \beta_{k,0} \). By Lemma \text{SC20} and Assumption \text{SC3 iv), } \nu_{1,i}(\hat{\beta}_k) \in \Omega_{\epsilon}(\beta_{k,0}) \text{ for any } i = 1, \ldots, n \text{ wpa1. Therefore by the triangle inequality and the Cauchy-Schwarz inequality, Assumptions \text{SC1 i) and SC2 v), Lemma \text{SC20}, (SC.88) and (SC.127), we can show that}

\[
    (\hat{\beta}_k - \beta_{k,0}) n^{-1} \sum_{i=1}^{n} u_{2,i} \partial^1 \tilde{P}_2, i(\beta_{k,0})' \hat{\beta}_\varphi(\beta_{k,0}) = (\hat{\beta}_k - \beta_{k,0}) \omega_p(1)
\]

(\text{SC.192})

and

\[
    (\hat{\beta}_k - \beta_{k,0})^2 n^{-1} \sum_{i=1}^{n} u_{2,i} \partial^2 \tilde{P}_2, i(\beta_{k,0})' \hat{\beta}_\varphi(\beta_{k,0}) = (\hat{\beta}_k - \beta_{k,0}) \omega_p(1).
\]

(\text{SC.193})

Collecting the results in (SC.190), (SC.191), (SC.192) and (SC.193), we obtain

\[
    n^{-1} \sum_{i=1}^{n} u_{2,i}(\tilde{P}_2(\nu_{1,i}(\hat{\beta}_k); \beta_{k,0}) - \tilde{P}_2(\nu_{1,i}(\beta_{k,0}); \beta_{k,0}))' \hat{\beta}_\varphi(\beta_{k,0}) = (\hat{\beta}_k - \beta_{k,0}) \omega_p(1)
\]

which together with (SC.185), (SC.188) and (SC.189) finishes the proof.

(ii) By the mean value expansion,

\[
    n^{-1} L_2(\tilde{P}_2(\beta_{k,0})) \beta_\varphi(\beta_{k,0}) = (\hat{\beta}_k - \beta_{k,0}) n^{-1} \sum_{i=1}^{n} \xi_2, i \partial^1 \tilde{P}_2(\nu_{1,i}(\hat{\beta}_k); \beta_{k,0})' \hat{\beta}_\varphi(\beta_{k,0})
\]

(\text{SC.194})
where \( \tilde{\beta}_k \) is between \( \hat{\beta}_k \) and \( \beta_{k,0} \). By Assumption [SC3 iv], Lemma [SC4] and Lemma [SC20], \( \tilde{\nu}_{1,i}(\tilde{\beta}_k) \in \Omega_\epsilon(\beta_{k,0}) \) for any \( i = 1, \ldots, n \) wpa1. By the triangle inequality, the Cauchy-Schwarz inequality, (SC.72) and (SC.127),

\[
n^{-1} \sum_{i=1}^{n} l_i \partial^1 \hat{P}_2(\nu_{1,i}(\tilde{\beta}_k); \beta_{k,0})' \hat{\beta}_\varphi(\beta_{k,0}) = O_p(\xi_{1,m_2})
\]

which together with Assumption [SC3 iv] and Lemma [SC20] finishes the proof.

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

(iv) By the mean value expansion,

\[
n^{-1}(\mathbf{G}_n - \mathbf{G}_n)'(\hat{\mathbf{P}}_2(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\beta_{k,0})) \hat{\beta}_\varphi(\beta_{k,0})
\]

\[
= n^{-1} \sum_{i=1}^{n} (\tilde{g}_i(\tilde{\beta}_k) - g(\nu_{1,i}))(\hat{P}_2(\nu_{1,i}(\tilde{\beta}_k); \beta_{k,0}) - \hat{P}_2(\nu_{1,i}(\beta_{k,0}); \beta_{k,0})' \hat{\beta}_\varphi(\beta_{k,0})
\]

\[
= -(\hat{\beta}_k - \beta_{k,0}) n^{-1} \sum_{i=1}^{n} (\tilde{g}_i(\tilde{\beta}_k) - g(\nu_{1,i})) \partial^1 \hat{P}_2(\nu_{1,i}(\tilde{\beta}_k); \beta_{k,0})' \hat{\beta}_\varphi(\beta_{k,0})
\]

(SC.195)

where \( \tilde{\beta}_k \) is between \( \hat{\beta}_k \) and \( \beta_{k,0} \). By Assumption [SC3 iv], Lemma [SC4] and Lemma [SC20], \( \tilde{\nu}_{1,i}(\tilde{\beta}_k) \in \Omega_\epsilon(\beta_{k,0}) \) for any \( i = 1, \ldots, n \) wpa1. By Assumptions [SC2 v] and [SC3 iv], Lemma [SC15], Lemma [SC20] and (SC.127), we get

\[
n^{-1} \sum_{i=1}^{n} (\tilde{g}_i(\tilde{\beta}_k) - g(\nu_{1,i})) \partial^1 \hat{P}_2(\nu_{1,i}(\tilde{\beta}_k); \beta_{k,0})' \hat{\beta}_\varphi(\beta_{k,0}) = o_p(1)
\]

which together with (SC.195) finishes the proof. \( Q.E.D. \)

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

\[
n^{-1} \sum_{i=1}^{n} (\tilde{g}(\nu_{1,i}(\tilde{\beta}_k); \beta_{k,0}) - g(\nu_{1,i}(\beta_{k,0}); \beta_{k,0}))(k_{2,i} - k_{1,i}\tilde{g}_1(\tilde{\beta}_k))
\]

\[
= -(\hat{\beta}_k - \beta_{k,0}) [\mathbb{E}[k_{1,i}\nu_{1,i},s_{1,i}] + \mathbb{E}[k_{2,i}(\gamma_{2,i} - \gamma_{1,i}\nu_{1,i})] + o_p(1)] + o_p(n^{-1/2})
\]

where \( a_{j,i} \equiv \mathbb{E}[k_{j,i} | \nu_{1,i}] \) and \( g_{1,i} \equiv g_1(\nu_{1,i}). \)

**Proof of Lemma SC23.** In view of (SC.158), to prove the lemma it is sufficient to show that

\[
n^{-1} \sum_{i=1}^{n} (\tilde{g}(\nu_{1,i}(\tilde{\beta}_k); \beta_{k,0}) - g(\nu_{1,i}(\beta_{k,0}); \beta_{k,0}))(k_{2,i} - k_{1,i}\tilde{g}_1(\tilde{\beta}_k))
\]

\[
= -(\hat{\beta}_k - \beta_{k,0}) [\mathbb{E}[k_{1,i}\nu_{1,i},s_{1,i}] + \mathbb{E}[k_{2,i}(\gamma_{2,i} - \gamma_{1,i}\nu_{1,i})] + o_p(1)] + o_p(n^{-1/2}). \quad (SC.196)
\]
By the definition of \( \hat{g}(\hat{\nu}_{1,i}(\beta_k); \beta_k) \),
\[
\hat{g}(\hat{\nu}_{1,i}(\beta_k); \beta_k) = \hat{P}_{2,i}(\beta_k)' \beta_{g}(\beta_k) = \hat{P}_{2,i}^\star(\beta_k)(\hat{P}_{2}(\beta_k)' \hat{P}_{2}(\beta_k))^{-1} \hat{P}_{2}(\beta_k)' \hat{Y}_{2}(\beta_k)
\]
where \( \hat{P}_{2}^\star(\beta_k) \equiv (\hat{P}_{2,1}(\beta_k), \ldots, \hat{P}_{2,n}(\beta_k))' \) and \( \hat{P}_{2,i}(\beta_k) \equiv B(\beta_{k,0})P_{2}(\hat{\nu}_{1,i}(\beta_k)) \). Therefore we obtain the following decomposition
\[
\hat{g}(\hat{\nu}_{1,i}(\hat{\beta}_k); \hat{\beta}_k) - \hat{g}(\hat{\nu}_{1,i}(\beta_k); \beta_k) = (\hat{P}_{2,i}^\star(\beta_k) - \hat{P}_{2,i}(\beta_k,0))' (\hat{P}_{2}(\beta_k)' \hat{P}_{2}(\beta_k))^{-1} \hat{P}_{2}(\beta_k)' \hat{Y}_{2}(\beta_k)
\]
\[
+ \hat{P}_{2,i}(\beta_k,0)' \left[ (\hat{P}_{2}(\beta_k)' \hat{P}_{2}(\beta_k))^{-1} - (\hat{P}_{2}(\beta_k)' \hat{P}_{2}(\beta_k,0))^{-1} \right] \hat{P}_{2}(\beta_k)' \hat{Y}_{2}(\beta_k)
\]
\[
+ \hat{P}_{2,i}(\beta_k,0)' (\hat{P}_{2}(\beta_k,0)' \hat{P}_{2}(\beta_k,0))^{-1} (\hat{P}_{2}(\beta_k) - \hat{P}_{2}(\beta_k,0))' \hat{Y}_{2}(\beta_k) + \hat{P}_{2,i}(\beta_k,0)' (\hat{P}_{2}(\beta_k) - \hat{P}_{2}(\beta_k,0))'(\hat{Y}_{2}(\beta_k) - \hat{Y}_{2}(\beta_k)).
\]

(SC.197)

The proof is divided into 4 steps. The claim in (SC.196) follows from the results in (SC.198), (SC.212), (SC.225) and (SC.227).

**Step 1.** In this step, we show that
\[
n^{-1} \sum_{i=1}^{n} \beta_{g,i}(\hat{P}_{2,i}^\star(\beta_k) - \hat{P}_{2,i}(\beta_k,0))(k_{2,i} - k_{1,i}g_{1,i}) = -(\hat{\beta}_k - \beta_k) (E[k_{1,i}g_{1,i}(k_{2,i} - k_{1,i}g_{1,i})] + o_p(1)) + o_p(n^{-1/2}).
\]

(SC.198)

where \( \beta_{g,i} \equiv (\hat{P}_{2}(\beta_k)' \hat{P}_{2}(\beta_k))^{-1} \hat{P}_{2}(\beta_k)' \hat{Y}_{2}(\beta_k) \).

By Lemma SC.21 and (SC.78),
\[
\|\beta_{g,i}\| = O_p(1).
\]

(SC.199)

By the second order expansion,
\[
\beta_{g,i}' \left( \hat{P}_{2,i}^\star(\beta_k) - \hat{P}_{2,i}(\beta_k,0) - \partial^1 \hat{P}_{2,i}(\beta_k,0)(\hat{\nu}_{1,i}(\beta_k) - \nu_{1,i}(\beta_k)) \right)
\]
\[
= \beta_{g,i}' \partial^2 \hat{P}_{2}(\hat{\nu}_{1,i}; \beta_{k,0})(\hat{\nu}_{1,i}(\hat{\beta}_k) - \nu_{1,i}(\beta_k))^2
\]

(SC.200)

where \( \hat{\nu}_{1,i} \) lies between \( \hat{\nu}_{1,i}(\beta_k) \) and \( \nu_{1,i}(\beta_k) \). By Assumption SC.3 iv, Lemma SC.4 and Lemma SC.20, \( \hat{\nu}_{1,i} \in \Omega_k(\beta_{k,0}) \) for any \( i = 1, \ldots, n \) wpa1. Since \( \hat{\nu}_{1,i}(\beta_k) - \nu_{1,i}(\beta_k) = (\hat{\phi}_1 - \phi_i) - k_{1,i}(\hat{\beta}_k - \beta_k,0) \),
by Assumption [SC2-v], [SC.199] and (SC.200) we have

\[
\begin{align*}
& n^{-1} \sum_{i=1}^{n} \beta_{g,i} \left( - \hat{P}_{2,i}(\beta_k) - \check{P}_{2,i}(\beta_{k,0}) \right) (k_{2,i} - k_{1,i}g_{1,i}) \\
& \leq O_p(\varepsilon_{2,m_2}) \left( n^{-1} \sum_{i=1}^{n} (\phi_i - \phi_i)^2 + (\beta_k - \beta_{k,0})^2 n^{-1} \sum_{i=1}^{n} k_{1,i}^2 \right) \\
& = (\beta_k - \beta_{k,0}) O_p(\varepsilon_{2,m_2}(m_1^{1/2} + m_2)n^{-1/2}) + O_p(\varepsilon_{2,m_2}m_1n^{-1}) \\
& = (\beta_k - \beta_{k,0}) o_p(1) + o_p(n^{-1/2})
\end{align*}
\] (SC.201)

where the first equality is by Lemma [SC4] and Lemma [SC20] and the second equality is by Assumption [SC3-iv]. Similarly, we can show that

\[
\begin{align*}
& n^{-1} \sum_{i=1}^{n} \beta_{g,i} \left( - \hat{P}_{2,i}(\beta_k) - \check{P}_{2,i}(\beta_{k,0}) \right) (k_{2,i} - k_{1,i}g_{1,i}) \\
& = -(\beta_k - \beta_{k,0}) n^{-1} \sum_{i=1}^{n} \beta_{g,i} \partial^1 \check{P}_{2,i}(\beta_{k,0}) k_{1,i}(k_{2,i} - k_{1,i}g_{1,i}) + (\beta_k - \beta_{k,0}) o_p(1) + o_p(n^{-1/2})
\end{align*}
\] (SC.202)

Since \( \nu_{1,i}(\hat{\beta}_k) - \nu_{1,i}(\beta_{k,0}) = -k_{1,i}(\hat{\beta}_k - \beta_{k,0}) \), using (SC.201) and (SC.202) we get

\[
\begin{align*}
& n^{-1} \sum_{i=1}^{n} \beta_{g,i} \left( - \hat{P}_{2,i}(\beta_k) - \check{P}_{2,i}(\beta_{k,0}) \right) (k_{2,i} - k_{1,i}g_{1,i}) \\
& \quad = -(\beta_k - \beta_{k,0}) n^{-1} \sum_{i=1}^{n} \beta_{g,i} \partial^1 \check{P}_{2,i}(\beta_{k,0}) k_{1,i}(k_{2,i} - k_{1,i}g_{1,i}) + (\beta_k - \beta_{k,0}) o_p(1) + o_p(n^{-1/2})
\end{align*}
\] (SC.203)

By the definition of \( \tilde{\beta}_{g,*} \), we can write \( \tilde{\beta}_{g,*} \partial \check{P}_{2,i}(\beta_{k,0}) = \hat{\beta}_g(\hat{\beta}_k) \partial \check{P}_2(\nu_{1,i}(\beta_{k,0}); \hat{\beta}_k) \). Therefore

\[
\begin{align*}
& n^{-1} \sum_{i=1}^{n} \beta_{g,i} \partial \check{P}_{2,i}(\beta_{k,0}) k_{1,i}(k_{2,i} - k_{1,i}g_{1,i}) \\
& = \mathbb{E}[g_{1,i}k_{1,i}(k_{2,i} - k_{1,i}g_{1,i})] \\
& \quad + n^{-1} \sum_{i=1}^{n} (g_{1,i}k_{1,i}(k_{2,i} - k_{1,i}g_{1,i}) - \mathbb{E}[g_{1,i}k_{1,i}(k_{2,i} - k_{1,i}g_{1,i})]) \\
& \quad + n^{-1} \sum_{i=1}^{n} (\hat{\beta}_g(\hat{\beta}_k) \partial \check{P}_2(\nu_{1,i}(\beta_{k,0}); \hat{\beta}_k) - g_{1,i}) k_{1,i}(k_{2,i} - k_{1,i}g_{1,i}).
\end{align*}
\] (SC.204)

By Assumption [SC3-iv], Lemma [SC14] and Lemma [SC20]

\[
\hat{\beta}_g(\hat{\beta}_k) - \tilde{\beta}_{g,m_2}(\hat{\beta}_k) = O_p((m_2 + m_1^{1/2})n^{-1/2}).
\] (SC.205)
By the Markov inequality, Assumptions \( \text{SC1}(i) \) and \( \text{SC2}(i, ii, iii, v) \)

\[
n^{-1} \sum_{i=1}^{n} k_{2,i}^2 (k_{2,i} - k_{1,i} g_{1,i})^2 = O_p(1). \tag{SC.206}
\]

By the mean value expansion,

\[
n^{-1} \sum_{i=1}^{n} \beta_g(\hat{\beta}_k)'(\partial^1 \tilde{P}_2(\nu_{1,i}(\hat{\beta}_k); \tilde{\beta}_k) - \partial^1 \tilde{P}_2(\nu_{1,i}(\hat{\beta}_k); \tilde{\beta}_k)) k_{1,i}(k_{2,i} - k_{1,i} g_{1,i})
= (\hat{\beta}_k - \beta_{k,0}) n^{-1} \sum_{i=1}^{n} \beta_g(\hat{\beta}_k)'(\partial^2 \tilde{P}_2(\nu_{1,i}(\hat{\beta}_k); \tilde{\beta}_k)) k_{1,i}(k_{2,i} - k_{1,i} g_{1,i}) \tag{SC.207}
\]

where \( \hat{\beta}_k \) is between \( \beta_k \) and \( \beta_{k,0} \). By Lemma \( \text{SC20} \) and Assumption \( \text{SC3}(iv) \), \( \nu_{1,i}(\hat{\beta}_k) \in \Omega(\hat{\beta}_k) \) for any \( i = 1, \ldots, n \) wpa1. Therefore by the Cauchy-Schwarz inequality, Assumptions \( \text{SC2}(v) \) and Assumption \( \text{SC3}(iv) \), Lemma \( \text{SC20} \) (SC.78), (SC.206) and (SC.207),

\[
n^{-1} \sum_{i=1}^{n} \beta_g(\hat{\beta}_k)'(\partial^1 \tilde{P}_2(\nu_{1,i}(\hat{\beta}_k); \tilde{\beta}_k) - \partial^1 \tilde{P}_2(\nu_{1,i}(\hat{\beta}_k); \tilde{\beta}_k)) (k_{2,i} - k_{1,i} g_{1,i}) = o_p(1). \tag{SC.208}
\]

By the triangle inequality and the Cauchy-Schwarz inequality, Assumptions \( \text{SC2}(ii, iii, v) \) and \( \text{SC3}(iv) \), Lemma \( \text{SC20} \) (SC.205) and (SC.206),

\[
\left| n^{-1} \sum_{i=1}^{n} \beta_g(\hat{\beta}_k)'(\partial^1 \tilde{P}_2(\nu_{1,i}(\hat{\beta}_k); \tilde{\beta}_k) - g_{1,i}) k_{1,i}(k_{2,i} - k_{1,i} g_{1,i}) \right|
\leq n^{-1} \sum_{i=1}^{n} \left| \beta_g(\hat{\beta}_k) - \beta_{g,m_2}(\hat{\beta}_k)'(\partial^1 \tilde{P}_2(\nu_{1,i}(\hat{\beta}_k); \tilde{\beta}_k)) k_{1,i}(k_{2,i} - k_{1,i} g_{1,i}) \right|
+ n^{-1} \sum_{i=1}^{n} \left| \beta_{g,m_2}(\hat{\beta}_k)'(\partial^1 \tilde{P}_2(\nu_{1,i}(\hat{\beta}_k); \tilde{\beta}_k) - g_{1,i}(\nu_{1,i}(\hat{\beta}_k); \tilde{\beta}_k)) k_{1,i}(k_{2,i} - k_{1,i} g_{1,i}) \right|
+ n^{-1} \sum_{i=1}^{n} \left| (g_{1,i}(\nu_{1,i}(\hat{\beta}_k); \tilde{\beta}_k) - g_{1,i}(\nu_{1,i}(\beta_{k,0}; \beta_{k,0})) k_{1,i}(k_{2,i} - k_{1,i} g_{1,i}) \right|
= O_p(\xi_{1,m_2}(m_2 + m_1^{1/2})n^{-1/2}) + O_p(m^{-r_0}) + (\hat{\beta}_k - \beta_{k,0}) O_p(1) = o_p(1)
\]

which together with (SC.208) implies that

\[
n^{-1} \sum_{i=1}^{n} \beta_g(\hat{\beta}_k)'(\partial^1 \tilde{P}_2(\nu_{1,i}(\hat{\beta}_k); \tilde{\beta}_k) - g_{1,i}) k_{1,i}(k_{2,i} - k_{1,i} g_{1,i}) = o_p(1). \tag{SC.209}
\]
By Assumptions SC1(i) and SC2(i, ii), and the Markov inequality,

$$n^{-1} \sum_{i=1}^{n} g_{1,i} k_{1,i} (k_{2,i} - k_{1,i} g_{1,i}) - E[k_{1,i} g_{1,i} (k_{2,i} - k_{1,i} g_{1,i})] = O_p(n^{-1/2})$$  \hspace{1cm} (SC.210)

which together with (SC.204), (SC.209) and (SC.210) implies that

$$n^{-1} \sum_{i=1}^{n} \beta_{g,\psi} \partial_1 \tilde{P}_{2,i}(\beta_{k,0}) k_{1,i} (k_{2,i} - k_{1,i} g_{1,i}) = E[k_{1,i} g_{1,i} (k_{2,i} - k_{1,i} g_{1,i})] + o_p(1).$$  \hspace{1cm} (SC.211)

The claim in (SC.198) follows from (SC.203) and (SC.211).

**Step 2.** In this step, we show that

$$\beta_{g,\psi}^* \left[ \tilde{P}_2(\beta_{k,0}) \partial_1 \tilde{P}_2(\beta_{k,0},0) - \tilde{P}_2^*(\beta_k) \partial_1 \tilde{P}_2^*(\beta_k) \right] \beta_{\psi}(\beta_{k,0})$$

$$= (\hat{\beta}_k - \beta_{k,0}) \left( E[g_{1,i} k_{1,i} (\gamma_{2,i} - \gamma_{1,i} g_{1,i})] + o_p(1) \right)$$

$$+ n^{-1} \beta_{g,\psi}^* \tilde{P}_2(\beta_{k,0}) \partial_1 \tilde{P}_2(\beta_{k,0}) (\hat{\beta}_k - \beta_{k,0}) \beta_{\psi}(\beta_{k,0}) + o_p(n^{-1/2})$$  \hspace{1cm} (SC.212)

where \( \hat{\beta}_{\psi}(\beta_{k,0}) \equiv (\tilde{P}_2(\beta_{k,0}) \partial_1 \tilde{P}_2(\beta_{k,0},0))^{-1} \sum_{i=1}^{n} \tilde{P}_{2,i}(\beta_{k,0}) (k_{2,i} - k_{1,i} g_{1,i}) \).

First note that

$$\beta_{g,\psi}^* \left[ \tilde{P}_2(\beta_{k,0}) \partial_1 \tilde{P}_2(\beta_{k,0},0) - \tilde{P}_2^*(\beta_k) \partial_1 \tilde{P}_2^*(\beta_k) \right] \beta_{\psi}(\beta_{k,0})$$

$$= \beta_{g,\psi}^* (\tilde{P}_2(\beta_{k,0}) - \tilde{P}_2^*(\beta_k)) \partial_1 \tilde{P}_2(\beta_{k,0}) \beta_{\psi}(\beta_{k,0}) + \beta_{g,\psi}^* \tilde{P}_2^*(\beta_k) \partial_1 \tilde{P}_2(\beta_{k,0}) (\hat{\beta}_k - \beta_{k,0}) \beta_{\psi}(\beta_{k,0}).$$

Therefore to prove (SC.212), it is sufficient to show that

$$n^{-1} \beta_{g,\psi}^* (\tilde{P}_2(\beta_{k,0}) - \tilde{P}_2^*(\beta_k)) \partial_1 \tilde{P}_2(\beta_{k,0}) \beta_{\psi}(\beta_{k,0}) = (\hat{\beta}_k - \beta_{k,0}) \left( E[g_{1,i} k_{1,i} \varphi_i] + o_p(1) \right) + o_p(n^{-1/2}).$$  \hspace{1cm} (SC.213)

where \( \varphi_i \equiv \gamma_{2,i} - \gamma_{1,i} g_{1,i} \).

By the Cauchy-Schwarz inequality, Assumptions SC3(iv), (SC.66), (SC.126) and (SC.183),

$$n^{-1} \left| \beta_{g,\psi}^* (\tilde{P}_2(\beta_{k,0}) - \tilde{P}_2^*(\beta_k)) \partial_1 \tilde{P}_2(\beta_{k,0}) \beta_{\psi}(\beta_{k,0}) \right|$$

$$\leq n^{-1} \left| \beta_{g,\psi}^* (\tilde{P}_2(\beta_{k,0}) - \tilde{P}_2^*(\beta_k)) \right| \left\| \tilde{P}_2(\beta_{k,0}) \beta_{\psi}(\beta_{k,0}) - \tilde{\beta}_{\psi,m_2}(\beta_{k,0}) \right\|$$

$$\leq \left| n^{-1/2} \beta_{g,\psi}^* (\tilde{P}_2(\beta_{k,0}) - \tilde{P}_2^*(\beta_k)) \right| \left\| \beta_{\psi}(\beta_{k,0}) - \tilde{\beta}_{\psi,m_2}(\beta_{k,0}) \right\|$$

$$\leq \left( \lambda_{\max}(n^{-1/2} \tilde{P}_2(\beta_{k,0})) \right)^{-1/2} n^{-1/2} \left| \beta_k - \beta_{k,0} \right| O_p(1) = |\hat{\beta}_k - \beta_{k,0}| o_p(1).$$  \hspace{1cm} (SC.214)

By the Cauchy-Schwarz inequality, Assumptions SC2(iv) and SC3(iv), (SC.61), (SC.127) and
which together with (SC.214) and (SC.215) implies that

\[ \phi \]

Since \( \tilde{\beta}_g \), and (SC.205), we have

By the first-order expansion, the triangle inequality and the Cauchy-Schwarz inequality, Assumptions SC3(iv), Lemma SC4 and Lemma SC20. From (SC.217) and (SC.218),

\[ n^{-1} |\tilde{\beta}_{g,*}'(\tilde{P}_2(\beta_{k,0}) - \tilde{P}_2^*(\tilde{\beta}_k))| \leq n^{-1} \left\| (\tilde{P}_2(\beta_{k,0}) - \tilde{P}_2^*(\tilde{\beta}_k)) \beta_{g,m_2} \right\| \left\| (\tilde{P}_2(\beta_{k,0}) - \tilde{P}_2^*(\tilde{\beta}_k)) \tilde{\beta}_{g,m_2} \right\| = |\tilde{\beta}_k - \beta_{k,0}| O_p(\xi_{1,m_2}^{1/2} n^{-1/2}) = |\tilde{\beta}_k - \beta_{k,0}| o_p(1). \]  

(SC.215)

By Assumption Assumptions SC2(vi) and SC3(ii, iv), and (SC.183), where \( \varphi_n \equiv (\varphi_1, \ldots, \varphi_n)' \), which together with (SC.214) and (SC.215) implies that

\[ n^{-1} \tilde{\beta}_{g,*}'(\tilde{P}_2(\beta_{k,0}) - \tilde{P}_2^*(\tilde{\beta}_k)) = n^{-1} \tilde{\beta}_{g,*}'(\tilde{P}_2(\beta_{k,0}) - \tilde{P}_2^*(\tilde{\beta}_k)) + (\tilde{\beta}_k - \beta_{k,0}) o_p(1). \]  

(SC.216)

Since \( \tilde{\beta}_{g,*} = (B(\beta_{k,0})')^{-1} B(\tilde{\beta}_k)' \beta_g(\tilde{\beta}_k) \), we can write

\[ \tilde{\beta}_{g,*}'(\tilde{P}_2(\beta_{k,0}) - \tilde{P}_2^*(\tilde{\beta}_k)') \varphi_n = \sum_{i=1}^{n} \hat{\beta}_g(\hat{\beta}_k)'(\hat{P}_2(\hat{\nu}_{1,i}(\beta_{k,0}); \hat{\beta}_k) - \hat{P}_2(\hat{\nu}_{1,i}(\hat{\beta}_k); \hat{\beta}_k)) \varphi_i. \]  

(SC.217)

By the first-order expansion, the triangle inequality and the Cauchy-Schwarz inequality, Assumptions SC1(i) and SC3(i, iv), and (SC.205), we have

\[ n^{-1} \sum_{i=1}^{n} (\hat{\beta}_g(\hat{\beta}_k) - \hat{\beta}_{g,m_2}(\hat{\beta}_k))'(\hat{P}_2(\hat{\nu}_{1,i}(\beta_{k,0}); \hat{\beta}_k) - \hat{P}_2(\hat{\nu}_{1,i}(\hat{\beta}_k); \hat{\beta}_k)) \varphi_i = (\hat{\beta}_k - \beta_{k,0}) n^{-1} \sum_{i=1}^{n} (\hat{\beta}_g(\hat{\beta}_k) - \hat{\beta}_{g,m_2}(\hat{\beta}_k))' \partial^1 \hat{P}_2(\hat{\nu}_{1,i}; \hat{\beta}_k) k_{1,i} \varphi_i = (\hat{\beta}_k - \beta_{k,0}) O_p(m_1^{1/2} + m_2) n^{-1/2} O_p(\xi_{1,m_2}) = (\hat{\beta}_k - \beta_{k,0}) o_p(1) \]  

(SC.218)

where \( \hat{\nu}_{1,i} \) is between \( \hat{\nu}_{1,i}(\beta_{k,0}) \) and \( \nu_{1,i}(\hat{\beta}_k) \) and it is in \( \Omega_\varepsilon(\hat{\beta}_k) \) for any \( i = 1, \ldots, n \) wp 1 by Assumption SC3(iv), Lemma SC4 and Lemma SC20. From (SC.217) and (SC.218),

\[ \tilde{\beta}_{g,*}'(\tilde{P}_2(\beta_{k,0}) - \tilde{P}_2^*(\tilde{\beta}_k))' \varphi_n = \sum_{i=1}^{n} \hat{\beta}_{g,m_2}(\hat{\beta}_k)'(\hat{P}_2(\hat{\nu}_{1,i}(\beta_{k,0}); \hat{\beta}_k) - \hat{P}_2(\hat{\nu}_{1,i}(\hat{\beta}_k); \hat{\beta}_k)) \varphi_i + (\hat{\beta}_k - \beta_{k,0}) o_p(1). \]  

(SC.219)
By the second order expansion,

\[ n^{-1} \sum_{i=1}^{n} \tilde{\beta}_{g,m_2}(\hat{\beta}_k)'(P_2(\hat{\nu}_{1,i}(\beta_k,0); \hat{\beta}_k) - \tilde{P}_2(\nu_{1,i}(\hat{\beta}_k); \hat{\beta}_k))\varphi_i \]

\[ = n^{-1} \sum_{i=1}^{n} \tilde{\beta}_{g,m_2}(\hat{\beta}_k)'\partial^1 \tilde{P}_2(\nu_{1,i}(\hat{\beta}_k); \hat{\beta}_k)(\hat{\nu}_{1,i}(\beta_k,0) - \nu_{1,i}(\hat{\beta}_k))\varphi_i \]

\[ + n^{-1} \sum_{i=1}^{n} \tilde{\beta}_{g,m_2}(\hat{\beta}_k)'\partial^2 \tilde{P}_2(\nu_{1,i}; \hat{\beta}_k)(\hat{\nu}_{1,i}(\beta_k,0) - \nu_{1,i}(\hat{\beta}_k))^2 \varphi_i \]  \hspace{1cm} (SC.220)

where \( \hat{\nu}_{1,i} \) lies between \( \hat{\nu}_{1,i}(\beta_k,0) \) and \( \nu_{1,i}(\hat{\beta}_k) \). By Assumption \( SC3(iv) \), Lemma \( SC4 \) and Lemma \( SC20 \), \( \hat{\nu}_{1,i} \in \Omega_{\text{c}}(\beta_k) \) for any \( i = 1, \ldots, n \) wpa1. By Assumptions \( SC2(iii, vi) \) and \( SC3(i, iv) \), Lemma \( SC4 \) and Lemma \( SC20 \),

\[ n^{-1} \sum_{i=1}^{n} \tilde{\beta}_{g,m_2}(\hat{\beta}_k)'\partial^1 \tilde{P}_2(\nu_{1,i}(\hat{\beta}_k); \hat{\beta}_k)(\hat{\nu}_{1,i}(\beta_k,0) - \nu_{1,i}(\hat{\beta}_k))\varphi_i \]

\[ = n^{-1} \sum_{i=1}^{n} g_1(\nu_{1,i}(\hat{\beta}_k); \hat{\beta}_k)(\hat{\nu}_{1,i}(\beta_k,0) - \nu_{1,i}(\hat{\beta}_k))\varphi_i + o_p(n^{-1/2}). \]  \hspace{1cm} (SC.221)

By Assumptions \( SC2(v, vi) \) and \( SC3(i, iv) \), Lemma \( SC4 \) and Lemma \( SC20 \) and \( SC.69 \),

\[ n^{-1} \sum_{i=1}^{n} \tilde{\beta}_{g,m_2}(\hat{\beta}_k)'\partial^2 \tilde{P}_2(\nu_{1,i}; \hat{\beta}_k)(\hat{\nu}_{1,i}(\beta_k,0) - \nu_{1,i}(\hat{\beta}_k))^2 \varphi_i = (\hat{\beta}_k - \beta_k,0) + o_p(n^{-1/2}) \]

which together with \( SC.220 \) and \( SC.221 \) implies that

\[ n^{-1} \sum_{i=1}^{n} \tilde{\beta}_{g,m_2}(\hat{\beta}_k)'(\tilde{P}_2(\hat{\nu}_{1,i}(\beta_k,0); \hat{\beta}_k) - \tilde{P}_2(\nu_{1,i}(\hat{\beta}_k); \hat{\beta}_k))\varphi_i \]  \hspace{1cm} (SC.222)

\[ = n^{-1} \sum_{i=1}^{n} g_1(\nu_{1,i}(\hat{\beta}_k); \hat{\beta}_k)(\hat{\nu}_{1,i}(\beta_k,0) - \nu_{1,i}(\hat{\beta}_k))\varphi_i + (\hat{\beta}_k - \beta_k,0) + o_p(n^{-1/2}). \]

Similarly, we can show that

\[ n^{-1} \sum_{i=1}^{n} \tilde{\beta}_{g,m_2}(\hat{\beta}_k)'(\tilde{P}_2(\hat{\nu}_{1,i}(\beta_k,0); \hat{\beta}_k) - \tilde{P}_2(\nu_{1,i}(\hat{\beta}_k); \hat{\beta}_k))\varphi_i \]

\[ = n^{-1} \sum_{i=1}^{n} g_1(\nu_{1,i}(\hat{\beta}_k); \hat{\beta}_k)(\hat{\nu}_{1,i}(\beta_k,0) - \nu_{1,i}(\hat{\beta}_k))\varphi_i + o_p(n^{-1/2}) \]
which together with (SC.222) implies that

\[ n^{-1} \sum_{i=1}^{n} \hat{c}_{g,m} ((\hat{P}_2(\hat{\nu}_{1,i}(\hat{\beta}_k); \hat{\beta}_k) - \hat{P}_2(\nu_{1,i}(\hat{\beta}_k); \hat{\beta}_k)) \varphi_i \]

\[ = (\hat{\beta}_k - \beta_{k,0}) n^{-1} \sum_{i=1}^{n} g_1(\nu_{1,i}(\hat{\beta}_k); \hat{\beta}_k) k_{1,i} \varphi_i + (\hat{\beta}_k - \beta_{k,0}) o_p(1) + o_p(n^{-1/2}) \quad (SC.223) \]

By Assumptions SC2(ii, vi) and SC3(i, iv), and Lemma SC20

\[ n^{-1} \sum_{i=1}^{n} g_1(\nu_{1,i}(\hat{\beta}_k); \hat{\beta}_k) k_{1,i} \varphi_i = n^{-1} \sum_{i=1}^{n} g_1 k_{1,i} \varphi_i + o_p(1) \]

which combined with (SC.223) implies that

\[ n^{-1} \sum_{i=1}^{n} \hat{c}_{g,m} ((\hat{P}_2(\hat{\nu}_{1,i}(\hat{\beta}_k); \hat{\beta}_k) - \hat{P}_2(\hat{\nu}_{1,i}(\hat{\beta}_k); \hat{\beta}_k)) \varphi_i \]

\[ = (\hat{\beta}_k - \beta_{k,0}) n^{-1} \sum_{i=1}^{n} g_1 k_{1,i} \varphi_i + (\hat{\beta}_k - \beta_{k,0}) o_p(1) + o_p(n^{-1/2}) \]

\[ = (\hat{\beta}_k - \beta_{k,0}) \mathbb{E} [g_1 k_{1,i} \varphi_i] + (\hat{\beta}_k - \beta_{k,0}) o_p(1) + o_p(n^{-1/2}) \quad (SC.224) \]

where the second equality is by the Markov inequality. The claim in (SC.213) now follows from (SC.219) and (SC.224).

**Step 3.** In this step, we show that

\[ n^{-1}(\hat{Y}_2^*(\hat{\beta}_k) - \hat{P}_2(\hat{\beta}_k) \hat{c}_g(\hat{\beta}_k))'(\hat{P}_2^*(\hat{\beta}_k) - \hat{P}_2(\beta_{k,0})) \hat{\beta}_\varphi(\beta_{k,0}) = (\hat{\beta}_k - \beta_{k,0}) o_p(1) + o_p(n^{-1/2}). \quad (SC.225) \]

Since \( \hat{\beta}_k^* = y_2^* - l_{2,i} \hat{\beta}_l - k_{2,i} \hat{\beta}_k \), we can write

\[ \hat{y}_2^*(\hat{\beta}_k) - \hat{P}_2(\hat{\beta}_k)' \hat{c}_g(\hat{\beta}_k) = y_2^* - l_{2,i} \hat{\beta}_l - k_{2,i} \hat{\beta}_k - \hat{g}(\hat{\nu}_{1,i}(\hat{\beta}_k); \hat{\beta}_k) \]

\[ = u_{2,i} - l_{2,i} (\hat{\beta}_l - \beta_{l,0}) - k_{2,i} (\hat{\beta}_k - \beta_{k,0}) - (\hat{g}(\hat{\nu}_{1,i}(\hat{\beta}_k); \hat{\beta}_k) - g(\nu_{1,i})). \]

Therefore,

\[ n^{-1}(\hat{Y}_2^*(\hat{\beta}_k) - \hat{P}_2(\hat{\beta}_k) \hat{c}_g(\hat{\beta}_k))'(\hat{P}_2^*(\hat{\beta}_k) - \hat{P}_2(\beta_{k,0})) \hat{\beta}_\varphi(\beta_{k,0}) \]

\[ = n^{-1} U_2' \hat{P}_2^*(\hat{\beta}_k) - \hat{P}_2(\beta_{k,0})) \hat{\beta}_\varphi(\beta_{k,0}) \]

\[ - (\hat{\beta}_l - \beta_{l,0}) n^{-1} L_2'(\hat{P}_2^*(\hat{\beta}_k) - \hat{P}_2(\beta_{k,0})) \hat{\beta}_\varphi(\beta_{k,0}) \]

\[ - n^{-1}(\hat{\beta}_k - \beta_{k,0}) K_2' (\hat{P}_2^*(\hat{\beta}_k) - \hat{P}_2(\beta_{k,0})) \hat{\beta}_\varphi(\beta_{k,0}) \]

\[ - n^{-1}(G_2 - G_2)' (\hat{P}_2^*(\hat{\beta}_k) - \hat{P}_2(\beta_{k,0})) \hat{\beta}_\varphi(\beta_{k,0}) \quad (SC.226) \]
which combined with Lemma SC22 proves (SC.225).

**Step 4.** In this step, we show that

\[ n^{-1}(\hat{Y}_2^*(\hat{\beta}_k) - \hat{Y}_2^*(\beta_{k,0}))'\hat{P}_2(\beta_{k,0})\hat{\beta}_\varphi(\beta_{k,0}) = -(\hat{\beta}_k - \beta_{k,0})(E[k_{2,i}(\gamma_{2,i} - \gamma_{1,i}g_{1,i})] + o_p(1)). \]  

(SC.227)

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

\[ n^{-1}(\hat{Y}_2^*(\hat{\beta}_k) - \hat{Y}_2^*(\beta_{k,0}))'\hat{P}_2(\beta_{k,0})\hat{\beta}_\varphi(\beta_{k,0}) = n^{-1}(\hat{\beta}_k - \beta_{k,0})'K_2'\hat{P}_2(\beta_{k,0})\hat{\beta}_\varphi(\beta_{k,0}) \]  

(SC.228)

and

\[
n^{-1}K_2'\hat{P}_2(\beta_{k,0})\hat{\beta}_\varphi(\beta_{k,0}) = E[k_{2,i}\varphi(\nu_{1,i})] + n^{-1}\sum_{i=1}^{n}(k_{2,i}\varphi(\nu_{1,i}) - E[k_{2,i}\varphi(\nu_{1,i})]) \\
+ n^{-1}\sum_{i=1}^{n}k_{2,i}(\varphi(\nu_{1,i}) - \hat{P}_{2,i}(\beta_{k,0})'\hat{\beta}_{\varphi, m_2}(\beta_{k,0})) \\
+ n^{-1}\sum_{i=1}^{n}k_{2,i}\hat{P}_{2,i}(\beta_{k,0})'(\hat{\beta}_\varphi(\beta_{k,0}) - \hat{\beta}_{\varphi, m_2}(\beta_{k,0})). \]  

(SC.229)

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

\[ n^{-1}\sum_{i=1}^{n}(k_{2,i}\varphi(\nu_{1,i}) - E[k_{2,i}\varphi(\nu_{1,i})]) = o_p(1) \]  

(SC.230)

By the mean value expansion, Assumptions SC1(i), SC2(i) and SC3(i, iv)

\[
n^{-1}\sum_{i=1}^{n}k_{2,i}(\varphi(\nu_{1,i}) - \hat{P}_{2,i}(\beta_{k,0})'\hat{\beta}_{\varphi, m_2}(\beta_{k,0})) \\
= n^{-1}\sum_{i=1}^{n}k_{2,i}(\hat{P}_{2}(\nu_{1,i}(\beta_{k,0}); \beta_{k,0}) - \hat{P}_{2}(\nu_{1,i}(\beta_{k,0}); \beta_{k,0}))'\hat{\beta}_{\varphi, m_2}(\beta_{k,0}) \\
+ n^{-1}\sum_{i=1}^{n}k_{2,i}(\varphi(\nu_{1,i}) - \hat{P}_{2}(\nu_{1,i}(\beta_{k,0}); \beta_{k,0})')\hat{\beta}_{\varphi, m_2}(\beta_{k,0})) \\
= -n^{-1}\sum_{i=1}^{n}k_{2,i}(\phi_i - \phi_i)\partial_1\hat{P}_{2}(\nu_{1,i}(\beta_{k,0}); \beta_{k,0})'\hat{\beta}_{\varphi, m_2}(\beta_{k,0}) + o_p(1) \]  

(SC.231)

where \( \hat{\nu}_{1,i}(\beta_{k,0}) \) lies between \( \hat{\nu}_{1,i}(\beta_{k,0}) \) and \( \nu_{1,i}(\beta_{k,0}) \). By (SC.59), \( \hat{\nu}_{1,i}(\beta_{k,0}) \in \Omega_\varepsilon(\beta_{k,0}) \) for any \( i = 1, \ldots, n \) wpa1. By the triangle inequality and the Cauchy-Schwarz inequality, Assumptions SC1(i), SC2(i, v) and SC3(iv), and (SC.127)
n^{-1} \sum_{i=1}^{n} k_{2,i}(\hat{\phi}_i - \phi_i)\partial^2 \tilde{P}_2(\tilde{v}_{1,i}(\beta_{k,0}); \beta_{k,0})' \hat{\beta}_{\varphi,m_2}(\beta_{k,0}) = O_p(\xi_{1,m_2}m_1^{1/2}n^{-1/2}) = o_p(1)

which together with (SC.231) implies that

\[
n^{-1} \sum_{i=1}^{n} k_{2,i}(\varphi(\nu_{1,i}) - \hat{P}_{2,i}(\beta_{k,0})' \hat{\beta}_{\varphi,m_2}(\beta_{k,0})) = o_p(1).
\]

(SC.232)

By the Cauchy-Schwarz inequality, Assumptions SC1(i) and SC2(i) (SC.66) and (SC.126)

\[
\left| n^{-1} \sum_{i=1}^{n} k_{2,i} \hat{P}_{2,i}(\beta_{k,0})'(\hat{\beta}_{\varphi}(\beta_{k,0}) - \hat{\beta}_{\varphi,m_2}(\beta_{k,0})) \right|
\]

\[
\leq \left( \lambda_{\max}(n^{-1} \hat{P}_2(\beta_{k,0})' \hat{P}_2(\beta_{k,0})) \right)^{1/2} \left( n^{-1} \sum_{i=1}^{n} k_{2,i}^2 \right)^{1/2} \left\| \hat{\beta}_{\varphi}(\beta_{k,0}) - \hat{\beta}_{\varphi,m_2}(\beta_{k,0}) \right\| = o_p(1).
\]

The claim in (SC.227) follows from (SC.228), (SC.229), (SC.230), (SC.232) and (SC.233). Q.E.D.

SC.4 Auxiliary results for the standard error estimation

Assumption SC4. (i) There exist \( \hat{\varepsilon}_{1,i} \) 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 \( \tau_a > 0 \) and \( \beta_{a_2,m} \in \mathbb{R}^m \) such that \( \sup_{x \in \mathcal{X}} |a_{2,m}(x) - a_2(x)| = O(m^{-\tau_a}) \) where \( a_{2,m}(x) \equiv P_1(x)' \beta_{a_2,m} \) and \( \xi_{0,m_1}m^{-\tau_a} = o(1) \); (iii) \( \Omega > 0 \); (iv) \( \xi_{0,m_1}m_1^{1/2}m_2^{3n^{-1/2}} = o(1) \).

Assumption SC4(i) assumes the existence of estimators of the random variables \( \varepsilon_{1,i} \) in the linear representation of the estimation error in \( \hat{\beta}_t \). Specific estimator \( \hat{\varepsilon}_{1,i} \) can be constructed using the form of \( \varepsilon_{1,i} \). Assumption SC4(ii) requires that the unknown function \( a_2(x_{1,i}) = \mathbb{E}[k_{2,i}x_{1,i}] \) can be well approximated by the approximating functions \( P_1(x_{1,i}) \). Assumption SC4(iii) requires that the asymptotic variance \( \Omega \) is bounded away from zero. Assumption SC4(iv) restricts the numbers of the approximation functions used in the multi-step estimation procedure.

The following lemma is useful to show the consistency of the standard error estimator.

**Lemma SC24.** Under Assumptions SC1, SC2, SC3 and SC4, we have

(i) \( n^{-1}\hat{P}_2(\hat{\beta}_k)'\hat{P}_2(\beta_k) - n^{-1}\hat{P}_2(\beta_{k,0})'\hat{P}_2(\beta_{k,0}) = O_p(\xi_{1,m_2}n^{-1/2}) \);

(ii) \( \max_{i \leq n} |\hat{g}_{1,i} - g_{1,i}|^4 = o_p(1) \);

(iii) \( n^{-1}\sum_{i=1}^{n}(\hat{\varepsilon}_{1,i} - \varepsilon_{1,i})^4 = o_p(1) \);

(iv) \( n^{-1}\sum_{i=1}^{n}(\hat{u}_{2,i} - u_{2,i})^4 = o_p(1) \);

(v) \( \max_{i \leq n} |\hat{\varepsilon}_{1,i} - \varepsilon_{1,i}| = o_p(1) \).

\(^5\)See (SC.271) in Subsection SC.5 for the form of \( \hat{\varepsilon}_{1,i} \) when \( \beta_{t,0} \) is estimated by the partially linear regression proposed in Olley and Pakes (1996).
Proof of Lemma SC24

(i) For any $b \in \mathbb{R}^{m_2}$, by the mean value expansion, Assumption SC2(v) and (SB.16)

$$b'((\hat{P}_2(\beta_k) - \hat{P}_2(\beta_{k,0})))'((\hat{P}_2(\beta_k) - \hat{P}_2(\beta_{k,0}))b)$$

$$= \sum_{i=1}^{n} (b'(P_{2,i}(\beta_k) - \hat{P}_{2,i}(\beta_{k,0})))^2$$

$$= (\hat{\beta}_k - \beta_{k,0})^2 \sum_{i=1}^{n} (b'(\nu_{1,i}(\beta_k); \hat{\beta}_k)/\partial \beta_k)^2 = \|b\|^2 O_p(\xi_{1,m_2}^2).$$

where $\hat{\beta}_k$ is between $\tilde{\beta}_k$ and $\beta_{k,0}$, and by (SC.58) and Assumption SC2(vi), $\nu_{1,i}(\tilde{\beta}_k) \in \Omega_\varepsilon(\tilde{\beta}_k)$ for any $i = 1, \ldots, n$ wpa1. Therefore

$$\|\hat{P}_2(\beta_k) - \hat{P}_2(\beta_{k,0})\|_S = O_p(\xi_{1,m_2}).$$  \hfill (SC.234)

By the triangle inequality and the Cauchy-Schwarz inequality, Assumptions SC2(vi) and SC3(iv), (SC.66) and (SC.234)

$$n^{-1} \left\| \hat{P}_2(\beta_k)'\hat{P}_2(\beta_k) - \hat{P}_2(\beta_{k,0})'\hat{P}_2(\beta_{k,0}) \right\|_S$$

$$\leq n^{-1} \left\| (\hat{P}_2(\beta_k)' - \hat{P}_2(\beta_{k,0})')\hat{P}_2(\beta_{k,0})\right\|_S$$

$$+ n^{-1} \left\| \hat{P}_2(\beta_{k,0})'(\hat{P}_2(\beta_k) - \hat{P}_2(\beta_{k,0}))\right\|_S$$

$$+ n^{-1} \left\| (\hat{P}_2(\beta_k) - \hat{P}_2(\beta_{k,0}))(\hat{P}_2(\beta_k) - \hat{P}_2(\beta_{k,0}))\right\|_S = O_p(\xi_{1,m_2}n^{-1/2})$$  \hfill (SC.235)

which finishes the proof.

(ii) By triangle inequality and the Cauchy-Schwarz inequality, Assumption SC2(iii, v), and (SC.205)

$$\max_{i \leq n} \left\| \partial^1 \hat{P}_{2,i}(\tilde{\beta}_k)'\hat{g}(\tilde{\beta}_k) - g_1(\nu_{1,i}(\tilde{\beta}_k); \tilde{\beta}_k) \right\|_S$$

$$\leq \max_{i \leq n} \left\| (\partial^1 \hat{P}_{2,i}(\tilde{\beta}_k) - \partial^1 \hat{P}_{2,i}(\tilde{\beta}_k))'\hat{g}(\tilde{\beta}_k) \right\|_S$$

$$+ \max_{i \leq n} \left\| \partial^1 \hat{P}_{2,i}(\tilde{\beta}_k)'\hat{g}(\tilde{\beta}_k) - \hat{g}_{m_2}(\tilde{\beta}_k) \right\|_S$$

$$+ \max_{i \leq n} \left\| \partial^1 \hat{P}_{2,i}(\tilde{\beta}_k)'\hat{g}_{m_2}(\tilde{\beta}_k) - g_1(\nu_{1,i}(\tilde{\beta}_k); \tilde{\beta}_k) \right\|_S$$

$$\leq \max_{i \leq n} \left\| (\partial^1 \hat{P}_{2,i}(\tilde{\beta}_k) - \partial^1 \hat{P}_{2,i}(\tilde{\beta}_k))'\hat{g}(\tilde{\beta}_k) \right\| + O_p(\xi_{1,m_2}(m_2 + m_1^{1/2})n^{-1/2}).$$  \hfill (SC.236)

By the mean value expansion,

$$\partial^1 \hat{P}_{2,i}(\tilde{\beta}_k) - \partial^1 \hat{P}_{2,i}(\tilde{\beta}_k) = (\hat{\phi}_i - \phi_i)\partial^2 \hat{P}_{2}(\nu_{1,i}(\tilde{\beta}_k); \tilde{\beta}_k)'\hat{g}(\tilde{\beta}_k) \hfill (SC.237)$$
By Assumptions SC2(vi) and SC3(iv), (SC.152), (SC.153) and (SC.205), \( \nu_{1,i}(\hat{\beta}_k) \) is between \( \nu_{1,i}(\hat{\beta}_k) \) and \( \nu_{1,i}(\hat{\beta}_k) \). By Assumption SC2(ii) and (SB.16), we have Lemma SC4, (SC.78), (SC.236) and (SC.237), where \( \tilde{\nu} \) which together with (SC.239) implies that

\[
\max_{i \leq n} \left| \partial^1 \hat{P}_{2,i}(\hat{\beta}_k) \beta'_{g}(\hat{\beta}_k) - g_1(\nu_{1,i}(\hat{\beta}_k); \hat{\beta}_k) \right| \\
\leq O_p(\xi_{2,m_2}) \max_{i \leq n} \left| \hat{\phi}_i - \phi_i \right| + O_p(\xi_{1,m_2}(m_2 + m_1^{1/2})n^{-1/2}) \\
= O_p(\xi_{0,m_1}(m_1^{1/2} n^{-1/2}) + O_p(\xi_{1,m_2}(m_2 + m_1^{1/2})n^{-1/2}) = o_p(1). \quad \text{(SC.238)}
\]

By Assumption SC2(ii) and (SB.16), we have

\[
\max_{i \leq n} \left| g_1(\nu_{1,i}(\hat{\beta}_k); \hat{\beta}_k) - g_1(\nu_{1,i}(\hat{\beta}_k,0); \hat{\beta}_k,0) \right| = O_p(n^{-1/2}) \quad \text{(SC.239)}
\]

which together with (SC.238) proves the second claim of the lemma.

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

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

Recall that \( \Delta k_{2,i} \equiv k_{2,i} - k_{1,i}g_{1,i} \) and \( \bar{k}_{2,i} \equiv k_{2,i} - k_{1,i}\hat{g}_{1,i}(\hat{\beta}_k) \). Since \( s_{1,i} = \Delta k_{2,i} - \varphi_i \) and \( \hat{s}_{1,i} = \bar{k}_{2,i} - \hat{\varphi}_i \), we have

\[
n^{-1} \sum_{i = 1}^{n} (\hat{s}_{1,i} - s_{1,i})^4 \leq Cn^{-1} \sum_{i = 1}^{n} (\Delta k_{2,i} - \Delta k_{2,i})^4 + Cn^{-1} \sum_{i = 1}^{n} (\hat{\varphi}_i - \varphi_i)^4. \quad \text{(SC.240)}
\]

By Lemma SC24(ii),

\[
n^{-1} \sum_{i = 1}^{n} (\Delta k_{2,i} - \Delta k_{2,i})^4 = n^{-1} \sum_{i = 1}^{n} k_{1,i}^4(\hat{g}_{1,i}(\hat{\beta}_k) - g_1(\nu_{1,i}(\hat{\beta}_k,0); \hat{\beta}_k,0))^4 = o_p(1). \quad \text{(SC.241)}
\]

By Assumptions SC2(vi) and SC3(iv), (SC.152), (SC.153) and (SC.205)

\[
n^{-1} \sum_{i = 1}^{n} \left| \hat{g}_{1,i}(\hat{\beta}_k) - g_1(\nu_{1,i}(\hat{\beta}_k); \hat{\beta}_k) \right|^2 = O_p(m_1 m_2^6 n^{-1})
\]

which together with (SC.239) implies that

\[
n^{-1} \sum_{i = 1}^{n} (\hat{g}_{1,i}(\hat{\beta}_k) - g_1(\nu_{1,i}(\hat{\beta}_k,0); \hat{\beta}_k,0))^2 = O_p(m_1 m_2^6 n^{-1}). \quad \text{(SC.242)}
\]
Therefore,
\[ n^{-1} \sum_{i=1}^{n} (\Delta \hat{k}_{2,i} - \Delta k_{2,i})^2 \leq C n^{-1} \sum_{i=1}^{n} (\hat{g}_{1,i}(\hat{\beta}_k) - g_1(v_{1,i}(\beta; \beta_{k,0}))^2 = O_p(m_1 n_2^6 n^{-1}). \] (SC.243)

By the definition of \( \varphi_i \), we can write
\[
\varphi_i - \varphi = \hat{P}_{2,i}(\hat{\beta}_k)'(\hat{P}_2(\hat{\beta}_k)'\hat{P}_2(\hat{\beta}_k))^{-1}\hat{P}_2(\hat{\beta}_k)(\Delta \hat{K}_2 - \Delta K_2) \\
+ \hat{P}_{2,i}(\hat{\beta}_k)'(\hat{P}_2(\hat{\beta}_k)'\hat{P}_2(\hat{\beta}_k))^{-1}(\hat{P}_2(\hat{\beta}_k) - \hat{P}_2(\beta_{k,0}))\Delta K_2 \\
+ \hat{P}_{2,i}(\hat{\beta}_k)'[(\hat{P}_2(\beta_{k,0})'\hat{P}_2(\beta_{k,0}))^{-1} - (\hat{P}_2(\beta_{k,0})'\hat{P}_2(\beta_{k,0}))^{-1}]\hat{P}_2(\beta_{k,0})\Delta K_2 \\
+ (\hat{P}_{2,i}(\hat{\beta}_k) - \hat{P}_{2,i}(\beta_{k,0}))'(\hat{P}_2(\beta_{k,0})'\hat{P}_2(\beta_{k,0}))^{-1}\hat{P}_2(\beta_{k,0})\Delta K_2 \\
+ \hat{P}_{2,i}(\beta_{k,0})'(\hat{P}_2(\beta_{k,0})'\hat{P}_2(\beta_{k,0}))^{-1}\hat{P}_2(\beta_{k,0})\Delta K_2 - \varphi_i. \] (SC.244)

where \( \Delta \hat{K}_2 \equiv (\Delta \hat{k}_{2,1}, \ldots, \Delta \hat{k}_{2,n})' \) and \( \Delta K_2 \equiv (\Delta k_{2,1}, \ldots, \Delta k_{2,n})' \). By Assumption [SC2](v) and (SC.59),
\[
\max_{1 \leq i \leq n} \| \hat{P}_{2,i}(\hat{\beta}_k) \| = O_p(\xi_{0,m_2}). \] (SC.245)

By Assumption [SC2](v, vi), (SC.66), (SC.243) and (SC.245),
\[
\begin{align*}
& n^{-1} \sum_{i=1}^{n} (\hat{P}_{2,i}(\hat{\beta}_k)'(\hat{P}_2(\hat{\beta}_k)'\hat{P}_2(\hat{\beta}_k))^{-1}\hat{P}_2(\hat{\beta}_k)(\Delta \hat{K}_2 - \Delta K_2))^4 \\
& \leq \xi_{0,m_2}^2 \| (\hat{P}_2(\hat{\beta}_k)'\hat{P}_2(\hat{\beta}_k))^{-1}\hat{P}_2(\hat{\beta}_k)(\Delta \hat{K}_2 - \Delta K_2) \|^2 \\
& \times n^{-1} \sum_{i=1}^{n} (\hat{P}_{2,i}(\hat{\beta}_k)'(\hat{P}_2(\hat{\beta}_k)'\hat{P}_2(\hat{\beta}_k))^{-1}\hat{P}_2(\hat{\beta}_k)(\Delta \hat{K}_2 - \Delta K_2))^2 \\
& \leq (\lambda_{\min}(n^{-1}\hat{P}_2(\hat{\beta}_k)'\hat{P}_2(\hat{\beta}_k)))^{-1} \xi_{0,m_2}^2 \left( n^{-1} \sum_{i=1}^{n} (\Delta \hat{k}_{2,i} - \Delta k_{2,i})^2 \right)^2 \\
& = O_p \left( m_1 \xi_{0,m_2}^2 n_2^6 n^{-1} \right) = o_p(1). \] \] (SC.246)

By the Cauchy-Schwarz inequality and Assumptions [SC1](i), [SC2](i, vi) and [SC3](iv), (SC.66),
By the Cauchy-Schwarz inequality, Lemma (SC.24(i), (SC.78), (SC.127) and (SC.245),

\[
\begin{align*}
&n^{-1} \sum_{i=1}^{n} \left| \hat{P}_{2,i}(\hat{\beta}_k)'(\hat{P}_2(\hat{\beta}_k)'\hat{P}_2(\hat{\beta}_k))^{-1}(\hat{P}_2(\hat{\beta}_k) - \hat{P}_2(\beta_{k,0}))\Delta K_2 \right|^4 \\
\leq & n^{-1} \xi_{0,m2}^2 \left\| \hat{P}_2(\hat{\beta}_k)'(\hat{P}_2(\hat{\beta}_k)'\hat{P}_2(\hat{\beta}_k))^{-1}(\hat{P}_2(\hat{\beta}_k) - \hat{P}_2(\beta_{k,0}))\Delta K_2 \right\|^2 \\
&\times \left\| (\hat{P}_2(\hat{\beta}_k)'\hat{P}_2(\hat{\beta}_k))^{-1/2}(\hat{P}_2(\hat{\beta}_k) - \hat{P}_2(\beta_{k,0}))\Delta K_2 \right\|^2 \\
\leq & \frac{\xi_{0,m2}^2 \left\| \hat{P}_2(\hat{\beta}_k) - \hat{P}_2(\beta_{k,0}) \right\|^4}{n^2(\lambda_{\min}(n^{-1}\hat{P}_2(\beta_k)'\hat{P}_2(\beta_k)))^3} \left| n^{-1} \sum_{i=1}^{n} (\Delta k_{2,i})^2 \right|^2 = O_p(\xi_{0,m2}^2 \xi_{1,m2}^4 n^{-2}) = o_p(1).
\end{align*}
\]

By the Cauchy-Schwarz inequality, Lemma (SC.24(i), (SC.78), (SC.127) and (SC.245),

\[
\begin{align*}
&n^{-1} \sum_{i=1}^{n} \left| \hat{P}_{2,i}(\hat{\beta}_k)'(\hat{P}_2(\hat{\beta}_k)'\hat{P}_2(\hat{\beta}_k))^{-1} - (\hat{P}_2(\beta_{k,0})')\hat{P}_2(\beta_{k,0})\right| \Delta K_2 \right|^4 \\
= & n^{-1} \sum_{i=1}^{n} \left| \hat{P}_{2,i}(\hat{\beta}_k)'(\hat{P}_2(\hat{\beta}_k)'\hat{P}_2(\hat{\beta}_k))^{-1}\hat{P}_2(\hat{\beta}_k) - \hat{P}_2(\beta_{k,0})'\hat{P}_2(\beta_{k,0})\right|^4 \\
\leq & \frac{\xi_{0,m2}^2 \left\| \hat{P}_2(\hat{\beta}_k) - \hat{P}_2(\beta_{k,0}) \right\|^4}{n^2(\lambda_{\min}(n^{-1}\hat{P}_2(\beta_k)'\hat{P}_2(\beta_k)))^3} \left| n^{-1} \sum_{i=1}^{n} (\Delta k_{2,i})^2 \right|^2 = O_p(\xi_{0,m2}^2 \xi_{1,m2}^4 n^{-2}) = o_p(1)
\end{align*}
\]

where the second equality is by Assumptions (SC2 vi) and (SC3 iv). By the first order expansion, (SB.16) in Theorem SB1, Assumption (SC3 ii) and (SC.127),

\[
\begin{align*}
n^{-1} \sum_{i=1}^{n} ((\hat{P}_2,i(\hat{\beta}_k) - \hat{P}_2,i(\beta_{k,0}))')\hat{\beta}_\varphi(\beta_{k,0})^4 \\
= & (\hat{\beta}_k - \beta_{k,0})^4 n^{-1} \sum_{i=1}^{n} (\partial\hat{P}_2(\hat{1},i(\beta_k);\hat{\beta}_k)/\partial\beta_k')\hat{\beta}_\varphi(\beta_{k,0})^4 \\
\leq & (\hat{\beta}_k - \beta_{k,0})^4 \xi_{0,m2} \left\| \hat{\beta}_\varphi(\beta_{k,0}) \right\|^4 = O_p(\xi_{0,m2}^4 n^{-2}) = o_p(1)
\end{align*}
\]

where the second equality is by Assumptions (SC2 vi) and (SC3 iv). By Assumptions (SC2 v) and
which implies that where the equality is by Assumptions SC1(i, iii) and SC2(i, ii), and (SB.16). Using similar argu-

lemma. which together with Assumption SC3(iv), (SC.240) and (SC.241) proves the third claim of the 

where the second equality is by Assumptions SC2(vi) and SC3(iv). Collecting the results in (SC.244), (SC.246), (SC.247), (SC.248), (SC.249) and (SC.250), we get 

where the equality is by Assumptions SC1(i, iii) and SC2(i, ii), and (SB.16). Using similar argu-
ments for proving (SC.238), we can show that
\[
\max_{i \leq n} \left| \hat{g}_i(\hat{\beta}_k) - g(\nu_{1,i}(\beta_{k,0}); \beta_{k,0}) \right| = O_p(\xi_0, m_1/2 n_{1/2}) + O_p(\xi_0, m_2 + m_{1/2}) n_{1/2}) = o_p(1)
\] (SC.252)
where the second equality is by Assumption SC2(vi). By Assumption SC2(ii) and (SB.16), we have
\[
\max_{i \leq n} \left| g(\nu_{1,i}(\hat{\beta}_k); \hat{\beta}_k) - g(\nu_{1,i}(\beta_{k,0}); \beta_{k,0}) \right| = O_p(n_{1/2})
\] which together with (SC.252) shows that
\[
n^{-1} \sum_{i=1}^{n} (\hat{g}_i(\hat{\beta}_k) - g(\nu_{1,i}(\beta_{k,0}); \beta_{k,0}))^4 = o_p(1).
\] (SC.253)
The claim of the lemma follows from (SC.251) and (SC.253).

(v) Let \( \hat{\beta}_{a_2} \equiv (P_1' P_1)^{-1} \sum_{i=1}^{n} P_1(x_{1,i}) k_{2,i} \). By Assumptions SC1 and SC4(ii), we can use similar arguments for proving (SC.55) to show
\[
\hat{\beta}_{a_2} - \beta_{a_2, m_1} = O_p(m_{1/2} n_{1/2}^{-1/2} + m_{1/2}^{-r_a}).
\] (SC.254)
Therefore by the triangle inequality, Assumption SC1(vi) and (SC.254),
\[
\max_{i \leq n} |\hat{\varsigma}_{2,i} - \varsigma_{2,i}| \leq \xi_{0, m_1} \left\| \hat{\beta}_{a_2} - \beta_{a_2, m_1} \right\| + \max_{i \leq n} \left| a_{2,m_1}(x_{1,i}) - a_2(x_{1,i}) \right|
\] = \( O_p(\xi_{0, m_1} m_{1/2} n_{1/2}^{-1/2} + \xi_{0, m_1} m_{1/2}^{-r_a}) = o_p(1) \)
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) \( \hat{\Omega}_n - \Omega = o_p(1) \);
(iii) \( \hat{\Omega}_n - \Omega = o_p(1) \).

**Proof of Lemma SC25.** (i) By Assumptions SC1(i) and SC2(i, ii), and the Markov inequality
\[
n^{-1} \sum_{i=1}^{n} \varsigma_{1,i}^2 = \Upsilon + O_p(n^{-1/2}) = O_p(1)
\] (SC.255)
which together with Lemma SC24(iii) proves the first claim of the lemma.

(ii) Let \( \hat{\Gamma}_n = \sum_{i=1}^{n} ((l_{2,i} - l_{1,i} g_1(\nu_1,i)) \varsigma_{1,i} + l_{1,i} g_1(\nu_1,i) \varsigma_{2,i}) \). Then by Assumptions SC1(ii) and
(SC 2(i, ii), and the Markov inequality, we have
\[ E \left[ l_{1,i}^4 + l_{2,i}^4 + \varsigma_{1,i}^4 + \varsigma_{2,i}^4 + g_1(\nu_{1,i})^4 \right] \leq C \quad \text{(SC.256)} \]

which together with Assumption (SC 1(i) and the Markov inequality implies that
\[
\frac{1}{n} \sum_{i=1}^{n} \left[ (l_{2,i} - l_{1,i} g_1(\nu_{1,i})) \varsigma_{1,i} + l_{1,i} g_1(\nu_{1,i}) \varsigma_{2,i} \right] = E \left[ (l_{2,i} - l_{1,i} g_1(\nu_{1,i})) \varsigma_{1,i} + l_{1,i} g_1(\nu_{1,i}) \varsigma_{2,i} \right] + O_p(n^{-1/2}) \quad \text{(SC.257)}
\]

Therefore
\[ \hat{\Gamma}_n = \Gamma + o_p(1). \quad \text{(SC.258)} \]

By the definition of \( \hat{\Gamma}_n \), we can write
\[
\hat{\Gamma}_n - \tilde{\Gamma}_n = -n^{-1} \sum_{i=1}^{n} l_{1,i} (\hat{g}_1 - g_1)(\hat{\varsigma}_{1,i} - \varsigma_{1,i}) - n^{-1} \sum_{i=1}^{n} l_{1,i} (\hat{g}_1 - g_1) \varsigma_{1,i}
\]
\[ + n^{-1} \sum_{i=1}^{n} (l_{2,i} - l_{1,i} g_1) (\hat{\varsigma}_{1,i} - \varsigma_{1,i}) + n^{-1} \sum_{i=1}^{n} l_{1,i} (g_1 \varsigma_{2,i} - \hat{g}_1 \hat{\varsigma}_{2,i}). \quad \text{(SC.259)}
\]

The second claim of the lemma follows from Assumption (SC 1(i), Lemma (SC 24(ii, iii, v), (SC.256), (SC.258) and (SC.259).

(iii) Since \( \hat{\eta}_{1,i} = \eta_{1,i} - l_{1,i} (\hat{\beta}_t - \beta_{t,0}) - (\hat{\phi}_i - \phi_i) \), by the Markov inequality, Assumptions (SC 1(i, iii) and (SC 2), Lemma (SC 4) and (SC.256),
\[
n^{-1} \sum_{i=1}^{n} (\hat{\eta}_{1,i} - \eta_{1,i})^4 \leq C (\hat{\beta}_t - \beta_{t,0})^4 n^{-1} \sum_{i=1}^{n} l_{1,i}^4 + \max_{i \leq n} (\hat{\phi}_i - \phi_i)^2 n^{-1} \sum_{i=1}^{n} (\hat{\phi}_i - \phi_i)^2
\]
\[ = O_p(n^{-2}) + O_p(\xi_{0,m_1}^2 m_1^2 n^{-2}) = O_p(\xi_{0,m_1}^2 m_1^2 n^{-2}) = o_p(1). \quad \text{(SC.260)}
\]

By Assumption (SC 2(ii) and Lemma (SC 24(ii)
\[
\max_{i \leq n} g_{1,i}^4 \leq C \max_{i \leq n} (\hat{g}_1 - g_1)^4 + C \max_{i \leq n} g_{1,i}^4 = O_p(1). \quad \text{(SC.261)}
\]

65
By Assumption [SC1(i, ii), Lemma [SC24(ii, iv), (SC.260) and (SC.261)], we get

$$n^{-1} \sum_{i=1}^{n} (\hat{u}_{2,i} - \hat{\eta}_{1,i} \hat{g}_{1,i} - u_{2,i} + \eta_{1,i} g_{1,i})^4 \leq C n^{-1} \sum_{i=1}^{n} (\hat{u}_{2,i} - u_{2,i})^4 + C \max_{i \leq n} \hat{g}_{1,i} n^{-1} \sum_{i=1}^{n} (\hat{\eta}_{1,i} - \eta_{1,i})^4 + C \max_{i \leq n} (\hat{g}_{1,i} - g_{1,i})^4 n^{-1} \sum_{i=1}^{n} \hat{\eta}_{1,i} = o_p(1)$$  

(SC.262)

which together with Assumption [SC1(i, ii) and SC2(i, ii), Lemma [SC24(iii), (SC.88) and (SC.256)] implies that

$$n^{-1} \sum_{i=1}^{n} ((\hat{u}_{2,i} - \hat{\eta}_{1,i} \hat{g}_{1,i}) \hat{\varsigma}_{1,i} - (u_{2,i} - \eta_{1,i} g_{1,i}) \varsigma_{1,i})^2 = o_p(1).$$  

(SC.263)

By Assumptions [SC1(i, ii, iii) and SC4(i), and (SC.260), we have

$$n^{-1} \sum_{i=1}^{n} \hat{\varepsilon}_{1,i}^4 + n^{-1} \sum_{i=1}^{n} \hat{\eta}_{1,i}^2 = O_p(1)$$  

(SC.264)

which combined with Lemma [SC25(ii), (SC.260) and Assumption [SC4(i)] implies that

$$n^{-1} \sum_{i=1}^{n} (\hat{\Gamma}_{1,i} \hat{\varepsilon}_{1,i} - \Gamma \varepsilon_{1,i} \eta_{1,i})^2 \leq C (\hat{\Gamma}_{n} - \Gamma)^2 n^{-1} \sum_{i=1}^{n} \hat{\varepsilon}_{1,i}^2 \hat{\eta}_{1,i}^2 + C T^2 n^{-1} \sum_{i=1}^{n} (\hat{\varepsilon}_{1,i} - \varepsilon_{1,i})^2 \hat{\eta}_{1,i}^2 + C T^2 n^{-1} \sum_{i=1}^{n} \varepsilon_{1,i}^2 (\hat{\eta}_{1,i} - \eta_{1,i})^2 = o_p(1).$$  

(SC.265)

By Assumptions [SC1(i, ii) and SC2(ii), Lemma [SC24(ii, v) and (SC.256), we have

$$n^{-1} \sum_{i=1}^{n} (\hat{\eta}_{1,i} \hat{g}_{1,i} \hat{\varsigma}_{2,i} - \eta_{1,i} g_{1,i} \varsigma_{2,i})^2 \leq \max_{i \leq n} \hat{g}_{1,i} n^{-1} \sum_{i=1}^{n} (\hat{\eta}_{1,i} - \eta_{1,i})^2 \hat{\varsigma}_{2,i}^2 + \max_{i \leq n} |\hat{g}_{1,i} - g_{1,i}| n^{-1} \sum_{i=1}^{n} \eta_{1,i}^2 \hat{\varsigma}_{2,i}^2 + \max_{i \leq n} |\hat{\varsigma}_{2,i} - \varsigma_{2,i}| n^{-1} \sum_{i=1}^{n} \eta_{1,i}^2 \hat{g}_{1,i}^2 = o_p(1).$$  

(SC.266)

Let $\hat{\Omega}_{n} = n^{-1} \sum_{i=1}^{n} ((u_{2,i} - \eta_{1,i} g_{1,i}) \varsigma_{1,i} - \Gamma \varepsilon_{1,i} + \eta_{1,i} g_{1,i} \varsigma_{2,i})^2$. Then by Assumptions [SC1(i) and
\text{(SC.267)}

\[
\tilde{\Omega}_n = \Omega + O_p(n^{-1/2}).
\]

By the definition of \(\tilde{\Omega}_n\) and \(\hat{\Omega}_n\), the triangle inequality and the Cauchy-Schwarz inequality, \((\text{SC.263})\), \((\text{SC.265})\) and \((\text{SC.266})\), we get

\[
\left| \hat{\Omega}_n - \tilde{\Omega}_n \right| \leq Cn^{-1} \sum_{i=1}^{n} \left( (\hat{u}_{2,i} - \hat{\eta}_{1,i}\hat{g}_{1,i})\hat{s}_{1,i} - (u_{2,i} - \eta_{1,i}g_{1,i})s_{1,i} \right)^2
\]

\[
+ Cn^{-1} \sum_{i=1}^{n} \left( \hat{\Gamma}_{n,1,i}\hat{\eta}_{1,i} - \Gamma\varepsilon_{1,i}\eta_{1,i} \right)^2 + Cn^{-1} \sum_{i=1}^{n} \left( \hat{\eta}_{1,i}\hat{g}_{1,i}\hat{s}_{2,i} - \eta_{1,i}g_{1,i}s_{2,i} \right)^2
\]

\[
+ C\hat{\Omega}_n^{1/2} \left( n^{-1} \sum_{i=1}^{n} \left( (\hat{u}_{2,i} - \hat{\eta}_{1,i}\hat{g}_{1,i})\hat{s}_{1,i} - (u_{2,i} - \eta_{1,i}g_{1,i})s_{1,i} \right)^2 \right)^{1/2}
\]

\[
+ C\hat{\Omega}_n^{1/2} \left( n^{-1} \sum_{i=1}^{n} \left( \hat{\Gamma}_{n,1,i}\hat{\eta}_{1,i} - \Gamma\varepsilon_{1,i}\eta_{1,i} \right)^2 \right)^{1/2}
\]

\[
+ C\hat{\Omega}_n^{1/2} \left( n^{-1} \sum_{i=1}^{n} \left( \hat{\eta}_{1,i}\hat{g}_{1,i}\hat{s}_{2,i} - \eta_{1,i}g_{1,i}s_{2,i} \right)^2 \right)^{1/2} = o_p(1)
\]

which together with \((\text{SC.267})\) proves the third claim of the Lemma. \(Q.E.D.\)

\section{SC.5 Partially linear regression}

In this subsection, we provide the preliminary estimator of \(\hat{\beta}_i\) when \(\beta_{i,0}\) is estimated together with \(\phi(\cdot)\) in the partially linear regression proposed in \text{[Olley and Pakes (1996)]}. Define \(\tilde{x}_{1,i} \equiv (l_{1,i}, i_{1,i}, k_{1,i})'\) and \(P_1(\tilde{x}_{1,i}) \equiv (l_{1,i}, P_1(x_{1,i}))'\). Let \(\hat{\beta}_i\) and \(\hat{\beta}_{\phi pl}\) be the first element and the last \(m_1\) elements of \(\hat{\beta}_1\) respectively, where

\[
\hat{\beta}_1 \equiv (P_1 P_1)^{-1}(P_1 Y_1)
\]

where \(P_1 \equiv (P_1(\tilde{x}_{1,1}), \ldots, P_1(\tilde{x}_{1,n}))'\) and \(Y_1 \equiv (y_{1,1}, \ldots, y_{1,n})'\). The unknown function \(\phi(\cdot)\) is estimated by \(\hat{\phi}_{pl}(\cdot) \equiv P_1(\cdot)'\hat{\beta}_{\phi pl}\).

Let \(\hat{Q}_{m_1} \equiv \mathbb{E} [P_1(\tilde{x}_{1,1})P_1(\tilde{x}_{1,1})']\) and \(h_1(x_{1,i}) \equiv \mathbb{E}[l_{1,i}\mid x_{1,i}]\). The following assumptions are needed.

\textbf{Assumption SC5.} (i) there exist \(r_h > 0\) and \(\beta_{h_{1,m}} \in \mathbb{R}^{m}\) such that \(\sup_{x \in \mathcal{X}} |h_{1,m}(x) - h_1(x)| = O(m^{-r_h})\) where \(h_{1,m}(\cdot) \equiv P_1(\cdot)'\beta_{h_{1,m}}\) and \(n^{1/2}m_1^{-r_h} = O(1)\); (ii) \(C^{-1} \leq \lambda_{\min}(\hat{Q}_{m_1})\) uniformly over \(m_1\).

Assumption \textbf{SC5(i)} the unknown function \(h_1(x_{1,i})\) can be well approximated by the approxi-
matting functions $P_1(x_{1,i})$. Assumption SC5(ii) imposes a uniform lower bound on the eigenvalues of $\overline{Q}_{m_1}$. This condition implicitly imposes a identification condition on the unknown parameter $\beta_{l,0}$. That is in Lemma SC28 below, we show that

$$\|l_{1,i} - h_1(x_{1,i})\|_2 \geq C^{-1}$$  \hspace{1cm} (SC.268)

which together with (SA.1) implies that

$$\beta_{l,0} = \frac{E[(l_{1,i} - h_1(x_{1,i}))(y_{1,i} - E[y_{1,i} | x_{1,i}])]}{E[|l_{1,i} - h_1(x_{1,i})|^2]}.$$  \hspace{1cm} (SC.269)

We shall show below that Assumption SC1(iii) holds

$$\varepsilon_{1,i} = \frac{l_{1,i} - h_1(x_{1,i})}{E[|l_{1,i} - h_1(x_{1,i})|^2]} \eta_{1,i}.$$  \hspace{1cm} (SC.270)

Let $\hat{h}_{1,i} \equiv P_1(x_{1,i})' \left( \mathbf{P}_1' \mathbf{P}_1 \right)^{-1} \mathbf{P}_1' \mathbf{L}_1$ where $\mathbf{L}_1 \equiv (l_{1,1}, \ldots, l_{1,n})'$. Then $\varepsilon_{1,i}$ can be estimated by

$$\hat{\varepsilon}_{1,i} = \frac{l_{1,i} - \hat{h}_{1,i}}{n^{-1} \sum_{i=1}^{n} (l_{1,i} - \hat{h}_{1,i})^2} \hat{\eta}_{1,i}.$$  \hspace{1cm} (SC.271)

where $\hat{\eta}_{1,i} \equiv y_{1,i} - l_{1,i} \hat{\beta}_l - \hat{\phi}(x_{1,i})$ is defined in Subsection SB.2

**Lemma SC26.** Under Assumptions SC1(i, ii, iv, v, vi) and SC5 we have

$$\hat{\beta}_l - \beta_{l,0} = n^{-1} \sum_{i=1}^{n} \frac{l_{1,i} - h_1(x_{1,i})}{E[|l_{1,i} - h_1(x_{1,i})|^2]} \eta_{1,i} + o_p(n^{-1/2}).$$  \hspace{1cm} (SC.272)

**Proof of Lemma SC26** First note that we can write $\hat{\beta}_l = (\mathbf{L}_1' \mathbf{M}_1 \mathbf{L}_1)^{-1} (\mathbf{L}_1' \mathbf{M}_1 \mathbf{Y}_1)$. Therefore

$$\hat{\beta}_l - \beta_{l,0} = (\mathbf{L}_1' \mathbf{M}_1 \mathbf{L}_1)^{-1} (\mathbf{L}_1' \mathbf{M}_1 \mathbf{Y}_1 - \beta_{l,0} \mathbf{L}_1)$$

$$= n^{-1} \sum_{i=1}^{n} \frac{l_{1,i} - h_1(x_{1,i})}{E[|l_{1,i} - h_1(x_{1,i})|^2]} \eta_{1,i} + O_p(m_1 n^{-1})$$  \hspace{1cm} (SC.273)

where the second equality is by Lemma SC30(i, ii). The claim in (SC.272) follows by Assumption SC1(vi) and (SC.273).

**Lemma SC27.** Under Assumption SC1(ii, v), we have $\lambda_{\max}(\overline{Q}_{m_1}) \leq C$.

**Proof of Lemma SC27** Consider any $b \equiv (b_1, b_2')' \in \mathbb{R}^{m_1+1}$ with $b'b = 1$ where $b_2 \in \mathbb{R}^{m_1}$. Then

$$b' \overline{Q}_{m_1} b = b_1^2 E[l_{1,i}^2] + 2b_1 b'E[P_1(x_{1,i}) l_{1,i}] + b'Q_{m_1}b \leq C + 2b_1 b'E[P_1(x_{1,i}) l_{1,i}]$$  \hspace{1cm} (SC.274)
where the second inequality is by Assumption SC1(ii, v). Moreover by Assumption SC1(ii, v)
\[ \|E[P_1(x_{1,i})l_{1,i}]\| \leq E[l_{1,i}^2] \leq C \]
which together with the Cauchy-Schwarz inequality and (SC.274) implies that
\[ b'Q_m b \leq C. \]
Q.E.D.

Lemma SC28. Under Assumptions SC1(ii, v) and SC5, we have
\[ \|l_{1,i} - h_1(x_{1,i})\| \geq C_0. \]

Proof of Lemma SC28. By Assumption SC5(ii), there exists a fixed \( m_c \) such that
\[ \sup_{x \in X} |h_1,m(x) - h_1(x)| \leq (2C)^{-1} \] (SC.275)
for any \( m \geq m_c \). Consider any \( m \geq m_c \). By the triangle inequality and (SC.275)
\[ \|l_{1,i} - h_1(x_{1,i})\| \geq \|l_{1,i} - h_1,m(x)\| - (2C^{1/2})^{-1}. \] (SC.276)
Let \( \beta_{h_1,m}^* \equiv Q_m^{-1}E[P_1(x_{1,i})l_{1,i}] \). Then \( P_1(x_{1,i})' \beta_{h_1,m}^* \) is the projection of \( l_{1,i} \) on \( P_1(x_{1,i}) \) under the \( L_2 \)-norm. Therefore
\[ \|l_{1,i} - h_1,m(x_{1,i})\| \geq \|l_{1,i} - P_1(x_{1,i})' \beta_{h_1,m}^*\| \geq (\lambda_{\min}(Q_m))^{1/2} \geq C^{1/2} \]
which together with (SC.276) finishes the proof. Q.E.D.

Lemma SC29. Under Assumption SC1(i, ii, v, vi), we have
\[ \left\| n^{-1} \bar{P}_1' \bar{P}_1 - Q_m \right\|_S = O_p((\log m_1)^{1/2} \xi_0,m_1^{-1/2}) = o_p(1) \] (SC.277)
and
\[ C^{-1} \leq \lambda_{\min}(n^{-1} \bar{P}_1' \bar{P}_1) \leq \lambda_{\max}(n^{-1} \bar{P}_1' \bar{P}_1) \leq C \] wpa1. (SC.278)

Proof of Lemma SC29. By Assumption SC1(i, ii, v) and the Markov inequality, we have
\[ n^{-1} \sum_{i=1} l_{1,i}^2 = O_p(n^{-1/2}) \] (SC.279)
and
\[ n^{-1} \sum_{i=1} P_{1,i}l_{1,i} - E[P_{1,i}l_{1,i}] = O_p(m_1^{1/2} n^{-1/2}). \] (SC.280)
Let \( A_{11,n} = n^{-1} \sum_{i=1} l_{1,i}^2, A_{12,n} = n^{-1} \sum_{i=1} l_{1,i}P_{1,i} \) and \( A_{22,n} = n^{-1} \bar{P}_1' \bar{P}_1 \). Consider any
\(b = (b_1, b'_2)^\prime \in \mathbb{R}^{m_1 + 1}\) with \(bb' = 1\) where \(b_2 \in \mathbb{R}^{m_1}\). By the Cauchy-Schwarz inequality,

\[
\left( \begin{array}{c} b_1 \\ b_2 \end{array} \right) \left( \begin{array}{cc} A_{11,n} - \mathbb{E}[A_{11,n}] & A_{12,n} - \mathbb{E}[A_{12,n}] \\ A_{21,n} - \mathbb{E}[A_{21,n}] & A_{22,n} - \mathbb{E}[A_{22,n}] \end{array} \right)^2 \left( \begin{array}{c} b_1 \\ b_2 \end{array} \right)
\leq C \left( \|A_{11,n} - \mathbb{E}[A_{11,n}]\|^2 + \|A_{12,n} - \mathbb{E}[A_{12,n}]\|^2 + \|A_{22,n} - \mathbb{E}[A_{22,n}]\|^2 \right)
\]

which combined with (SC.46), (SC.279) and (SC.280) implies that

\[
\|n^{-1} \mathbf{P}'_1 \mathbf{P}_1 - \overline{Q}_{m_1}\|_S^2 = O_p((\log m_1)^{1/2} (\xi_0, m_1 + m_1^{1/2}) n^{-1/2}).
\] (SC.281)

By (SC.281) and Assumption SC1(vi), we have

\[
\|n^{-1} \mathbf{P}'_1 \mathbf{P}_1 - \overline{Q}_{m_1}\|_S = o_p(1)
\] (SC.282)

which together with Assumption SC1(v) proves (SC.278). Q.E.D.

**Lemma SC30.** Let \(\mathbf{M}_1 \equiv \mathbf{I}_n - \mathbf{P}_1 (\mathbf{P}'_1 \mathbf{P}_1)^{-1} \mathbf{P}'_1\). Under Assumptions SC1(i, ii, iv, v, vi) and SC5, we have

(i) \(n^{-1} \mathbf{L}'_1 \mathbf{M}_1 \mathbf{L}_1 = \mathbb{E}[l_{1,i} - h_1(x_{1,i})]^2 + O_p(m_1^{1/2} n^{-1/2});\)

(ii) \(n^{-1} \mathbf{L}'_1 \mathbf{M}_1 (\mathbf{Y}_1 - \mathbf{L}_1 \beta_{l,0}) = n^{-1} \sum_{i=1}^n (l_{1,i} - h_1(x_{1,i})) \eta_{1,i} + o_p(n^{-1/2}).\)

**Proof of Lemma SC30**

(i) By Assumption SC1(ii) and Hölder’s inequality,

\[
h_1^2(x_{1,i}) = (\mathbb{E}[l_{1,i} | x_{1,i}])^2 \leq \mathbb{E}[l_{1,i}^2 | x_{1,i}] \leq C
\] (SC.283)

which together with Assumption SC1(ii) implies that

\[
\mathbb{E}[\epsilon_{1,i}^2 | x_{1,i}] \leq 2 \mathbb{E}[l_{1,i}^2 | x_{1,i}] + 2h_1^2(x_{1,i}) \leq C
\] (SC.284)

where \(\epsilon_{1,i} \equiv l_{1,i} - h_1(x_{1,i})\). Let \(\hat{\beta}_{h_1} \equiv (\mathbf{P}'_1 \mathbf{P}_1)^{-1} \mathbf{P}'_1 \mathbf{L}_1\). Then

\[
\hat{\beta}_{h_1} - \beta_{h_1,m_1} = (\mathbf{P}'_1 \mathbf{P}_1)^{-1} \sum_{i=1}^n P_{1,i} \epsilon_{1,i} + (\mathbf{P}'_1 \mathbf{P}_1)^{-1} \sum_{i=1}^n P_{1,i} (h_1(x_{1,i}) - h_{1,m_1}(x_{1,i})).
\]
Therefore by Assumptions SC1(i, v) and SC5(i), (SC.47) and (SC.284), we obtain

\[ \| \hat{\beta}_{h_1} - \beta_{h_1,m_1} \|^2 \]
\[ \leq 2 \left( \sum_{i=1}^{n} \epsilon_{i,P_1(x_{i,1})} \right) (P_1'P_1)^{-2} \left( \sum_{i=1}^{n} P_1(x_{i,1}) \epsilon_{i,i} \right) \]
\[ + 2 \left( \sum_{i=1}^{n} (h_{1,m_1}(x_{i,1}) - h_1(x_{i,1})) P_1(x_{i,1})' \right) (P_1'P_1)^{-2} \left( \sum_{i=1}^{n} P_1(x_{i,1}) (h_{1,m_1}(x_{i,1}) - h_1(x_{i,1})) \right) \]
\[ \leq \frac{2}{\lambda_{\min}(n^{-1}P_1'P_1)^2} \left( \sum_{i=1}^{n} (h_{1,m_1}(x_{i,1}) - h_1(x_{i,1}))^2 \right) + \frac{2}{n\lambda_{\min}(n^{-1}P_1'P_1)} = O_p(m_1^{-1}) \] (SC.285)

which together with Assumption SC5(i) and (SC.47) further implies that

\[ n^{-1} \sum_{i=1}^{n} (h_1(x_{i,1}) - \hat{h}_1(x_{i,1}))^2 = O_p(m_1^{-1}). \] (SC.286)

By (SC.283) and (SC.286)

\[ \left| n^{-1} \sum_{i=1}^{n} h_1^2(x_{i,1}) - n^{-1} \sum_{i=1}^{n} \hat{h}_1^2(x_{i,1}) \right| \]
\[ \leq n^{-1} \sum_{i=1}^{n} (h_1(x_{i,1}) - \hat{h}_1(x_{i,1}))^2 \]
\[ + \left( n^{-1} \sum_{i=1}^{n} h_1^2(x_{i,1}) \right)^{1/2} \left( n^{-1} \sum_{i=1}^{n} (h_1(x_{i,1}) - \hat{h}_1(x_{i,1}))^2 \right)^{1/2} = O_p(m_1^{1/2}n^{-1/2}), \] (SC.287)

Therefore by the Markov inequality, Assumption SC1(i, ii), (SC.283) and (SC.287)

\[ n^{-1}L_1'M_1L_1 - (E[l_{1,i}^2] - E[h_1(x_{i,1})^2]) \]
\[ = n^{-1} \sum_{i=1}^{n} (l_{1,i}^2 - E[l_{1,i}^2]) + n^{-1} \sum_{i=1}^{n} (h_1^2(x_{i,1}) - \hat{h}_1^2(x_{i,1})) \]
\[ - n^{-1} \sum_{i=1}^{n} (h_1(x_{i,1})^2 - E[h_1(x_{i,1})^2]) = O_p(m_1^{1/2}n^{-1/2}). \] (SC.288)

Since \( E[l_{1,i} - h_1(x_{i,1})^2] = E[h_1(x_{i,1})^2], \) the first claim of the lemma follows from (SC.288).

(ii) Since \( Y_1 - L_1 \beta_{l,0} = \phi + \eta_1 \) where \( \phi \equiv (\phi_1, \ldots, \phi_n)' \) and \( \eta_1 \equiv (\eta_{1,1}, \ldots, \eta_{1,n})' \), we can write

\[ n^{-1}L_1'M_1(Y_1 - L_1 \beta_{l,0}) = n^{-1}L_1'M_1\phi + n^{-1}L_1'M_1\eta_1, \] (SC.289)

Let \( \phi_{m_1} \equiv (\phi_{m_1}(x_{1,1}), \ldots, \phi_{m_1}(x_{1,n}))' \). Then \( \phi_{m_1} = P_1\beta_{m_1} \) and \( M_1P_1 = 0 \). Therefore by
Assumption SC1(iv, vi) and (SC.47)

\[ n^{-1} \phi' M_1 \phi = n^{-1} (\phi - \phi_{m_1})' M_1 (\phi - \phi_{m_1}) \leq n^{-1} \sum_{i=1}^n (\phi(x_{1,i}) - \phi_{m_1}(x_{1,i}))^2 = O(n^{-1}). \]  
(SC.290)

Let \( h_1 \equiv (h_1(x_{1,1}), \ldots, h_1(x_{1,n}))' \). Then by the similar arguments of showing (SC.290), we get

\[ n^{-1} h_1' M_1 h_1 = O(n^{-1}). \]  
(SC.291)

By Assumption SC1(i), (SC.284) and (SC.290)

\[ \mathbb{E} \left[ \left\| n^{-1} \epsilon_1' M_1 \phi \right\|^2 \mid \{x_{1,i}\}_{i=1}^n \right] = n^{-2} \phi' M_1 \mathbb{E} [\epsilon_1 \epsilon_1' \mid \{x_{1,i}\}_{i=1}^n] M_1 \phi \leq C n^{-2} \phi' M_1 \phi = O(n^{-2}) \]

which together with the Markov inequality implies that

\[ n^{-1} \epsilon_1' M_1 \phi = O_p(n^{-1}). \]  
(SC.292)

Similarly, we can show that

\[ n^{-1} h_1' M_1 \eta_1 = O_p(n^{-1}). \]  
(SC.293)

Collecting the results in (SC.289), (SC.290), (SC.291), (SC.292) and (SC.293), we obtain

\[ n^{-1} L_1' M_1 (Y_1 - L_1 \beta_{l,0}) = n^{-1} \epsilon_1' M_1 \eta_1 + O_p(n^{-1}). \]  
(SC.294)

Since \( n^{-1} \sum_{i=1}^n P_{1,i} \epsilon_{1,i} = O_p(m_1^{1/2} n^{-1/2}) \) and \( n^{-1} \sum_{i=1}^n P_{1,i} \eta_{1,i} = O_p(m_1^{1/2} n^{-1/2}) \) by Assumption SC1(i, ii), (SC.284) and the Markov inequality, we can use (SC.47) to deduce that

\[ n^{-1} \epsilon_1' P_1 P_1' \eta_1 = O_p(m_1 n^{-1}) \]

which together with (SC.294) proves the second claim of the lemma. \( Q.E.D. \)

SC.6 Preliminary results

Lemma SC31 (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 $\xi$ is a finite constant. Introduce the random matrix $D_n = \sum_{i=1}^n d_i$. Compute the variance parameter

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

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 SC32.** Let $S_{2,i}(\beta_k) = \tilde{P}_{2,i}(\beta_k) \tilde{P}_{2,i}(\beta_k)'$ where $\tilde{P}_{2,i}(\beta_k) = \tilde{P}_2(\nu_1,i(\beta_k), \beta_k)$ for any $\beta_k \in \Theta_k$. Then under Assumptions SC1(i) and SC2(iv, v), 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 SC32.** For any $\beta_k \in \Theta_k$, by the triangle inequality and Assumptions SC2(iv, v),

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

(SC.295)

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

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

(SC.296)

Therefore we can use Lemma SC31 to deduce that

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

(SC.297)

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 $b \in \mathbb{R}^{m_2}$ with $\|b\| = 1$. By the triangle inequality,

$$
\|S_{2,i}(\beta_{k,1}) - S_{2,i}(\beta_{k,2})\|_S \leq \left\| S_{2,i}(\beta_{k,1}) - \tilde{P}_{2,i}(\beta_{k,2}) \tilde{P}_{2,i}(\beta_{k,1})' \right\|_S + \left\| \tilde{P}_{2,i}(\beta_{k,2}) \tilde{P}_{2,i}(\beta_{k,1})' - S_{2,i}(\beta_{k,2}) \right\|_S.
$$

(SC.298)
By the mean value expansion and the Cauchy-Schwarz inequality, and Assumption \( \text{SC2(v)} \)

\[
|b'(\tilde{P}_{2,i}(\beta_{k,1})\tilde{P}_{2,i}(\beta_{k,1})' - \tilde{P}_{2,i}(\beta_{k,2})\tilde{P}_{2,i}(\beta_{k,1})')|^2
\]

\[
= |\tilde{P}_{2,i}(\beta_{k,1})|^2 |b'(\tilde{P}_{2,i}(\beta_{k,1}) - \tilde{P}_{2,i}(\beta_{k,2}))|^2
\]

\[
= |\tilde{P}_{2,i}(\beta_{k,1})|^2 b'\tilde{P}_2 \left( \nu_{i,1}(\tilde{\beta}_{k,12}); \tilde{\beta}_{k,12} \right) / \partial \beta_k (\beta_{k,1} - \beta_{k,2})^2
\]

\[
\leq \|b\|^2 \epsilon_{i,m_{2}}^{2} (\beta_{k,1} - \beta_{k,2})^2
\]

where \( \tilde{\beta}_{k,12} \) lies between \( \beta_{k,1} \) and \( \beta_{k,2} \), which together with Assumption \( \text{SC2(vi)} \) implies that

\[
\left\| S_{2,i}(\beta_{k,1}) - S_{2,i}(\beta_{k,2}) \right\| \leq Cm_2^2 |\beta_{k,2} - \beta_{k,1}|. \quad (\text{SC.299})
\]

The same upper bound can be established for the second term in the right hand side of the inequality of \( \text{SC.298} \). Therefore,

\[
\left\| S_{2,i}(\beta_{k,1}) - S_{2,i}(\beta_{k,1}) \right\| \leq Cm_2^2 |\beta_{k,2} - \beta_{k,1}|. \quad \text{(SC.300)}
\]

Similarly, we can show that

\[
\| \mathbb{E} [S_{2,i}(\beta_{k,1})] - \mathbb{E} [S_{2,i}(\beta_{k,2})] \| \leq Cm_2^2 |\beta_{k,2} - \beta_{k,1}|. \quad \text{(SC.301)}
\]

Combining the results in \( \text{SC.300} \) and \( \text{SC.301} \), and applying the triangle inequality, we get

\[
\left\| n^{-1} \sum_{i=1}^{n} (S_{2,i}(\beta_{k,1}) - \mathbb{E} [S_{2,i}(\beta_{k,1})]) - n^{-1} \sum_{i=1}^{n} (S_{2,i}(\beta_{k,2}) - \mathbb{E} [S_{2,i}(\beta_{k,2})]) \right\| \leq C_{m_2} m_2^3 |\beta_{k,2} - \beta_{k,1}| \quad \text{(SC.302)}
\]

where \( C_{m_2} \) is a finite fixed constant. Since the parameter space \( \Theta_{k} \) is compact, there exist \( \{\beta_{k}(l)\}_{l=1,...,K_n} \)

such that for any \( \beta_{k} \in \Theta_{k} \)

\[
\min_{l=1,...,K_n} |\beta_{k} - \beta_{k}(l)| \leq (C_{m_2} m_2^{3} n^{1/2})^{-1} \quad \text{(SC.303)}
\]

where \( K_n \leq 2C_{m_2} m_2^{3} n^{1/2} \). For any \( \beta_{k} \in \Theta_{k} \), by \( \text{SC.302} \) and \( \text{SC.303} \)

\[
\left\| n^{-1} \sum_{i=1}^{n} S_{2,i}(\beta_{k}) - \mathbb{E} [S_{2,i}(\beta_{k})] \right\| \leq \max_{l=1,...,K_n} \left\| n^{-1} \sum_{i=1}^{n} S_{2,i}(\beta_{k}(l)) - \mathbb{E} [S_{2,i}(\beta_{k}(l))] \right\| + n^{-1/2}. \quad \text{(SC.304)}
\]
Therefore for any $B > 1,$

$$
P \left( \sup_{\beta_k \in \Theta_k} \left\| n^{-1} \sum_{i=1}^{n} S_{2,i}(\beta_k) - \mathbb{E} [S_{2,i}(\beta_k)] \right\| \geq B(\xi_{0,m_2}^2 \log(n)n^{-1/2}) \right) \\
\leq P \left( \max_{l=1, \ldots, K_n} \left\| n^{-1} \sum_{i=1}^{n} S_{2,i}(\beta_k(l)) - \mathbb{E} [S_{2,i}(\beta_k(l))] \right\| \geq (B - 1)(\xi_{0,m_2}^2 \log(n)n^{-1/2}) \right) \\
\leq \sum_{l=1}^{K_n} P \left( \left\| n^{-1} \sum_{i=1}^{n} S_{2,i}(\beta_k(l)) - \mathbb{E} [S_{2,i}(\beta_k(l))] \right\| \geq (B - 1)(\xi_{0,m_2}^2 \log(n)n^{-1/2}) \right) \\
\leq 2K_n m_2 \exp \left( - \frac{B}{C} \frac{\log(n)}{1 + (\xi_{0,m_2}^2 \log(n)n^{-1/2})} \right) \quad \text{(SC.305)}
$$

where the last inequality is by (SC.297). The claim of the theorem follows from (SC.305) and Assumption SC2[vi]. \[ Q.E.D. \]

**Lemma SC33.** Let $w_{2,i}(\beta_k) = y_{2,i}^* - k_{2,i} \beta_k - g(\nu_{1,i}(\beta_k), \beta_k)$. Then under Assumptions SC1 and SC2(ii, iii, v, vi), we have

$$
\sup_{\beta_k \in \Theta_k} \left\| n^{-1} \sum_{i=1}^{n} \tilde{P}_{2,i}(\beta_k)w_{2,i}(\beta_k) \right\| = O_p(n_2^{5/4}n^{-1/2}).
$$

**Proof of Lemma SC33.** Define $\pi_n(\beta_k) = n^{-1/2} \sum_{i=1}^{n} \tilde{P}_{2,i}(\beta_k)w_{2,i}(\beta_k)$. For any $\beta_k \in \Theta_k$, by Assumption SC2(i) and (SC.68),

$$
\mathbb{E} \left[ (u_{2,i}(\beta_k))^4 \mid \nu_{1,i}(\beta_k) \right] \leq C \mathbb{E} \left[ (y_{2,i}^*)^4 + \nu_{2,i}^4 \right] \nu_{1,i}(\beta_k) + C \left| g(\nu_{1,i}(\beta_k); \beta_k) \right|^4 \leq C. \quad \text{(SC.306)}
$$

For any $\beta_{k,1}, \beta_{k,2} \in \Theta_k$, by the i.i.d. assumption and the Cauchy-Schwarz inequality

$$
\mathbb{E} \left[ \left\| \pi_n(\beta_{k,1}) - \pi_n(\beta_{k,2}) \right\|^2 \right] \\
= \mathbb{E} \left[ \left\| \tilde{P}_{2,i}(\beta_{k,1})u_{2,i}(\beta_{k,1}) - \tilde{P}_{2,i}(\beta_{k,2})u_{2,i}(\beta_{k,2}) \right\|^2 \right] \\
\leq 2\mathbb{E} \left[ (u_{2,i}(\beta_{k,1}))^2 \left\| \tilde{P}_{2,i}(\beta_{k,1}) - \tilde{P}_{2,i}(\beta_{k,2}) \right\|^2 \right] \\
+ 2\mathbb{E} \left[ \left\| \tilde{P}_{2,i}(\beta_{k,1}) \right\|^2 (u_{2,i}(\beta_{k,2}) - u_{2,i}(\beta_{k,1}))^2 \right]. \quad \text{(SC.307)}
$$

Consider any $b \in \mathbb{R}^m$. By the mean value expansion and Assumption SC2(v)

$$
\left| b' (\tilde{P}_{2,i}(\beta_{k,1}) - \tilde{P}_{2,i}(\beta_{k,2})) \right|^2 = \left| b' \partial \tilde{P}_{2,i} \left( \beta_{k,12} \right) / \partial \beta_k \right|^2 (\beta_{k,1} - \beta_{k,2})^2 \leq \| b \|^2 \xi_{1,m_2}^2 (\beta_{k,1} - \beta_{k,2})^2
$$

75
where \( \tilde{\beta}_{k,12} \) lies between \( \beta_{k,1} \) and \( \beta_{k,2} \), which implies that

\[
\left\| \tilde{P}_{2,i}(\beta_{k,1}) - \tilde{P}_{2,i}(\beta_{k,2}) \right\|^2 \leq \xi^2_{1,m2}(\beta_{k,1} - \beta_{k,2})^2. \tag{SC.308}
\]

Therefore, by (SC.306) and (SC.308),

\[
E \left[ (u_{2,i}(\beta_{k,2}))^2 \left\| \tilde{P}_{2,i}(\beta_{k,1}) - \tilde{P}_{2,i}(\beta_{k,2}) \right\|^2 \right] \leq C \xi^2_{1,m2}(\beta_{k,2} - \beta_{k,1})^2. \tag{SC.309}
\]

By the definition of \( u_{2,i}(\beta_k) \), we can write

\[
u_1,i(\beta_{k,2}) - u_1,i(\beta_{k,1}) = g(\nu_1,i(\beta_{k,1}), \beta_{k,1}) - g(\nu_1,i(\beta_{k,2}), \beta_{k,2}) + k_2,i(\beta_{k,2} - \beta_{k,1}).
\]

Therefore by Assumption SC2(i, ii, iv), we have

\[
E \left[ \left\| \tilde{P}_{2,i}(\beta_{k,1}) \right\|^2 (u_{1,i}(\beta_{k,2}) - u_{1,i}(\beta_{k,1}))^2 \right] \leq Cm_2(\beta_{k,2} - \beta_{k,1})^2 \tag{SC.310}
\]

which together with Assumption SC2(vi), (SC.307) and (SC.309) implies that

\[
\|\|\pi_n(\beta_{k,1}) - \pi_n(\beta_{k,2})\|\|_2 \leq Cm_2^2 |\beta_{k,2} - \beta_{k,1}| \tag{SC.311}
\]

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 for 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 |\( \beta_{k,j} - \beta'_{k,j} \)| > 2^{-j}. Since \( \Theta_k \) 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 \( J_n = \min\{j : 2^{-j} \leq Cm_2^{-3/2}\} \). 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 \( \beta_{k,J_n} \) in \( \Theta_{k,J_n} \). For arbitrary points \( \beta_{k,J+1}, \beta'_{k,J+1} \in \Theta_{k,J+1} \), by the triangle inequality

\[
\left\| \pi_n(\beta_{k,J+1}) - \pi_n(\beta_{k,J_n}) - [\pi_n(\beta'_{k,J+1}) - \pi_n(\beta'_{k,J_n})] \right\|
\]

\[
= \left\| \sum_{j = J_n}^{J} [\pi_n(\beta_{k,j+1}) - \pi_n(\beta_{k,j})] - \sum_{j = J_n}^{J} [\pi_n(\beta'_{k,j+1}) - \pi_n(\beta'_{k,j})] \right\|
\]

\[
\leq 2 \sum_{j = J_n}^{J} \max \left\| \pi_n(\beta_{k,j+1}) - \pi_n(\beta_{k,j}) \right\| \tag{SC.312}
\]

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 $j$th maximum is taken over at most $C2^{j+1}$ many links. By Assumption \(\text{SC2 vii}\), \(\text{SC3.11}\), \(\text{SC.312}\), the triangle inequality and the finite maximum inequality,

\[
\|\max \|\pi_n(\beta_{k,J+1}) - \pi_n(\beta_{k,J}) - [\pi_n(\beta'_{k,J+1}) - \pi_n(\beta'_{k,J})]\|_2 \leq 2 \sum_{j=J_n} \|\max \|\pi_n(\beta_{k,J+1}) - \pi_n(\beta_{k,J})\|_2
\]

\[
\leq C \sum_{j=J_n} 2^{j/2} \max \|\pi_n(\beta_{k,J+1}) - \pi_n(\beta_{k,J})\|_2 \leq Cm_2^2 \sum_{j=J_n} 2^{-j/2} \leq Cm_2^{5/4} \quad (\text{SC.313})
\]

where $\beta_{k,J}$ and $\beta'_{k,J}$ are the endpoints of the chains starting at $\beta_{k,J+1}$ and $\beta'_{k,J+1}$ respectively. Since the set $\Theta_{k,J}$ has at most $Cm_2^{3/2}$ many elements, by the finite maximum inequality, the triangle inequality, \(\text{SC.306}\) and Assumption \(\text{SC2 iv}\)

\[
\|\max \|\pi_n(\beta_{k,J}) - \pi_n(\beta'_{k,J})\|_2 \leq Cm_2^{3/4} \max \|\pi_n(\beta_{k,J})\|_2 \leq Cm_2^{5/4}. \quad (\text{SC.314})
\]

Therefore, by the triangle inequality, \(\text{SC.313}\) and \(\text{SC.314}\),

\[
\|\max \|\pi_n(\beta_{k,J+1}) - \pi_n(\beta'_{k,J+1})\|_2 \leq \|\max \|\pi_n(\beta_{k,J+1}) - \pi_n(\beta_{k,J}) - [\pi_n(\beta'_{k,J+1}) - \pi_n(\beta'_{k,J})]\|_2
\]

\[
+ \|\max \|\pi_n(\beta_{k,J}) - \pi_n(\beta'_{k,J})\|_2 \leq Cm_2^{5/4}. \quad (\text{SC.315})
\]

Let $J$ go to infinity, by \(\text{SC.315}\) we deduce that

\[
\left\| \sup_{\beta_k,\beta'_k \in \Theta_k} \|\pi_n(\beta_k) - \pi_n(\beta'_k)\|_2 \right\| \leq Cm_2^{5/4}. \quad (\text{SC.316})
\]

By \(\text{SC.314}\), \(\text{SC.316}\) and the triangle inequality,

\[
\left\| \sup_{\beta_k \in \Theta_k} \|\pi_n(\beta_k)\|_2 \right\| \leq \left\| \sup_{\beta_k \in \Theta_k} \|\pi_n(\beta_k) - \pi_n(\beta_{k,0})\|_2 \right\| + \left\| \|\pi_n(\beta_{k,0})\|_2 \right\| \leq Cm_2^{5/4} \quad (\text{SC.317})
\]

which finishes the proof. \(Q.E.D.\)

**Lemma SC34.** Under Assumptions \(\text{SC1}\) and \(\text{SC2 (ii, iii, v, vi)}\), we have

\[
\sup_{\beta_k \in \Theta_k} \left\| n^{-1} \sum_{i=1}^n u_{2,i,k_1,i} \partial^1 \hat{P}_{2,i}(\beta_k) \right\| = O_p(m_2^{5/2} n^{-1/2}).
\]

**Proof of Lemma SC34** Define $\pi_n(\beta_k) = n^{-1/2} \sum_{i=1}^n u_{2,i,k_1,i} \partial^1 \hat{P}_{2,i}(\beta_k)$ for any $\beta_k \in \Theta_k$. By
Assumptions SC1(i) and Assumption SC2(v, vi), and (SC.88)

\[ \sup_{\beta_k \in \Theta_k} \| \pi_n(\beta_k) \|_2 \leq C \xi_{1,m_2} \leq C m_2^2 \]  

(SC.318)

Moreover for any \( \beta_{k,1} \) and \( \beta_{k,2} \), we can use similar arguments in showing (SC.309) to obtain

\[ \| \pi_n(\beta_{k,1}) - \pi_n(\beta_{k,2}) \|_2 \leq C \xi_{1,m_2} |\beta_{k,1} - \beta_{k,2}| \leq C m_2^3 |\beta_{k,1} - \beta_{k,2}| \]  

(SC.319)

Consider the same nested sets \( \Theta_{k,j} \) \((j = 1, 2, \ldots)\) constructed in the proof of lemma SC34. Let \( J_n = \min\{j : 2^{-j} \leq C m_2^{-1}\} \). Then for any positive integer \( J > J_n \) using the similar arguments in the proof of Lemma SC33, we obtain

\[ \| \max \| \pi_n(\beta_{k,J+1}) - \pi_n(\beta_{k,J}) \| - \| \pi_n(\beta'_{k,J+1}) - \pi_n(\beta'_{k,J}) \| \|_2 \| \leq m_2^3 \sum_{j=J_n}^{\infty} 2^{-j/2} \leq C m_2^{5/2} \]  

(SC.320)

where \( \beta_{k,J} \) and \( \beta'_{k,J} \) are the endpoints of the chains starting at \( \beta_{k,J+1} \) and \( \beta'_{k,J+1} \) respectively. Since the set \( \Theta_{k,J_n} \) has at most \( C m_2 \) many elements, by the finite maximum inequality, the triangle inequality and (SC.318)

\[ \| \max \| \pi_n(\beta_{k,J}) - \pi_n(\beta'_{k,J}) \|_2 \| \leq C m_2^{1/2} \sup_{\beta_k \in \Theta_k} \| \pi_n(\beta_k) \|_2 \leq C m_2^{5/2} \]  

(SC.321)

Then the claim of the lemma follows by applying the chaining arguments in the proof of Lemma SC33.  

Q.E.D.

Lemma SC35. Under Assumptions SC1 and SC2(ii, iii, v, vi), we have

\[ \sup_{\beta_k \in \Theta_k} \left\| n^{-1} \sum_{i=1}^{n} u_{2,i} \partial^1 \hat{P}_{2,i}(\beta_k) P_1(x_{1,i})' \right\| = O_p(m_2^{5/2} m_1^{1/2} n^{-1/2}). \]

Proof of Lemma SC35 Define \( \pi_n(\beta_k) = n^{-1/2} \sum_{i=1}^{n} u_{2,i} \partial^1 \hat{P}_{2,i}(\beta_k) P_1(x_{1,i})' \) for any \( \beta_k \in \Theta_k \). By Assumptions SC1(i) and Assumption SC2(v, vi), and (SC.88)

\[ \sup_{\beta_k \in \Theta_k} \| \pi_n(\beta_k) \|_2 \leq C \xi_{1,m_2} m_1^{1/2} \leq C m_1^{1/2} m_2. \]  

(SC.322)

Moreover for any \( \beta_{k,1} \) and \( \beta_{k,2} \), we can use similar arguments in showing (SC.309) to obtain

\[ \| \pi_n(\beta_{k,1}) - \pi_n(\beta_{k,2}) \|_2 \leq C m_1^{1/2} m_2^3 |\beta_{k,1} - \beta_{k,2}|. \]  

(SC.323)

Consider the same nested sets \( \Theta_{k,j} \) \((j = 1, 2, \ldots)\) constructed in the proof of lemma SC34. Let
where $\beta$ inequality and (SC.326) Since the set $\Theta$ the proof of Lemma SC33, we obtain

Consider the same nested sets $\Theta_{k,j}$ Moreover for any $\beta$ Consider the same nested sets $\Theta_{k,j}$

Proof of Lemma SC36

Under Assumptions SC1 and SC2(ii, iii, v, vi), we have Lemma SC36. Define $\pi_n(\beta_k) = n^{-1/2} \sum_{i=1}^{n} u_{2,i} (\phi_{m_2}(x_{1,i}) - \phi(x_{1,i})) \partial^{1} \tilde{P}_{2,i}(\beta_k)$ for any $\beta_k \in \Theta_k$. By Assumptions SC1(i) and Assumption SC2(iii, v, vi), and (SC.88)

$$\sup_{\beta_k \in \Theta_k} \left\| n^{-1} \sum_{i=1}^{n} u_{2,i} (\phi_{m_2}(x_{1,i}) - \phi(x_{1,i})) \partial^{1} \tilde{P}_{2,i}(\beta_k) \right\| = O_p(m_2^{5/2} n^{-1}).$$

Proof of Lemma SC36

Define $\pi_n(\beta_k) = n^{-1/2} \sum_{i=1}^{n} u_{2,i} (\phi_{m_2}(x_{1,i}) - \phi(x_{1,i})) \partial^{1} \tilde{P}_{2,i}(\beta_k)$ for any $\beta_k \in \Theta_k$. By Assumptions SC1(i) and Assumption SC2(iii, v, vi), and (SC.88)

$$\sup_{\beta_k \in \Theta_k} \left\| \pi_n(\beta_k) \right\| \leq C \xi_{1,m_2} n^{-1/2} \leq C m_2^{2} n^{-1/2}. \quad \text{(SC.326)}$$

Moreover for any $\beta_{k,1}$ and $\beta_{k,2}$, we can use similar arguments in showing (SC.309) to obtain

$$\left\| \pi_n(\beta_{k,1}) - \pi_n(\beta_{k,2}) \right\| \leq C m_2^{3} n^{-1/2} |\beta_{k,1} - \beta_{k,2}|. \quad \text{(SC.327)}$$

Consider the same nested sets $\Theta_{k,j}$ ($j = 1, 2, \ldots$) constructed in the proof of lemma SC34. Let $J_n = \min\{ j : 2^{-j} \leq C m_2^{-1} \}$. Then for any positive integer $J > J_n$ using the similar arguments in the proof of Lemma SC33 we obtain

$$\left\| \max \left\{ \pi_n(\beta_{k,J+1}) - \pi_n(\beta_{k,J}) - \left[ \pi_n(\beta_{k,J+1}) - \pi_n(\beta_{k,J}) \right] \right\} \right\|_2 \leq C m_2^{5/2} n^{-1/2} \quad \text{(SC.328)}$$

where $\beta_{k,J}$ and $\beta'_{k,J}$ are the endpoints of the chains starting at $\beta_{k,J+1}$ and $\beta'_{k,J+1}$ respectively. Since the set $\Theta_{k,J}$ has at most $C m_2$ many elements, by the finite maximum inequality, the triangle inequality and (SC.326)

$$\left\| \max \left\{ \pi_n(\beta_{k,J}) - \pi_n(\beta'_{k,J}) \right\} \right\|_2 \leq C m_2^{1/2} \sup_{\beta_k \in \Theta_k} \left\| \pi_n(\beta_k) \right\|_2 \leq C m_2^{5/2} n^{-1/2}. \quad \text{(SC.329)}$$

Q.E.D.
Then the claim of the lemma follows by applying the chaining arguments in the proof of Lemma SC33. Q.E.D.

References


