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

Jinyong Hahn* Zhipeng Liao† Geert Ridder‡

UCLA UCLA USC

First Version: March 2020; This Version: May 2022

Abstract

This supplemental appendix contains additional technical details of Hahn, Liao, and Ridder (2022). Section SA presents the asymptotic properties of the three-step estimator and the consistency of the variance estimator proposed in Appendix D of Hahn, Liao, and Ridder (2022). The detailed proofs of the asymptotic properties of the three-step estimator and the consistency of the asymptotic variance estimator are included in Section SB.

*Department of Economics, UCLA, Los Angeles, CA 90095-1477 USA. Email: hahn@econ.ucla.edu

†Department of Economics, UCLA, Los Angeles, CA 90095-1477 USA. Email: zhipeng.liao@econ.ucla.edu

‡Department of Economics, University of Southern California, Los Angeles, CA 90089. Email: ridder@usc.edu

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

In this section, we derive the asymptotic properties of $\hat{\beta}_k$ and the consistency of the asymptotic variance estimator defined in Appendix D of Hahn, Liao, and Ridder (2022). The consistency and the asymptotic distribution of $\hat{\beta}_k$ are presented in Subsection SA.1. In Subsection SA.2, we present the consistency of the estimator of the asymptotic variance of $\hat{\beta}_k$ which can be used to construct confidence intervals 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 SA.3.

SA.1 Consistency and Asymptotic Normality

To show the consistency of $\hat{\beta}_k$, we use the standard arguments for showing the consistency of the extremum estimator which requires two primitive conditions: (i) the identification uniqueness condition of the unknown parameter $\beta_{k,0}$; and (ii) the convergence of the estimation criterion function $n^{-1}\sum_{i=1}^n \hat{\tau}_i(\beta_k)^2$ to the population criterion function uniformly over $\beta_k \in \Theta_k$. We impose the identification uniqueness condition of $\beta_{k,0}$ in condition (SA.1) below, which can be verified under low-level sufficient conditions. The uniform convergence of the estimation criterion function is proved in Lemma SA1 in Subsection SA.3.

Lemma SA1 *Let $\tau_i(\beta_k) \equiv y_{2,i} - l_{2,i}\beta_{l,0} - \beta_k k_{2,i} - g(\omega_{1,i}(\beta_k);\beta_k)$ for any $\beta_k \in \Theta_k$. Suppose that for any $\varepsilon > 0$, there exists a constant $\delta_\varepsilon > 0$ such that*

$$\inf_{\{\beta_k\in\Theta_k:|\beta_k-\beta_{k,0}|\geq\varepsilon\}} \mathbb{E}\left[\tau_i(\beta_k)^2 - \tau_i(\beta_{k,0})^2\right] > \delta_\varepsilon. \tag{SA.1}$$

Then under Assumptions SB1 and SB2 in Section SB, we have $\hat{\beta}_k = \beta_{k,0} + o_p(1)$.

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

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

where for any $\beta_k \in \Theta_k$

$$\frac{\partial \hat{g}(\hat{\omega}_{1,i}(\beta_k);\beta_k)}{\partial \beta_k} = \hat{\beta}_g(\beta_k)'\frac{\partial P_2(\hat{\omega}_{1,i}(\beta_k))}{\partial \beta_k} + P_2(\hat{\omega}_{1,i}(\beta_k))'\frac{\partial \hat{\beta}_g(\beta_k)}{\partial \beta_k}. \tag{SA.3}$$

By the definition of $\hat{g}(\hat{\omega}_{1,i}(\hat{\beta}_k); \hat{\beta}_k)$ in (34) of Hahn, Liao, and Ridder (2022), we can write

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

which implies that

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

Therefore, the first-order condition (SA.2) can be reduced to

$$n^{-1}\sum_{i=1}^{n} \hat{\tau}_i(\hat{\beta}_k)\left(k_{2,i} - k_{1,i}\hat{\beta}_g(\hat{\beta}_k)'\frac{\partial P_2(\hat{\omega}_{1,i}(\hat{\beta}_k))}{\partial \omega}\right) = 0 \tag{SA.4}$$

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

Theorem SA1 *Let $g_1(\omega) \equiv \partial g(\omega)/\partial\omega$. Suppose that*

$$\Upsilon \equiv \mathbb{E}\left[(v_{2,i} - v_{1,i}g_1(\omega_{1,i}))^2\right] > 0, \tag{SA.5}$$

where $v_{j,i} \equiv k_{j,i} - \mathbb{E}[k_{j,i}|\omega_{1,i}]$ for $j = 1, 2$. Then under (SA.1) in Lemma SA1, and Assumptions SB1, SB2 and SB3 in Section SB

$$\begin{aligned} n^{1/2}(\hat{\beta}_k - \beta_{k,0}) = {} & \Upsilon^{-1}n^{-1/2}\sum_{i=1}^{n} u_{2,i}\left(v_{2,i} - v_{1,i}g_1(\omega_{1,i})\right) \\ & - \Upsilon^{-1}n^{-1/2}\sum_{i=1}^{n} \eta_{1,i}g_1(\omega_{1,i})\left(v^*_{2,i} - v_{1,i}g_1(\omega_{1,i})\right) \\ & - \Upsilon^{-1}\Gamma n^{1/2}(\hat{\beta}_l - \beta_{l,0}) + o_p(1), \end{aligned} \tag{SA.6}$$

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

$$n^{1/2}(\hat{\beta}_k - \beta_{k,0}) \to_d N(0, \Upsilon^{-1}\Omega\Upsilon^{-1}), \tag{SA.7}$$

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

Remark. The local identification condition of $\beta_{k,0}$ is imposed in (SA.5) which is important to ensure the root-n consistency of $\hat{\beta}_k$. □

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

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

which is maintained in Assumption SB1(iii) in Section SB. Different estimation procedures of $\hat{\beta}_l$ may give different forms for $\varepsilon_{1,i}$. Therefore, the specific form of $\varepsilon_{1,i}$ has to be derived case by case. □

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

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

Therefore we can write

$$\Gamma = \mathbb{E}\left[\left(l_{2,i} - \mathbb{E}[l_{2,i}|\omega_{1,i}] - h(x_{1,i})g_1(\omega_{1,i})\right)\left(v_{2,i} - v_{1,i}g_1(\omega_{1,i})\right)\right], \tag{SA.8}$$

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

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

by the law of iterated expectation. Hence we deduce that

$$\Gamma = \mathbb{E}\left[\left(l_{2,i} - l_{1,i}g_1(\omega_{1,i})\right)\left(v_{2,i} - v_{1,i}g_1(\omega_{1,i})\right)\right] \tag{SA.9}$$

under PIM. □

REMARK. From the asymptotic expansion in (SA.6), we see that the asymptotic variance of $\hat{\beta}_k$ is determined by three components. The first component, $n^{-1/2}\sum_{i=1}^{n} u_{2,i}\left(v_{2,i} - v_{1,i}g_1(\omega_{1,i})\right)$ comes from the third-step estimation with known $\omega_{1,i}$. The second and the third components are from the first-step estimation. Specifically, the second one, $n^{-1/2}\sum_{i=1}^{n} \eta_{1,i}g_1(\omega_{1,i})\left(v_{2,i} - v_{1,i}g_1(\omega_{1,i})\right)$ is from estimating $\phi(\cdot)$ in the first step, while the third component $\Gamma n^{1/2}(\hat{\beta}_l - \beta_{l,0})$ is due to the estimation error in $\hat{\beta}_l$. □

SA.2 Consistent Variance Estimation

The asymptotic variance of $\hat{\beta}_k$ can be estimated using its explicit form and the estimators of $v_{1,i}$, $v_{2,i}$, $\varepsilon_{1,i}$, $\eta_{1,i}$, $u_{2,i}$, $v^*_{2,i}$, $h(x_{1,i})$ and $g_1(\omega_{1,i})$. The unknown functions in $v_{1,i}$, $v_{2,i}$, $\varepsilon_{1,i}$, $\eta_{1,i}$, $u_{2,i}$, $v^*_{2,i}$, $h(x_{1,i})$ and

$g_1(\omega_{1,i})$ can be estimated by the kernel or the series method. Since $\hat{\beta}_k$ is constructed using the series method and its asymptotic properties have been established in the previous subsection, in Appendix D.2 of Hahn, Liao, and Ridder (2022) we provide a estimator of the asymptotic variance based on the series approach. In this subsection, we provide the consistency of this variance estimator.

For ease of reference, we repeat the definition of the variance estimator in Appendix D.2 of Hahn, Liao, and Ridder (2022). First, it is clear that $g_1(\omega_{1,i})$ can be estimated by $\hat{g}_1(\hat{\omega}_{1,i}(\hat{\beta}_k); \hat{\beta}_k)$ where

$$\hat{g}_1(\hat{\omega}_{1,i}(\beta_k); \beta_k) \equiv \hat{\beta}_g(\beta_k)' \frac{\partial P_2(\hat{\omega}_{1,i}(\beta_k))}{\partial \omega} \text{ for any } \beta_k \in \Theta_k. \tag{SA.10}$$

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

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

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

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

$$\hat{\Upsilon}_n \equiv n^{-1} \sum_{i=1}^{n} \hat{\varsigma}_i^2. \tag{SA.11}$$

Moreover the Jacobian term Γ can be estimated by

$$\hat{\Gamma}_n \equiv n^{-1} \sum_{i=1}^{n} (l_{2,i} - \hat{h}_i \hat{g}_1(\hat{\omega}_{1,i}(\hat{\beta}_k); \hat{\beta}_k)) \hat{\varsigma}_i, \tag{SA.12}$$

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

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

Then Ω is estimated by

$$\hat{\Omega}_n \equiv n^{-1} \sum_{i=1}^{n} \left((\hat{u}_{2,i} - \hat{\eta}_{1,i} \hat{g}_1(\hat{\omega}_{1,i}(\hat{\beta}_k); \hat{\beta}_k)) \hat{\varsigma}_i - \hat{\Gamma}_n \hat{\varepsilon}_{1,i} \right)^2 \tag{SA.13}$$

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

Theorem SA2 *Under Assumptions SB1, SB2, SB3 and SB4 in Section SB, we have*

$$\hat{\Upsilon}_n = \Upsilon + o_p(1) \quad \text{and} \quad \hat{\Omega}_n = \Omega + o_p(1) \tag{SA.14}$$

and moreover

$$\frac{n^{1/2}(\hat{\beta}_k - \beta_{k,0})}{(\hat{\Upsilon}_n^{-1}\hat{\Omega}_n\hat{\Upsilon}_n^{-1})^{1/2}} \to_d N(0,1), \tag{SA.15}$$

where $\hat{\Omega}_n$ *is defined in (SA.13).*

SA.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 SA1. By (SB.67) in the proof of Lemma SB5 and Assumption SB2(i)

$$\sup_{\beta_k \in \Theta_k} \mathbb{E}\left[\tau_i(\beta_k)^2\right] \leq C \tag{SA.16}$$

which together with Lemma SB5 implies that

$$\sup_{\beta_k \in \Theta_k} n^{-1} \sum_{i=1}^{n} \tau_i(\beta_k)^2 = O_p(1). \tag{SA.17}$$

By the Markov inequality, Assumptions SB1(i, iii) and SB2(i), we obtain

$$n^{-1} \sum_{i=1}^{n} (\hat{y}_{2,i}^* - y_{2,i}^*)^2 = (\hat{\beta}_l - \beta_l)^2 n^{-1} \sum_{i=1}^{n} l_{2,i}^2 = O_p(n^{-1}). \tag{SA.18}$$

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

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

which together with Assumption SB2(vi), Lemma SB4, Lemma SB5, (SA.17), (SA.18) and the Cauchy-Schwarz inequality implies that

$$
\sup_{\beta_k\in\Theta_k}\left|n^{-1}\sum_{i=1}^{n}\hat{\tau}_i(\beta_k)^2 - \mathbb{E}\left[\tau_i(\beta_k)^2\right]\right| = o_p(1). \tag{SA.19}
$$

The consistency of $\hat{\beta}_k$ follows from its definition in (35) of Hahn, Liao, and Ridder (2022), (SA.19), the identification uniqueness condition of $\beta_{k,0}$ assumed in (SA.1) and the standard arguments of showing the consistency of the extremum estimator. *Q.E.D.*

Lemma SA2 *Let $g_{1,i} \equiv g_1(\omega_{1,i})$ and $\hat{J}_i(\beta_k) \equiv \hat{\tau}_i(\beta_k)\left(k_{2,i} - k_{1,i}\hat{g}_1(\hat{\omega}_{1,i}(\beta_k);\beta_k)\right)$ for any $\beta_k \in \Theta_k$, where $\hat{g}_1(\hat{\omega}_{1,i}(\beta_k);\beta_k)$ is defined in (SA.10). Then under Assumptions SB1, SB2 and SB3, we have*

$$
n^{-1}\sum_{i=1}^{n}\hat{J}_i(\beta_{k,0}) = n^{-1}\sum_{i=1}^{n}(u_{2,i} - \eta_{1,i}g_{1,i})(v_{2,i} - v_{1,i}g_{1,i}) - \Gamma(\hat{\beta}_l - \beta_{l,0}) + o_p(n^{-1/2}). \tag{SA.20}
$$

Proof of Lemma SA2. By the definition of $\hat{\tau}_i(\beta_{k,0})$ and Lemma SB7,

$$
\begin{aligned}
& n^{-1}\sum_{i=1}^{n}\hat{\tau}_i(\beta_{k,0})\left(k_{2,i} - k_{1,i}\hat{g}_1(\hat{\omega}_{1,i}(\beta_{k,0});\beta_{k,0})\right) \\
&= n^{-1}\sum_{i=1}^{n}(\hat{y}^*_{2,i}(\beta_{k,0}) - g(\omega_{1,i}))(k_{2,i} - k_{1,i}g_{1,i}) \\
&\quad - n^{-1}\sum_{i=1}^{n}(\hat{g}(\hat{\omega}_{1,i}(\beta_{k,0});\beta_{k,0}) - g(\omega_{1,i}))(k_{2,i} - k_{1,i}g_{1,i}) + o_p(n^{-1/2}),
\end{aligned} \tag{SA.21}
$$

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 SB9

$$
\begin{aligned}
& n^{-1}\sum_{i=1}^{n}(\hat{g}(\hat{\omega}_{1,i}(\beta_{k,0});\beta_{k,0}) - g(\omega_{1,i}))(k_{2,i} - k_{1,i}g_{1,i}) \\
& = n^{-1}\sum_{i=1}^{n} u_{2,i}\varphi(\omega_{1,i}) - \mathbb{E}[l_{2,i}\varphi(\omega_{1,i})](\hat{\beta}_l - \beta_{l,0}) \\
& + n^{-1}\sum_{i=1}^{n} g_{1,i}(\hat{\phi}(x_{1,i}) - \phi(x_{1,i}))(v_{2,i} - v_{1,i}g_{1,i}) + o_p(n^{-1/2}),
\end{aligned} \tag{SA.22}
$$

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

$$
\begin{aligned}
& n^{-1}\sum_{i=1}^{n}(\hat{y}^*_{2,i}(\beta_{k,0}) - g(\omega_{1,i}))(k_{2,i} - k_{1,i}g_{1,i}) \\
& = n^{-1}\sum_{i=1}^{n} u_{2,i}(k_{2,i} - k_{1,i}g_{1,i}) - (\hat{\beta}_l - \beta_{l,0})n^{-1}\sum_{i=1}^{n} l_{2,i}(k_{2,i} - k_{1,i}g_{1,i}) \\
& = n^{-1}\sum_{i=1}^{n} u_{2,i}(k_{2,i} - k_{1,i}g_{1,i}) - (\hat{\beta}_l - \beta_{l,0})\mathbb{E}[l_{2,i}(k_{2,i} - k_{1,i}g_{1,i})] + o_p(n^{-1/2})
\end{aligned} \tag{SA.23}
$$

where the second equality is by Assumption SB1(iii) and

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

which holds by the Markov inequality, Assumptions SB1(i) and SB2(i, ii). Therefore by (SA.21), (SA.22) and (SA.23), we obtain

$$
\begin{aligned}
& n^{-1}\sum_{i=1}^{n} \hat{\tau}_i(\beta_{k,0})\left(k_{2,i} - k_{1,i}\hat{g}_1(\hat{\omega}_{1,i}(\beta_{k,0});\beta_{k,0})\right) \\
& = n^{-1}\sum_{i=1}^{n} u_{2,i}(v_{2,i} - v_{1,i}g_{1,i}) - (\hat{\beta}_l - \beta_{l,0})\mathbb{E}[l_{2,i}(v_{2,i} - v_{1,i}g_{1,i})] \\
& - n^{-1}\sum_{i=1}^{n} g_{1,i}(\hat{\phi}(x_{1,i}) - \phi(x_{1,i}))(v_{2,i} - v_{1,i}g_{1,i}) + o_p(n^{-1/2}).
\end{aligned} \tag{SA.24}
$$

The claim of the lemma follows from (SA.24) and Lemma SB10. *Q.E.D.*

Lemma SA3 *Under Assumptions SB1, SB2 and SB3, we have*

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

PROOF OF LEMMA SA3. First note that by the definition of $\hat{J}_i(\beta_k)$ and $\hat{\tau}_i(\beta_k)$, we can write

$$\begin{aligned}
& n^{-1}\sum_{i=1}^{n}(\hat{J}_i(\hat{\beta}_k)-\hat{J}_i(\beta_{k,0})) \\
& = -(\hat{\beta}_k-\beta_{k,0})n^{-1}\sum_{i=1}^{n}k_{2,i}(k_{2,i}-k_{1,i}\hat{g}_1(\hat{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k)) \\
& - n^{-1}\sum_{i=1}^{n}(\hat{g}(\hat{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k)-\hat{g}(\hat{\omega}_{1,i}(\beta_{k,0});\beta_{k,0}))(k_{2,i}-k_{1,i}\hat{g}_1(\hat{\omega}_{1,i}(\beta_{k,0});\beta_{k,0})) \\
& - n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(\hat{g}_1(\hat{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k)-\hat{g}_1(\hat{\omega}_{1,i}(\beta_{k,0});\beta_{k,0})) \\
& + (\hat{\beta}_l-\beta_{l,0})n^{-1}\sum_{i=1}^{n}l_{2,i}k_{1,i}(\hat{g}_1(\hat{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k)-\hat{g}_1(\hat{\omega}_{1,i}(\beta_{k,0});\beta_{k,0})) \qquad \text{(SA.25)}
\end{aligned}$$

which together with Assumption SB1(iii), Lemma SB14, Lemma SB18 and Lemma SB20 implies that

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

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

PROOF OF THEOREM SA1. By Assumptions SB1(ii, iii) and SB2(i, ii), and Hölder's inequality

$$\Gamma = \mathbb{E}\left[(l_{2,i}-h_i g_1(\omega_{1,i}))(v_{2,i}-v_{1,i}g_1(\omega_{1,i}))\right] \leq C \qquad \text{(SA.26)}$$

and

$$\begin{aligned}
\Omega &= \mathbb{E}[((u_{2,i}-\eta_{1,i}g_1(\omega_{1,i}))(v_{2,i}-v_{1,i}g_1(\omega_{1,i}))-\Gamma\varepsilon_{1,i})^2] \\
&\leq C\mathbb{E}[u_{2,i}^4+\eta_{1,i}^4+v_{1,i}^4+v_{2,i}^4+\varepsilon_{1,i}^2] \leq C. \qquad \text{(SA.27)}
\end{aligned}$$

By Assumption SB1(i), (SA.27) and the Lindeberg–Lévy central limit theorem,

$$n^{-1/2}\sum_{i=1}^{n}((u_{2,i}-\eta_{1,i}g_1(\omega_{1,i}))(v_{2,i}-v_{1,i}g_1(\omega_{1,i}))-\Gamma\varepsilon_{1,i})\rightarrow_d N(0,\Omega). \tag{SA.28}$$

By (SA.4), Assumption SB1(iii), Lemma SA2 and Lemma SA3, we can write

$$\begin{aligned}0 &= n^{-1}\sum_{i=1}^{n}\hat{J}_i(\beta_{k,0})+n^{-1}\sum_{i=1}^{n}(\hat{J}_i(\hat{\beta}_k)-\hat{J}_i(\beta_{k,0}))\\ &= n^{-1}\sum_{i=1}^{n}(u_{2,i}-\eta_{1,i}g_1(\omega_{1,i}))(v_{2,i}-v_{1,i}g_1(\omega_{1,i}))-\Gamma n^{1/2}(\hat{\beta}_l-\beta_{l,0})\\ &\quad-(\hat{\beta}_k-\beta_{k,0})\left[\mathbb{E}[(v_{2,i}-v_{1,i}g_1(\omega_{1,i}))^2]+o_p(1)\right]+o_p(n^{-1/2})\end{aligned} \tag{SA.29}$$

which together with (SA.5) and (SA.28) implies that

$$n^{1/2}(\hat{\beta}_k-\beta_{k,0})=\Upsilon^{-1}n^{-1/2}\sum_{i=1}^{n}(u_{2,i}-\eta_{1,i}g_1(\omega_{1,i}))(v_{2,i}-v_{1,i}g_1(\omega_{1,i}))-\Upsilon^{-1}\Gamma n^{1/2}(\hat{\beta}_l-\beta_{l,0})+o_p(1). \tag{SA.30}$$

This proves (SA.6). The claim in (SA.7) follows from Assumption SB1(iii), (SA.28) and (SA.30).*Q.E.D.*

Proof of Theorem SA2. The results in (SA.14) are proved in Lemma SB22(i, iii), which together with Theorem SA1, Assumption SB4(iii) and the Slutsky Theorem proves the claim in (SA.15). *Q.E.D.*

SB Auxiliary Results

In this section, we provide the auxiliary results which are used to show Lemma SA1, Theorem SA1 and Theorem SA2. 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 section, 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.

SB.1 The Asymptotic Properties of the First-step Estimators

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

Assumption SB1 *(i) The data* $\{(y_{t,i}, i_{t,i}, k_{t,i}, l_{t,i})_{t=1,2}\}_{i=1}^{n}$ *are i.i.d.; (ii)* $\mathbb{E}[\eta_{1,i}|\, x_{1,i}] = 0$ *and* $\mathbb{E}[l_{1,i}^2 + \eta_{1,i}^4|x_{1,i}] \leq C$*; (iii) there exist i.i.d. random variables* $\varepsilon_{1,i}$ *with* $\mathbb{E}[\varepsilon_{1,i}^4] \leq C$ *such that*

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

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

Assumption SB1(iii) assumes that there exists a root-n consistent estimator $\hat{\beta}_l$ of $\beta_{l,0}$. Different estimation procedures of $\hat{\beta}_l$ may give different forms for $\varepsilon_{1,i}$. For example, $\hat{\beta}_l$ may be obtained together with the nonparametric estimator of $\phi(\cdot)$ in the partially linear regression proposed in Olley and Pakes (1996), or from the GMM estimation proposed in Ackerberg, Caves, and Frazer (2015). Therefore, the specific form of $\varepsilon_{1,i}$ has to be derived case by case. The rest of the conditions in Assumption SB1 are fairly standard in series estimation; see, for example, Andrews (1991), Newey (1997) and Chen (2007).[1] In particular, condition (iv) specifies the precision for approximating the unknown function $\phi(\cdot)$ via approximating functions, for which comprehensive results are available from numerical approximation theory.

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

Lemma SB1 *Under Assumption SB1, we have*

$$n^{-1} \sum_{i=1}^{n} |\hat{\phi}(x_{1,i}) - \phi(x_{1,i})|^2 = O_p(m_1 n^{-1}) \tag{SB.31}$$

and moreover

$$\sup_{x_1\in\mathcal{X}} |\hat{\phi}(x_1) - \phi(x_1)| = O_p(\xi_{0,m_1} m_1^{1/2} n^{-1/2}). \tag{SB.32}$$

Proof of Lemma SB1. Under Assumption SB1(i, v, vi), we can invoke Lemma 6.2 in Belloni, Chernozhukov, Chetverikov, and Kato (2015) to obtain

$$\left\|n^{-1}\mathbf{P}_1'\mathbf{P}_1 - Q_{m_1}\right\|_S = O_p((\log m_1)^{1/2}\xi_{0,m_1} n^{-1/2}) = o_p(1), \tag{SB.33}$$

[1] For some approximating functions such as power series, Assumptions SB1(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 SB1(v) implies that

$$C^{-1} \leq \lambda_{\min}(n^{-1}\mathbf{P}_1'\mathbf{P}_1) \leq \lambda_{\max}(n^{-1}\mathbf{P}_1'\mathbf{P}_1) \leq C \tag{SB.34}$$

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

$$\begin{aligned}\hat{\beta}_\phi - \beta_{\phi,m_1} = & (\mathbf{P}_1'\mathbf{P}_1)^{-1}\sum_{i=1}^n P_1(x_{1,i})\eta_{1,i} \\ & + (\mathbf{P}_1'\mathbf{P}_1)^{-1}\sum_{i=1}^n P_1(x_{1,i})(\phi(x_{1,i}) - \phi_{m_1}(x_{1,i})) \\ & - (\hat{\beta}_l - \beta_{l,0})(\mathbf{P}_1'\mathbf{P}_1)^{-1}\sum_{i=1}^n P_1(x_{1,i})l_{1,i}. \end{aligned} \tag{SB.35}$$

By Assumption SB1(i, ii, v) and the Markov inequality

$$n^{-1}\sum_{i=1}^n P_1(x_{1,i})\eta_{1,i} = O_p(m_1^{1/2}n^{-1/2}) \tag{SB.36}$$

which together with Assumption SB1(vi), (SB.33) and (SB.34) implies that

$$\left[(n^{-1}\mathbf{P}_1'\mathbf{P}_1)^{-1} - Q_{m_1}^{-1}\right] n^{-1}\sum_{i=1}^n P_1(x_{1,i})\eta_{1,i} = O_p((\log m_1)^{1/2}\xi_{0,m_1}m_1^{1/2}n^{-1}) = o_p(n^{-1/2}). \tag{SB.37}$$

By Assumption SB1(iv, vi) and (SB.34)

$$(\mathbf{P}_1'\mathbf{P}_1)^{-1}\sum_{i=1}^n P_1(x_{1,i})(\phi(x_{1,i}) - \phi_{m_1}(x_{1,i})) = O_p(m^{-r_\phi}) = O_p(n^{-1/2}). \tag{SB.38}$$

Under Assumption SB1(i, ii, v, vi), we can use similar arguments in showing (SB.36) to get

$$n^{-1}\sum_{i=1}^n P_1(x_{1,i})l_{1,i} - \mathbb{E}[P_1(x_{1,i})l_{1,i}] = O_p(m_1^{1/2}n^{-1/2}) = o_p(1). \tag{SB.39}$$

By Assumption SB1(i, ii, v),

$$\|\mathbb{E}[l_{1,i}P_1(x_{1,i})]\|^2 \leq \lambda_{\max}(Q_{m_1})\mathbb{E}[l_{1,i}P_1(x_{1,i})']Q_{m_1}^{-1}\mathbb{E}[P_1(x_{1,i})l_{1,i}] \leq C\mathbb{E}[l_{1,i}^2] \leq C \tag{SB.40}$$

which combined with (SB.39) implies that

$$n^{-1}\sum_{i=1}^{n}P_1(x_{1,i})l_{1,i}=O_p(1). \tag{SB.41}$$

By Assumption SB1(iii, v, vi), (SB.33), (SB.39), (SB.40) and (SB.41),

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

which combined with Assumption SB1(vi), (SB.35), (SB.37) and (SB.38) shows that

$$\hat{\beta}_\phi-\beta_{\phi,m_1}=Q_{m_1}^{-1}\left(\sum_{i=1}^{n}P_1(x_{1,i})\eta_{1,i}-\mathbb{E}[P_1(x_{1,i})l_{1,i}](\hat{\beta}_l-\beta_{l,0})\right)+O_p(n^{-1/2})=O_p(m_1^{1/2}n^{-1/2}), \tag{SB.42}$$

where the second equality follows from Assumptions SB1(iii, v), (SB.36) and (SB.40). By the Cauchy-Schwarz inequality

$$\begin{aligned}n^{-1}\sum_{i=1}^{n}|\hat{\phi}(x_{1,i})-\phi(x_{1,i})|^2&\leq 2n^{-1}\sum_{i=1}^{n}|\hat{\phi}(x_{1,i})-\phi_{m_1}(x_{1,i})|^2+2n^{-1}\sum_{i=1}^{n}|\phi_{m_1}(x_{1,i})-\phi(x_{1,i})|^2\\&\leq 2\lambda_{\max}(n^{-1}\mathbf{P}_1'\mathbf{P}_1)\left\|\hat{\beta}_\phi-\beta_{\phi,m_1}\right\|^2+2\sup_{x\in\mathcal{X}_1}|\phi_{m_1}(x)-\phi(x)|=O_p(m_1^{1/2}n^{-1/2}),\end{aligned} \tag{SB.43}$$

where the equality is by Assumptions SB1(iv, vi), (SB.34) and (SB.42), which proves (SB.31). By the triangle inequality, the Cauchy-Schwarz inequality, Assumption SB1(iv, vi) and (SB.42)

$$\begin{aligned}\sup_{x_1\in\mathcal{X}}|\hat{\phi}(x_1)-\phi(x_1)|&\leq\sup_{x_1\in\mathcal{X}}|\hat{\phi}(x_1)-\phi_{m_1}(x_1)|+\sup_{x_1\in\mathcal{X}}|\phi_{m_1}(x_1)-\phi(x_1)|\\&\leq\xi_{0,m_1}\left\|\hat{\beta}_\phi-\beta_{\phi,m_1}\right\|+O(m_1^{-r_\phi})=O_p(\xi_{0,m_1}m_1^{1/2}n^{-1/2})\end{aligned} \tag{SB.44}$$

which proves the claim in (SB.31). *Q.E.D.*

SB.2 Auxiliary Results for the Consistency of $\hat{\beta}_k$

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

Assumption SB2 *(i)* $\mathbb{E}[(y_{2,i}^*)^4 + l_{2,i}^4 + k_{2,i}^4|x_{1,i}] \leq C$*; (ii)* $g(\omega;\beta_k)$ *is continuously differentiable with uniformly bounded derivatives; (iii) for some* $d \geq 1$ *there exist* $\beta_{g,m_2}(\beta_k) \in \mathbb{R}^{m_2}$ *and* $r_g > 0$ *such that* $\sup_{\beta_k\in\Theta_k} |g(\beta_k) - g_{m_2}(\beta_k)|_d = O(m_2^{-r_g})$ *where* $g_{m_2}(\omega;\beta_k) \equiv P_2(\omega)' \beta_{g,m_2}(\beta_k)$*; (iv) for any* $\beta_k \in \Theta_k$ *there exists a nonsingular matrix* $B(\beta_k)$ *such that for* $\tilde{P}_2(\omega_1(\beta_k);\beta_k) \equiv B(\beta_k)P_2(\omega_1(\beta_k))$*,*

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

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

Assumption SB2(i) imposes an upper bound on the conditional moments of $y_{2,i}^*$, $l_{2,i}$ and $k_{2,i}$ given $x_{1,i}$. Assumptions SB2(ii, iii) require that the conditional moment function $g(\omega;\beta_k)$ is smooth and can be well approximated by linear combinations of $P_2(\omega)$. Assumption SB2(iv) imposes normalization on the approximating functions $P_2(\omega)$, and uniform lower and upper bounds on the eigenvalues of $Q_{m_2}(\beta_k)$. Assumption SB2(v, vi) restrict the magnitudes of the normalized approximating functions and their derivatives, and the convergence rate of the series approximation error.

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

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

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

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

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

Lemma SB2 *Under Assumptions SB1 and SB2, we have*

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

[2]Note that we define $\hat{P}_{2,i}(\beta_k) \equiv P_2(\hat{\omega}_{1,i}(\beta_k))$ in Section D.1 of Hahn, Liao, and Ridder (2022). We change its definition here since the asymptotic properties of the series estimator $\hat{g}(\hat{\omega}_{1,i}(\beta_k);\beta_k) = P_2(\hat{\omega}_{1,i}(\beta_k))'\hat{\beta}_g(\beta_k)$ shall be investigated under the new definition $\hat{P}_{2,i}(\beta_k) \equiv B(\beta_k)P_2(\hat{\omega}_{1,i}(\beta_k))$.

Proof of Lemma SB2. Since $\hat{\omega}_{1,i}(\beta_k) = \hat{\phi}(x_{1,i}) - \beta_k k_{1,i}$, by Lemma SB1

$$\sup_{\beta_k \in \Theta_k} \max_{i \leq n} |\hat{\omega}_{1,i}(\beta_k) - \omega_{1,i}(\beta_k)| = \max_{i \leq n} |\hat{\phi}(x_{1,i}) - \phi(x_{1,i})| = O_p(\xi_{0,m_1} m_1^{1/2} n^{-1/2}) = o_p(1) \tag{SB.45}$$

which together with Assumption SB2(vi) implies that

$$\hat{\omega}_{1,i}(\beta_k) \in \Omega_{\varepsilon}(\beta_k) \text{ wpa1} \tag{SB.46}$$

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

$$\left| \upsilon_2'(\tilde{P}_{2,i}(\beta_k) - \hat{P}_{2,i}(\beta_k)) \right| = \left| \upsilon_2' \partial^1 \tilde{P}_2 \left(\tilde{\omega}_{1,i}(\beta_k); \beta_k \right) \left(\hat{\omega}_{1,i}(\beta_k) - \omega_{1,i}(\beta_k) \right) \right|, \tag{SB.47}$$

where $\tilde{\omega}_{1,i}(\beta_k)$ lies between $\omega_{1,i}(\beta_k)$ and $\hat{\omega}_{1,i}(\beta_k)$. Since $\omega_{1,i}(\beta_k)$ and $\hat{\omega}_{1,i}(\beta_k)$ are in $\Omega_{\varepsilon}(\beta_k)$ uniformly over $\beta_k \in \Theta_k$ and for any $i = 1, \ldots, n$ wpa1, the same property holds for $\tilde{\omega}_{1,i}(\beta_k)$. By the Cauchy-Schwarz inequality, Assumption SB2(v) and (SB.47)

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

Therefore,

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

wpa1, which together with Lemma SB1 implies that

$$\sup_{\beta_k \in \Theta_k} ||\hat{\mathbf{P}}_2(\beta_k) - \tilde{\mathbf{P}}_2(\beta_k)||_S = O_p(\xi_{1,m_2} m_1^{1/2}). \tag{SB.48}$$

By Lemma SB24 and Assumption SB2(iv, vi), we have uniformly over $\beta_k \in \Theta_k$

$$C^{-1} \leq \lambda_{\min}(n^{-1} \tilde{\mathbf{P}}_2(\beta_k)' \tilde{\mathbf{P}}_2(\beta_k)) \leq \lambda_{\max}(n^{-1} \tilde{\mathbf{P}}_2(\beta_k)' \tilde{\mathbf{P}}_2(\beta_k)) \leq C \text{ wpa1.} \tag{SB.49}$$

By the triangle inequality, Assumption SB2(vi), (SB.48) and (SB.49), we get

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

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

Lemma SB3 *Under Assumptions SB1 and SB2, we have*

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

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

PROOF OF LEMMA SB3. By the Cauchy-Schwarz inequality and Assumption SB2(iii)

$$\begin{aligned}
&n^{-1} \sum_{i=1}^{n} \left| \tilde{P}_{2,i}(\beta_k)'\hat{\beta}_g(\beta_k) - g(\omega_{1,i}(\beta_k); \beta_k) \right|^2 \\
&\leq 2n^{-1} \sum_{i=1}^{n} \left| \tilde{P}_{2,i}(\beta_k)'\hat{\beta}_g(\beta_k) - g_{m_2}(\omega_{1,i}(\beta_k); \beta_k) \right|^2 \\
&+ 2n^{-1} \sum_{i=1}^{n} |g_{m_2}(\omega_{1,i}(\beta_k); \beta_k) - g(\omega_{1,i}(\beta_k); \beta_k)|^2 \\
&\leq 2\lambda_{\max}(n^{-1}\tilde{\mathbf{P}}_2(\beta_k)'\tilde{\mathbf{P}}_2(\beta_k))||\hat{\beta}_g(\beta_k) - \tilde{\beta}_{g,m_2}(\beta_k)||^2 + Cm_2^{-2r_g}
\end{aligned} \qquad \text{(SB.50)}$$

for any $\beta_k \in \Theta_k$, where $\tilde{\beta}_{g,m_2}(\beta_k) \equiv (B(\beta_k)')^{-1}\beta_{g,m_2}(\beta_k)$ and $\beta_{g,m_2}(\beta_k)$ is defined in Assumption SB2(iii). We next show that

$$\sup_{\beta_k \in \Theta_k} \left\| \hat{\beta}_g(\beta_k) - \tilde{\beta}_{g,m_2}(\beta_k) \right\|^2 = O_p((m_2^2 + \xi_{1,m_2}^2 m_1)n^{-1}) = o_p(1) \qquad \text{(SB.51)}$$

which together with (SB.49) and (SB.50) proves the claim of the lemma.

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

$$\begin{aligned}\hat{\beta}_g(\beta_k) - \tilde{\beta}_{g,m_2}(\beta_k) &= (\hat{\mathbf{P}}_2(\beta_k)'\hat{\mathbf{P}}_2(\beta_k))^{-1}\hat{\mathbf{P}}_2(\beta_k)'(\hat{\mathbf{Y}}^*_2(\beta_k) - \hat{\mathbf{P}}_2(\beta_k)'\tilde{\beta}_{g,m_2}(\beta_k)) \\ &= (\hat{\mathbf{P}}_2(\beta_k)'\hat{\mathbf{P}}_2(\beta_k))^{-1}\sum_{i=1}^n \hat{P}_{2,i}(\beta_k)(g(\omega_{1,i}(\beta_k), \beta_k) - g_{m_2}(\hat{\omega}_{1,i}(\beta_k), \beta_k)) \\ &- (\hat{\beta}_l - \beta_{l,0})(\hat{\mathbf{P}}_2(\beta_k)'\hat{\mathbf{P}}_2(\beta_k))^{-1}\sum_{i=1}^n \hat{P}_{2,i}(\beta_k)l_{2,i} \\ &+ (\hat{\mathbf{P}}_2(\beta_k)'\hat{\mathbf{P}}_2(\beta_k))^{-1}\sum_{i=1}^n \hat{P}_{2,i}(\beta_k)u_{2,i}(\beta_k), \end{aligned} \tag{SB.52}$$

where $g_{m_2}(\hat{\omega}_{1,i}(\beta_k), \beta_k) = \hat{P}_{2,i}(\beta_k)'\tilde{\beta}_{g,m_2}(\beta_k)$. By Assumption SB2(vi), Lemma SB2 and (SB.49), we have uniformly over $\beta_k \in \Theta_k$

$$C^{-1} \leq \lambda_{\min}(n^{-1}\hat{\mathbf{P}}_2(\beta_k)'\hat{\mathbf{P}}_2(\beta_k)) \leq \lambda_{\max}(n^{-1}\hat{\mathbf{P}}_2(\beta_k)'\hat{\mathbf{P}}_2(\beta_k)) \leq C \text{ wpa1} \tag{SB.53}$$

which implies that $\hat{\mathbf{P}}_2(\beta_k)(\hat{\mathbf{P}}_2(\beta_k)'\hat{\mathbf{P}}_2(\beta_k))^{-1}\hat{\mathbf{P}}_2(\beta_k)'$ is an idempotent matrix uniformly over $\beta_k \in \Theta_k$ wpa1. Therefore,

$$\begin{aligned}&\left\|(\hat{\mathbf{P}}_2(\beta_k)'\hat{\mathbf{P}}_2(\beta_k))^{-1}\sum_{i=1}^n \hat{P}_{2,i}(\beta_k)(g(\omega_{1,i}(\beta_k), \beta_k) - g_{m_2}(\hat{\omega}_{1,i}(\beta_k), \beta_k))\right\|^2 \\ &\leq O_p(1)n^{-1}\sum_{i=1}^n (g(\omega_{1,i}(\beta_k), \beta_k) - g_{m_2}(\hat{\omega}_{1,i}(\beta_k), \beta_k))^2, \end{aligned} \tag{SB.54}$$

uniformly over $\beta_k \in \Theta_k$. Since $\omega_{1,i}(\beta_k) = \phi(x_{1,i}) - k_{1,i}\beta_k$, we can use Assumptions SB1(i) and SB2(i) to deduce

$$\sup_{\beta_k \in \Theta_k} |g(\omega_{1,i}(\beta_k); \beta_k)| \leq C. \tag{SB.55}$$

Therefore,

$$\begin{aligned}\sup_{\beta_k \in \Theta_k} \left\|\tilde{\beta}_{g,m_2}(\beta_k)\right\|^2 &\leq \sup_{\beta_k \in \Theta_k} (\lambda_{\min}(Q_{m_2}(\beta_k)))^{-1} \left\|\tilde{P}_{2,i}(\beta_k)'\tilde{\beta}_{g,m_2}(\beta_k)\right\|_2^2 \\ &\leq C \sup_{\beta_k \in \Theta_k} \|g(\omega_{1,i}(\beta_k); \beta_k) - g_{m_2}(\omega_{1,i}(\beta_k); \beta_k)\|_2^2 \\ &+ C \sup_{\beta_k \in \Theta_k} \|g(\omega_{1,i}(\beta_k); \beta_k)\|_2^2 \leq C. \end{aligned} \tag{SB.56}$$

By the second order expansion, Assumption SB2(iii, v, vi), Lemma SB1, (SB.55) and (SB.56), we have uniformly over $\beta_k \in \Theta_k$,

$$
\begin{aligned}
& n^{-1}\sum_{i=1}^{n}(g_{m_2}(\omega_{1,i}(\beta_k),\beta_k)-g_{m_2}(\hat{\omega}_{1,i}(\beta_k),\beta_k))^2 \\
& \leq 2n^{-1}\sum_{i=1}^{n}(\partial^1\tilde{P}_{2,i}(\beta_k)'\tilde{\beta}_{g,m_2}(\beta_k)(\hat{\phi}(x_{1,i})-\phi(x_{1,i}))^2 \\
& + 2n^{-1}\sum_{i=1}^{n}(\partial^2\tilde{P}_2\left(\tilde{\omega}_{1,i}(\beta_k);\beta_k\right)'\tilde{\beta}_{g,m_2}(\beta_k)(\hat{\phi}(x_{1,i})-\phi(x_{1,i})^2)^2 \\
& = O_p(m_1n^{-1})+O_p(\xi_{2,m_2}^2\xi_{0,m_1}^2m_1^2n^{-2}) = O_p(m_1n^{-1}),
\end{aligned}
$$

where $\tilde{\omega}_{1,i}(\beta_k)$ is between $\omega_{1,i}(\beta_k)$ and $\hat{\omega}_{1,i}(\beta_k)$ and it lies in $\Omega_\varepsilon(\beta_k)$ uniformly over $\beta_k \in \Theta_k$ wpa1 by (SB.46), which together with Assumption SB2(iii, vi) implies that

$$
\begin{aligned}
& n^{-1}\sum_{i=1}^{n}(g(\omega_{1,i}(\beta_k),\beta_k)-g_{m_2}(\hat{\omega}_{1,i}(\beta_k),\beta_k))^2 \\
& \leq Cn^{-1}\sum_{i=1}^{n}(g(\omega_{1,i}(\beta_k),\beta_k)-g_{m_2}(\omega_{1,i}(\beta_k),\beta_k))^2 \\
& + Cn^{-1}\sum_{i=1}^{n}(g_{m_2}(\omega_{1,i}(\beta_k),\beta_k)-g_{m_2}(\hat{\omega}_{1,i}(\beta_k),\beta_k))^2 \\
& = O_p(m_1n^{-1}+m_2^{-2r_g}) = O_p(m_1n^{-1}).
\end{aligned} \tag{SB.57}
$$

From (SB.54) and (SB.57), we get uniformly over $\beta_k \in \Theta_k$

$$
(\hat{\mathbf{P}}_2(\beta_k)'\hat{\mathbf{P}}_2(\beta_k))^{-1}\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_k)(g(\omega_{1,i}(\beta_k),\beta_k)-g_{m_2}(\hat{\omega}_{1,i}(\beta_k),\beta_k)) = O_p(m_1^{1/2}n^{-1/2}). \tag{SB.58}
$$

By Assumptions SB1(i) and SB2(i), and the Markov inequality,

$$
n^{-1}\sum_{i=1}^{n}l_{2,i}^2 = O_p(1) \tag{SB.59}
$$

which together with Assumption SB1(iii) and (SB.53) implies that

$$
(\hat{\beta}_l-\beta_{l,0})(\hat{\mathbf{P}}_2(\beta_k)'\hat{\mathbf{P}}_2(\beta_k))^{-1}\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_k)l_{2,i} = O_p(n^{-1/2}) \tag{SB.60}
$$

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

$$\begin{aligned}
&\left| n^{-1} \sum_{i=1}^{n} \upsilon_2' (\hat{P}_{2,i}(\beta_k) - \tilde{P}_{2,i}(\beta_k)) u_{2,i}(\beta_k) \right| \\
&= \left| n^{-1} \sum_{i=1}^{n} \upsilon_2' \partial^1 \tilde{P}_2 \left(\tilde{\omega}_{1,i}(\beta_k); \beta_k\right) \left(\hat{\omega}_{1,i}(\beta_k) - \omega_{1,i}(\beta_k)\right) u_{2,i}(\beta_k) \right| \\
&\leq \|\upsilon_2\| \, \xi_{1,m_2} n^{-1} \sum_{i=1}^{n} \left| (\hat{\phi}(x_{1,i}) - \phi(x_{1,i})) u_{2,i}(\beta_k) \right| . \qquad \text{(SB.61)}
\end{aligned}$$

By the definition of $u_{2,i}(\beta_k)$, we can use Assumptions SB1(i) and SB2(i), (SB.55) and the Markov inequality to deduce

$$\sup_{\beta_k \in \Theta_k} n^{-1} \sum_{i=1}^{n} (u_{2,i}(\beta_k))^2 = O_p(1). \qquad \text{(SB.62)}$$

Thus by the Cauchy-Schwarz inequality, Lemma SB1 and (SB.62),

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

which together with (SB.53) and (SB.61) implies that

$$(\hat{\mathbf{P}}_2(\beta_k)' \hat{\mathbf{P}}_2(\beta_k))^{-1} \sum_{i=1}^{n} (\hat{P}_{2,i}(\beta_k) - \tilde{P}_{2,i}(\beta_k)) u_{2,i}(\beta_k) = O_p(\xi_{1,m_2} m_1^{1/2} n^{-1/2}) \qquad \text{(SB.63)}$$

uniformly over $\beta_k \in \Theta_k$. Applying Lemma SB25 and (SB.53) yields

$$(\hat{\mathbf{P}}_2(\beta_k)' \hat{\mathbf{P}}_2(\beta_k))^{-1} \sum_{i=1}^{n} \tilde{P}_{2,i}(\beta_k) u_{2,i}(\beta_k) = O_p(m_2 n^{-1/2}) \qquad \text{(SB.64)}$$

uniformly over $\beta_k \in \Theta_k$. The claim in (SB.51) then follows from Assumption SB2(vi), (SB.52), (SB.58), (SB.60), (SB.63) and (SB.64). *Q.E.D.*

Lemma SB4 *Under Assumptions SB1 and SB2, we have*

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

Proof of Lemma SB4. By the triangle inequality, (SB.51) and (SB.56)

$$\sup_{\beta_k\in\Theta_k}\left\|\hat{\beta}_g(\beta_k)\right\| \leq \sup_{\beta_k\in\Theta_k}\left\|\tilde{\beta}_{g,m_2}(\beta_k)\right\| + \sup_{\beta_k\in\Theta_k}\left\|\hat{\beta}_g(\beta_k)-\tilde{\beta}_{g,m_2}(\beta_k)\right\| = O_p(1). \tag{SB.65}$$

By the mean value expansion, the Cauchy-Schwarz inequality, Assumption SB2(v, vi), Lemma SB1 and (SB.65),

$$\begin{aligned}
&\sup_{\beta_k\in\Theta_k} n^{-1}\sum_{i=1}^{n}\left|(\hat{P}_{2,i}(\beta_k)-\tilde{P}_{2,i}(\beta_k))'\hat{\beta}_g(\beta_k)\right|^2 \\
&= \sup_{\beta_k\in\Theta_k} n^{-1}\sum_{i=1}^{n}\left|\partial^1\tilde{P}_2\left(\tilde{\omega}_{1,i}(\beta_k);\beta_k\right)'\hat{\beta}_g(\beta_k)(\hat{\omega}_{1,i}(\beta_k)-\omega_{1,i}(\beta_k)\right|^2 \\
&\leq \xi_{1,m_2}^2 n^{-1}\sum_{i=1}^{n}(\hat{\phi}(x_{1,i})-\phi(x_{1,i}))^2\sup_{\beta_k\in\Theta_k}\left\|\hat{\beta}_g(\beta_k)\right\| = O_p(\xi_{1,m_2}^2 m_1 n^{-1}) = o_p(1),
\end{aligned} \tag{SB.66}$$

where $\tilde{\omega}_{1,i}(\beta_k)$ is between $\hat{\omega}_{1,i}(\beta_k)$ and $\omega_{1,i}(\beta_k)$ and hence by (SB.46) it lies in $\Omega_\varepsilon(\beta_k)$ wpa1 for any $i \leq n$ and uniformly over $\beta_k \in \Theta_k$. The claim of the lemma directly follows from Lemma SB3 and (SB.66). *Q.E.D.*

Lemma SB5 *Under Assumptions SB1 and SB2, we have*

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

Proof of Lemma SB5. For any $\beta_k \in \Theta_k$, by the Cauchy-Schwarz inequality and (SB.55),

$$\tau_i(\beta_k)^2 \leq C\left[(y_{2,i}^*)^2 + k_{2,i}^2\beta_k^2 + g(\omega_{1,i}(\beta_k);\beta_k)^2\right] \leq C(1+(y_{2,i}^*)^2+k_{2,i}^2). \tag{SB.67}$$

For any $\beta_{k,1}$, $\beta_{k,2} \in \Theta_k$, by the triangle inequality and Assumption SB2(ii),

$$\left|\tau_i(\beta_{k,1})-\tau_i(\beta_{k,2})\right| \leq (C+k_{2,i})\left|\beta_{k,1}-\beta_{k,2}\right|. \tag{SB.68}$$

By Assumption SB2(ii), (SB.67) and (SB.68), we get

$$\mathbb{E}\left[\tau_i(\beta_k)^2\left|\tau_i(\beta_{k,1})-\tau_i(\beta_{k,2})\right|^2\right] \leq C(\beta_{k,1}-\beta_{k,2})^2 \tag{SB.69}$$

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

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

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

$$\left\|\tau_i(\beta_{k,1})^2 - \tau_i(\beta_{k,2})^2\right\|_2 \leq C\left|\beta_{k,1} - \beta_{k,2}\right|. \tag{SB.70}$$

By Assumptions SB1(i) and SB2(i), and (SB.55),

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

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

$$n^{-1/2}\sum_{i=1}^{n}\left(\tau_i(\beta_k)^2 - \mathbb{E}\left[\tau_i(\beta_k)^2\right]\right) = O_p(1) \tag{SB.71}$$

for any $\beta_k \in \Theta_k$. Moreover, by Assumption SB1(i) and (SB.70)

$$\begin{aligned}
&\mathbb{E}\left[\left|n^{-1/2}\sum_{i=1}^{n}\left(\tau_i(\beta_{k,1})^2 - \tau_i(\beta_{k,2})^2 - \mathbb{E}\left[\tau_i(\beta_{k,1})^2 - \tau_i(\beta_{k,1})^2\right]\right)\right|^2\right] \\
&\leq \mathbb{E}\left[\left|\tau_i(\beta_{k,1})^2 - \tau_i(\beta_{k,2})^2\right|^2\right] \leq C\left|\beta_{k,1} - \beta_{k,2}\right|^2.
\end{aligned} \tag{SB.72}$$

Collecting the results in (SB.71) and (SB.72), we can invoke Theorem 2.2.4 in van der Vaart and Wellner (1996) to deduce that

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

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

SB.3 Auxiliary Results for the Asymptotic Normality of $\hat{\beta}_k$

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

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

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

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

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

Assumption SB3 *(i) $\varphi(\omega)$ is continuously differentiable with uniformly bounded derivatives over $\omega \in \Omega(\beta_{k,0})$; (ii) there exist $\beta_{\varphi,m_2} \in \mathbb{R}^{m_2}$ and $r_\varphi > 0$ such that*

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

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

Assumptions SB3(i, ii) require that the function $\varphi(\omega)$ is smooth and can be well approximated by the approximating functions $P_2(\omega)$. Assumption SB3(iii) requires that any function of $x_{1,i}$ with finite L_2-norm can be approximated by the approximating functions $P_1(x_{1,i})$. Assumption SB3(iv) restricts the numbers of the approximating functions, and the smoothness of $\varphi(\omega)$.

Lemma SB6 *Under Assumptions SB1, SB2 and SB2(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 SB2(iii).

Proof of Lemma SB6. By the definition of $\hat{\beta}_g(\beta_k)$, we can utilize the decomposition in (SB.52), and the results in (SB.58) and (SB.60) to get

$$\hat{\beta}_g(\beta_{k,0}) - \tilde{\beta}_{g,m_2}(\beta_{k,0}) = (\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1} \sum_{i=1}^{n} \hat{P}_{2,i}(\beta_{k,0})u_{2,i} + O_p(m_1^{1/2}n^{-1/2}). \tag{SB.73}$$

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

$$n^{-1}\sum_{i=1}^{n} \upsilon_2'(\hat{P}_{2,i}(\beta_{k,0}) - \tilde{P}_{2,i}(\beta_{k,0}))u_{2,i} = n^{-1}\sum_{i=1}^{n} \upsilon_2'\partial^1\tilde{P}_{2,i}(\beta_{k,0})(\hat{\phi}_i - \phi_i)u_{2,i}$$
$$+ n^{-1}\sum_{i=1}^{n} \upsilon_2'\partial^2\tilde{P}_2(\tilde{\omega}_{1,i}; \beta_{k,0})(\hat{\phi}_i - \phi_i)^2 u_{2,i}. \tag{SB.74}$$

By Assumption SB2(i) and (SB.55),

$$\mathbb{E}\left[u_{2,i}^2 | x_{1,i}\right] \leq C. \tag{SB.75}$$

By Assumptions SB1(i, v) and SB2(vi), (SB.75) and the Markov inequality

$$\left\| n^{-1} \sum_{i=1}^{n} |u_{2,i}| \, P_1(x_{1,i}) P_1(x_{1,i})' - \mathbb{E}\left[|u_{2,i}| \, P_1(x_{1,i}) P_1(x_{1,i})' \right] \right\|_S = o_p(1). \tag{SB.76}$$

Since $\lambda_{\max}(\mathbb{E}\left[|u_{2,i}| \, P_1(x_{1,i}) P_1(x_{1,i})'\right]) \leq C$ by Assumption SB1(v) and (SB.75), from (SB.76) we deduce that

$$\lambda_{\max}\left(n^{-1} \sum_{i=1}^{n} |u_{2,i}| \, P_1(x_{1,i}) P_1(x_{1,i})' \right) \leq C \text{ wpa1}. \tag{SB.77}$$

By (SB.42) and (SB.77), we get

$$n^{-1} \sum_{i=1}^{n} \left| u_{2,i} (\hat{\phi}_i - \phi_{m_1,i})^2 \right| = O_p(m_1 n^{-1}). \tag{SB.78}$$

By Assumptions SB1(i, iv) and SB2(i), and the Markov inequality

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

which together with (SB.78) and Assumption SB2(vi) implies that

$$n^{-1} \sum_{i=1}^{n} |u_{2,i}| \, (\hat{\phi}_i - \phi_i)^2 = O_p(m_1 n^{-1}). \tag{SB.79}$$

By the Cauchy-Schwarz inequality and the triangle inequality, Assumption SB2(v) and (SB.79)

$$\left| n^{-1} \sum_{i=1}^{n} \upsilon_2' \partial^2 \tilde{P}_2 \left(\tilde{\omega}_{1,i}; \beta_{k,0} \right) (\hat{\phi}_i - \phi_i)^2 u_{2,i} \right| \leq \|\upsilon_2\| \, O_p(\xi_{2,m_2} m_1 n^{-1}). \tag{SB.80}$$

By Assumptions SB1(i, v) and SB2(v), and (SB.75),

$$\mathbb{E}\left[\left\| n^{-1} \sum_{i=1}^{n} u_{2,i} \partial^1 \tilde{P}_{2,i} \left(\beta_{k,0} \right) P_1(x_{1,i})' \right\|^2 \right] \leq C \xi_{1,m_2}^2 m_1 n^{-1}$$

which together with the Cauchy-Schwarz inequality, the Markov inequality and (SB.42) implies that

$$\left| n^{-1} \sum_{i=1}^{n} u_{2,i} \upsilon_2' \partial^1 \tilde{P}_{2,i} (\beta_{k,0}) P_1(x_{1,i})' (\hat{\beta}_\phi - \beta_{\phi,m_1}) \right|$$
$$\leq \|\upsilon_2\| \left\| \hat{\beta}_\phi - \beta_{\phi,m_1} \right\| \left\| n^{-1} \sum_{i=1}^{n} u_{2,i} \partial^1 \tilde{P}_{2,i} (\beta_{k,0}) P_1(x_{1,i})' \right\| = \|\upsilon_2\| O_p(\xi_{1,m_2} m_1 n^{-1}). \quad \text{(SB.81)}$$

By Assumptions SB1(i) and SB2(iii, v, vi), and (SB.75),

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

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

$$\left| n^{-1} \sum_{i=1}^{n} \upsilon_2' \partial^1 \tilde{P}_{2,i} (\beta_{k,0}) (\phi_{m_2,i} - \phi_i) u_{2,i} \right| \leq \|\upsilon_2\| O_p(\xi_{1,m_2} n^{-1}). \quad \text{(SB.82)}$$

Collecting the results in (SB.81) and (SB.82) obtains

$$\left| n^{-1} \sum_{i=1}^{n} \upsilon_2' \partial^1 \tilde{P}_{2,i} (\beta_{k,0}) (\hat{\phi}_i - \phi_i) u_{2,i} \right| \leq \|\upsilon_2\| O_p(\xi_{1,m_2} m_1 n^{-1}). \quad \text{(SB.83)}$$

Therefore, from Assumptions SB2(vi) and SB3(iv), (SB.53), (SB.74), (SB.80) and (SB.83) we can deduce

$$(\hat{\mathbf{P}}_2(\beta_{k,0})' \hat{\mathbf{P}}_2(\beta_{k,0}))^{-1} \sum_{i=1}^{n} (\hat{P}_{2,i}(\beta_{k,0}) - \tilde{P}_{2,i}(\beta_{k,0})) u_{2,i} = O_p(m_1^{1/2} n^{-1/2}). \quad \text{(SB.84)}$$

By Assumptions SB1(i) and SB2(v), and (SB.75),

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

which together with (SB.53) implies that

$$(\hat{\mathbf{P}}_2(\beta_{k,0})' \hat{\mathbf{P}}_2(\beta_{k,0}))^{-1} \sum_{i=1}^{n} \tilde{P}_{2,i}(\beta_{k,0}) u_{2,i} = O_p(m_2^{1/2} n^{-1/2}). \quad \text{(SB.85)}$$

The claim of the lemma follows from (SB.73), (SB.84) and (SB.85). *Q.E.D.*

Lemma SB7 *Under Assumptions SB1, SB2 and SB3, we have:*

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

Proof of Lemma SB7. By the definition of $\hat{\tau}_i(\beta_{k,0})$, we can write

$$\begin{aligned}
&n^{-1}\sum_{i=1}^{n}\hat{\tau}_i(\beta_{k,0})k_{1,i}(\hat{g}_{1,i}(\beta_{k,0})-g_1(\omega_{1,i}))\\
&=n^{-1}\sum_{i=1}^{n}k_{1,i}(g(\omega_{1,i})-\hat{g}_i(\beta_{k,0}))(\hat{g}_{1,i}(\beta_{k,0})-g_1(\omega_{1,i}))\\
&+n^{-1}\sum_{i=1}^{n}k_{1,i}(\hat{y}^{*}_{2,i}(\beta_{k,0})-g(\omega_{1,i}))(\hat{g}_{1,i}(\beta_{k,0})-g_1(\omega_{1,i})). \qquad \text{(SB.86)}
\end{aligned}$$

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, (SB.53), (SB.57) and Lemma SB6

$$\begin{aligned}
&n^{-1}\sum_{i=1}^{n}(\hat{g}_i(\beta_{k,0})-g(\omega_{1,i}))^2\\
&\leq Cn^{-1}\sum_{i=1}^{n}(\hat{P}_{2,i}(\beta_{k,0})'(\hat{\beta}_g(\beta_{k,0})-\tilde{\beta}_{g,m_2}(\beta_{k,0})))^2\\
&+Cn^{-1}\sum_{i=1}^{n}(\hat{P}_{2,i}(\beta_{k,0})'\tilde{\beta}_{g,m_2}(\beta_{k,0})-g(\omega_{1,i})))^2\\
&\leq 2\left\|\hat{\beta}_g(\beta_{k,0})-\tilde{\beta}_{g,m_2}(\beta_{k,0})\right\|^2\lambda_{\max}(n^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))+O_p(m_1n^{-1})\\
&=O_p((m_1+m_2)n^{-1}). \qquad \text{(SB.87)}
\end{aligned}$$

Similarly, we can show that

$$
\begin{aligned}
& n^{-1}\sum_{i=1}^{n}(\hat{g}_{1,i}(\beta_{k,0})-g_1(\omega_{1,i}))^2 \\
& \leq Cn^{-1}\sum_{i=1}^{n}(\partial^1\hat{P}_{2,i}(\beta_{k,0})'(\hat{\beta}_g(\beta_{k,0})-\tilde{\beta}_{g,m_2}(\beta_{k,0})))^2 \\
& + Cn^{-1}\sum_{i=1}^{n}((\partial^1\hat{P}_{2,i}(\beta_{k,0})-\partial^1\tilde{P}_{2,i}(\beta_{k,0}))'\tilde{\beta}_{g,m_2}(\beta_{k,0}))^2 \\
& + Cn^{-1}\sum_{i=1}^{n}(\partial^1\tilde{P}_{2,i}(\beta_{k,0})'\tilde{\beta}_{g,m_2}(\beta_{k,0})-g_1(\omega_{1,i}))^2 \\
& \leq C\xi_{1,m_2}^2\left\|\hat{\beta}_g(\beta_{k,0})-\tilde{\beta}_{g,m_2}(\beta_{k,0})\right\|^2+O_p(\xi_{2,m_2}^2 m_1 n^{-1}) \\
& = O_p(\xi_{1,m_2}^2(m_1+m_2)n^{-1}+\xi_{2,m_2}^2 m_1 n^{-1}). \qquad \text{(SB.88)}
\end{aligned}
$$

Therefore, by the Cauchy-Schwarz inequality, Assumption SB3(iv), (SB.87) and (SB.88),

$$
n^{-1}\sum_{i=1}^{n}k_{1,i}(\hat{g}_i(\beta_{k,0})-g(\omega_{1,i}))(\hat{g}_{1,i}(\beta_{k,0})-g_1(\omega_{1,i}))=o_p(n^{-1/2}). \qquad \text{(SB.89)}
$$

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

$$
\begin{aligned}
& n^{-1}\sum_{i=1}^{n}(\hat{y}^*_{2,i}(\beta_{k,0})-g(\omega_{1,i}))k_{1,i}(\hat{g}_{1,i}(\beta_{k,0})-g_1(\omega_{1,i})) \\
& = n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(\hat{g}_{1,i}(\beta_{k,0})-g_1(\omega_{1,i})) \\
& \quad -(\hat{\beta}_l-\beta_{l,0})n^{-1}\sum_{i=1}^{n}l_{1,i}k_{1,i}(\hat{g}_{1,i}(\beta_{k,0})-g_1(\omega_{1,i})). \qquad \text{(SB.90)}
\end{aligned}
$$

Since $k_{1,i}$ has bounded support, by Assumptions SB1(i, ii, iii), SB2(vi) and SB3(iv), (SB.88) and the Markov inequality,

$$
(\hat{\beta}_l-\beta_{l,0})n^{-1}\sum_{i=1}^{n}l_{1,i}k_{1,i}(\hat{g}_{1,i}(\beta_{k,0})-g_1(\omega_{1,i}))=o_p(n^{-1/2}). \qquad \text{(SB.91)}
$$

Let

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

Then we can write

$$
\begin{aligned}
& n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(\hat{g}_{1,i}(\beta_{k,0})-g_{1}(\omega_{1,i})) \\
& = n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(\partial^{1}\hat{P}_{2,i}(\beta_{k,0})-\partial^{1}\tilde{P}_{2,i}(\beta_{k,0}))'\hat{\beta}_{g}(\beta_{k,0}) \\
& + n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}\partial^{1}\tilde{P}_{2,i}(\beta_{k,0})'(\hat{\beta}_{g}(\beta_{k,0})-\tilde{\beta}_{g,m_2}(\beta_{k,0})) \\
& + n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}\left(\partial^{1}\tilde{P}_{2,i}(\beta_{k,0})'\tilde{\beta}_{g,m_2}(\beta_{k,0})-g_{1}(\omega_{1,i})\right). \qquad \text{(SB.92)}
\end{aligned}
$$

By Assumptions SB1(i) and SB2(iii), (SB.75) and the Markov inequality, we have

$$
n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}\left(\partial^{1}\tilde{P}_{2,i}(\beta_{k,0})'\tilde{\beta}_{g,m_2}(\beta_{k,0})-g_{1}(\omega_{1,i})\right)=o_{p}(n^{-1/2}). \qquad \text{(SB.93)}
$$

Similarly,

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

which together with Assumptions SB2(vi) and SB3(iv), and Lemma SB6 implies that

$$
n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}\partial^{1}\tilde{P}_{2,i}(\beta_{k,0})'(\hat{\beta}_{g}(\beta_{k,0})-\tilde{\beta}_{g,m_2}(\beta_{k,0}))=o_{p}(n^{-1/2}). \qquad \text{(SB.94)}
$$

By Assumption SB1(i), (SB.75) and the Markov inequality

$$
n^{-1}\sum_{i=1}^{n}u_{2,i}^{2}k_{1,i}^{2}=O_{p}(1). \qquad \text{(SB.95)}
$$

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

$$
\begin{aligned}
& n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(\partial^{1}\hat{P}_{2,i}(\beta_{k,0})-\partial^{1}\tilde{P}_{2,i}(\beta_{k,0}))'\hat{\beta}_{g}(\beta_{k,0}) \\
& = n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(\hat{\phi}_i-\phi_i)\partial^{2}\tilde{P}_{2,i}(\beta_{k,0})'\hat{\beta}_{g}(\beta_{k,0}) \\
& + n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(\hat{\phi}_i-\phi_i)^{2}\partial^{3}\tilde{P}_{2}(\tilde{\omega}_{1,i}(\beta_{k,0});\beta_{k,0})'\hat{\beta}_{g}(\beta_{k,0}), \qquad \text{(SB.96)}
\end{aligned}
$$

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

SB2(v), Lemma SB1, (SB.65) and (SB.95)

$$n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(\hat{\phi}_i-\phi_i)^2\partial^3\tilde{P}_2(\tilde{\omega}_{1,i}(\beta_{k,0});\beta_{k,0})'\hat{\beta}_g(\beta_{k,0})=O_p(\xi_{3,m_2}m_1n^{-1})=o_p(n^{-1/2}), \tag{SB.97}$$

where the second equality is by Assumptions SB2(vi) and SB3(iv). By Assumptions SB1(i, v) and SB2(v), and (SB.75)

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

which together with Lemma SB1 and (SB.65) implies that

$$n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(\hat{\phi}_i-\phi_i)\partial^2\tilde{P}_{2,i}(\beta_{k,0})'\hat{\beta}_g(\beta_{k,0})=O_p(\xi_{2,m_2}m_1n^{-1})=o_p(n^{-1/2}), \tag{SB.98}$$

where the second equality is by Assumptions SB2(vi) and SB3(iv). Similarly, we can show that

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

which together with (SB.98) implies that

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

Collecting the results in (SB.96), (SB.97) and (SB.99) we get

$$n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(\partial^1\hat{P}_{2,i}(\beta_{k,0})-\partial^1\tilde{P}_{2,i}(\beta_{k,0}))'\hat{\beta}_g(\beta_{k,0})=o_p(n^{-1/2}). \tag{SB.100}$$

By (SB.90), (SB.91), (SB.92), (SB.93), (SB.94) and (SB.100),

$$n^{-1}\sum_{i=1}^{n}(\hat{y}^*_{2,i}(\beta_{k,0})-g(\omega_{1,i}))k_{1,i}(\hat{g}_{1,i}(\beta_{k,0})-g_1(\omega_{1,i}))=o_p(n^{-1/2}). \tag{SB.101}$$

The claim of the lemma follows from (SB.86), (SB.89) and (SB.101). *Q.E.D.*

Lemma SB8 *Under Assumptions SB1, SB2 and SB3, we have*

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

Proof of Lemma SB8. First we write

$$n^{-1}\sum_{i=1}^{n}(\hat{P}_{2,i}(\beta_{k,0})'\tilde{\beta}_{g,m_2}(\beta_{k,0})-g(\omega_{1,i}))(k_{2,i}-k_{1,i}g_1(\omega_{1,i}))$$
$$=n^{-1}\sum_{i=1}^{n}(\hat{P}_{2,i}(\beta_{k,0})-\tilde{P}_{2,i}(\beta_{k,0}))'\tilde{\beta}_{g,m_2}(\beta_{k,0})(k_{2,i}-k_{1,i}g_1(\omega_{1,i}))$$
$$+n^{-1}\sum_{i=1}^{n}(\tilde{P}_{2,i}(\beta_{k,0})'\tilde{\beta}_{g,m_2}(\beta_{k,0})-g(\omega_{1,i}))(k_{2,i}-k_{1,i}g_1(\omega_{1,i})). \tag{SB.102}$$

By Assumptions SB1(i) and SB2(i, ii), and the Markov inequality

$$n^{-1}\sum_{i=1}^{n}(k_{2,i}-k_{1,i}g_1(\omega_{1,i}))^2=O_p(1). \tag{SB.103}$$

Therefore by Assumption SB2(iii, vi) and (SB.103), we have

$$n^{-1}\sum_{i=1}^{n}(\tilde{P}_{2,i}(\beta_{k,0})'\tilde{\beta}_{g,m_2}(\beta_{k,0})-g(\omega_{1,i}))(k_{2,i}-k_{1,i}g_1(\omega_{1,i}))=o_p(n^{-1/2}). \tag{SB.104}$$

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

$$n^{-1}\sum_{i=1}^{n}(\hat{P}_{2,i}(\beta_{k,0})-\tilde{P}_{2,i}(\beta_{k,0}))'\tilde{\beta}_{g,m_2}(\beta_{k,0})(k_{2,i}-k_{1,i}g_1(\omega_{1,i}))$$
$$=n^{-1}\sum_{i=1}^{n}\partial^1\tilde{P}_{2,i}(\beta_{k,0})'\tilde{\beta}_{g,m_2}(\beta_{k,0})(\hat{\phi}_i-\phi_i)(k_{2,i}-k_{1,i}g_1(\omega_{1,i}))$$
$$+n^{-1}\sum_{i=1}^{n}\partial^2\tilde{P}_2(\tilde{\omega}_{1,i}(\beta_{k,0});\beta_{k,0})'\tilde{\beta}_{g,m_2}(\beta_{k,0})(\hat{\phi}_i-\phi_i)^2(k_{2,i}-k_{1,i}g_1(\omega_{1,i})). \tag{SB.105}$$

By the Cauchy-Schwarz inequality and the triangle inequality, Assumption SB2(v), (SB.46) and (SB.56)

$$\left| n^{-1} \sum_{i=1}^{n} \partial^2 \tilde{P}_2(\tilde{\omega}_{1,i}(\beta_{k,0}); \beta_{k,0})' \tilde{\beta}_{g,m_2}(\beta_{k,0})(\hat{\phi}_i - \phi_i)^2 (k_{2,i} - k_{1,i} g_1(\omega_{1,i})) \right|$$
$$\leq O_p(\xi_{2,m_2}) n^{-1} \sum_{i=1}^{n} |k_{2,i} - k_{1,i} g_1(\omega_{1,i})| (\hat{\phi}_i - \phi_i)^2. \quad \text{(SB.106)}$$

Since by Assumption SB2(i, ii) $\mathbb{E}[|k_{2,i} - k_{1,i} g_1(\omega_{1,i})|^2 |x_{1,i}] \leq C$, we can use similar arguments for showing (SB.79) to get

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

which combined with Assumption SB3(iv) and (SB.106) implies that

$$n^{-1} \sum_{i=1}^{n} \partial^2 \tilde{P}_2(\tilde{\omega}_{1,i}(\beta_{k,0}); \beta_{k,0})' \tilde{\beta}_{g,m_2}(\beta_{k,0})(\hat{\phi}_i - \phi_i)^2 (k_{2,i} - k_{1,i} g_1(\omega_{1,i})) = o_p(n^{-1/2}). \quad \text{(SB.107)}$$

By the Cauchy-Schwarz inequality, Assumption SB2(iii), Lemma SB1 and (SB.103)

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

which together with (SB.105) and (SB.107) shows that

$$n^{-1} \sum_{i=1}^{n} (\hat{P}_{2,i}(\beta_{k,0}) - \tilde{P}_{2,i}(\beta_{k,0}))' \tilde{\beta}_{g,m_2}(\beta_{k,0})(k_{2,i} - k_{1,i} g_1(\omega_{1,i}))$$
$$= n^{-1} \sum_{i=1}^{n} g_1(\omega_{1,i})(\hat{\phi}_i - \phi_i)(k_{2,i} - k_{1,i} g_1(\omega_{1,i})) + o_p(n^{-1/2}). \quad \text{(SB.108)}$$

The claim of the lemma follows from (SB.102), (SB.104) and (SB.107). *Q.E.D.*

Lemma SB9 *Under Assumptions SB1, SB2 and SB3, we have*

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

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

Proof of Lemma SB9. By the definition of $\hat{g}_i(\beta_{k,0})$, we can write

$$n^{-1}\sum_{i=1}^{n}(\hat{g}_i(\beta_{k,0})-g(\omega_{1,i}))(k_{2,i}-k_{1,i}g_1(\omega_{1,i}))$$
$$=(\hat{\beta}_g(\beta_{k,0})-\tilde{\beta}_{g,m_2}(\beta_{k,0}))'n^{-1}\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_{k,0})(k_{2,i}-k_{1,i}g_1(\omega_{1,i}))$$
$$+n^{-1}\sum_{i=1}^{n}(\hat{P}_{2,i}(\beta_{k,0})'\tilde{\beta}_{g,m_2}(\beta_{k,0})-g(\omega_{1,i}))(k_{2,i}-k_{1,i}g_1(\omega_{1,i})). \tag{SB.109}$$

In view of Lemma SB8 and (SB.109), the claim of the lemma follows if

$$(\hat{\beta}_g(\beta_{k,0})-\tilde{\beta}_{g,m_2}(\beta_{k,0}))'n^{-1}\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_{k,0})(k_{2,i}-k_{1,i}g_1(\omega_{1,i}))$$
$$=n^{-1}\sum_{i=1}^{n}u_{2,i}\varphi(\omega_{1,i})-\mathbb{E}[l_{2,i}\varphi(\omega_{1,i})](\hat{\beta}_l-\beta_{l,0})$$
$$-n^{-1}\sum_{i=1}^{n}g_1(\omega_{1,i})(\hat{\phi}(x_{1,i})-\phi(x_{1,i}))\varphi(\omega_{1,i})+o_p(n^{-1/2}). \tag{SB.110}$$

We next prove (SB.110).

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

$$(\hat{\beta}_g(\beta_{k,0})-\tilde{\beta}_{g,m_2}(\beta_{k,0}))'n^{-1}\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_{k,0})(k_{2,i}-k_{1,i}g_1(\omega_{1,i}))$$
$$=(\hat{\beta}_g(\beta_{k,0})-\tilde{\beta}_{g,m_2}(\beta_{k,0}))'(n^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))\hat{\beta}_\varphi(\beta_{k,0}). \tag{SB.111}$$

Under Assumptions SB1, SB2 and SB3, we can use the same arguments for proving Lemma SB6 to show that

$$\hat{\beta}_{\varphi}(\beta_{k,0}) - \tilde{\beta}_{\varphi,m_2}(\beta_{k,0}) = O_p((m_1^{1/2} + m_2^{1/2})n^{-1/2}) = o_p(1), \tag{SB.112}$$

where $\tilde{\beta}_{\varphi,m_2}(\beta_{k,0}) \equiv (B(\beta_{k,0})')^{-1}\beta_{\varphi,m_2}$ and β_{φ,m_2} is defined in Assumption SB3(ii). By Assumptions SB2(v, vi) and SB3, Lemma SB1, Lemma SB6, (SB.53) and (SB.112)

$$\begin{aligned}
&(\hat{\beta}_g(\beta_{k,0}) - \tilde{\beta}_{g,m_2}(\beta_{k,0}))'(n^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))\hat{\beta}_{\varphi}(\beta_{k,0}) \\
&= (\hat{\beta}_g(\beta_{k,0}) - \tilde{\beta}_{g,m_2}(\beta_{k,0}))'(n^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))\tilde{\beta}_{\varphi,m_2}(\beta_{k,0}) + o_p(n^{-1/2}) \\
&= (\hat{\beta}_g(\beta_{k,0}) - \tilde{\beta}_{g,m_2}(\beta_{k,0}))'n^{-1}\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_{k,0})\varphi(\hat{\omega}_{1,i}(\beta_{k,0})) + o_p(n^{-1/2}) \\
&= (\hat{\beta}_g(\beta_{k,0}) - \tilde{\beta}_{g,m_2}(\beta_{k,0}))'n^{-1}\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_{k,0})\varphi(\omega_{1,i}) + o_p(n^{-1/2}).
\end{aligned} \tag{SB.113}$$

Using the decomposition in (SB.52), we can write

$$\begin{aligned}
&(\hat{\beta}_g(\beta_{k,0}) - \tilde{\beta}_{g,m_2}(\beta_{k,0}))'n^{-1}\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_{k,0})\varphi(\omega_{1,i}) \\
&= \frac{\varphi_n'\hat{\mathbf{P}}_2(\beta_{k,0})(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}}{n}\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_k)(g(\omega_{1,i}) - g_{m_2}(\hat{\omega}_{1,i}(\beta_{k,0}), \beta_{k,0})) \\
&- (\hat{\beta}_l - \beta_{l,0})\frac{\varphi_n'\hat{\mathbf{P}}_2(\beta_{k,0})(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}}{n}\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_k)l_{2,i} \\
&+ \frac{\varphi_n'\hat{\mathbf{P}}_2(\beta_{k,0})(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}}{n}\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_{k,0})u_{2,i},
\end{aligned} \tag{SB.114}$$

where $\varphi_n \equiv (\varphi(\omega_{1,1}), \ldots, \varphi(\omega_{1,n}))'$. The rest of the proof is divided into 3 steps. The claim in (SB.110) follows from (SB.111), (SB.113), (SB.114), (SB.115), (SB.119) and (SB.125) below.

Step 1. In this step, we show that

$$\begin{aligned}
&\frac{\varphi_n'\hat{\mathbf{P}}_2(\beta_{k,0})(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}}{n}\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_k)(g(\omega_{1,i}) - g_{m_2}(\hat{\omega}_{1,i}(\beta_{k,0}), \beta_{k,0})) \\
&= -n^{-1}\sum_{i=1}^{n}g_1(\omega_{1,i})(\hat{\phi}(x_{1,i}) - \phi(x_{1,i}))\varphi(\omega_{1,i}) + o_p(n^{-1/2}).
\end{aligned} \tag{SB.115}$$

Recall that $\hat{\phi}_i \equiv \hat{\phi}(x_{1,i})$ and $\phi_i \equiv \phi(x_{1,i})$. By the second order expansion, Assumptions SB2(iii, v, vi) and SB3(iv), Lemma SB1, (SB.46), (SB.53) and (SB.56),

$$
\begin{aligned}
&\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_{k,0})(g_{m_2}(\hat{\omega}_{1,i}(\beta_{k,0}),\beta_{k,0})-g(\omega_{1,i}))\\
&=\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_{k,0})(\hat{P}_{2,i}(\beta_{k,0})-\tilde{P}_{2,i}(\beta_{k,0}))'\tilde{\beta}_{g,m_2}(\beta_{k,0})+o_p(n^{-1/2})\\
&=\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_{k,0})(\hat{\phi}_i-\phi_i)\partial^1\tilde{P}_{2,i}(\beta_{k,0})'\tilde{\beta}_{g,m_2}(\beta_{k,0})\\
&+\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_{k,0})(\hat{\phi}_i-\phi_i)^2\partial^2\tilde{P}_2(\tilde{\omega}_{1,i}(\beta_{k,0});\beta_{k,0})'\tilde{\beta}_{g,m_2}(\beta_{k,0})+o_p(n^{-1/2})\\
&=\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_{k,0})(\hat{\phi}_i-\phi_i)g_1(\omega_{1,i})+o_p(n^{-1/2}). \qquad \text{(SB.116)}
\end{aligned}
$$

By Assumptions SB1(i) and SB3(i), and (SB.53),

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

which together with (SB.53) and (SB.116) implies that

$$
\begin{aligned}
&\frac{\varphi_n'\hat{\mathbf{P}}_2(\beta_{k,0})(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}}{n}\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_k)(g(\omega_{1,i})-g_{m_2}(\hat{\omega}_{1,i}(\beta_{k,0}),\beta_{k,0}))\\
&=\frac{\varphi_n'\hat{\mathbf{P}}_2(\beta_{k,0})(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}}{n}\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_{k,0})g_1(\omega_{1,i})(\hat{\phi}_i-\phi_i)+o_p(n^{-1/2}). \qquad \text{(SB.117)}
\end{aligned}
$$

By Assumptions SB2(ii) and SB3(ii, iv), Lemma SB1 and (SB.53)

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

which further implies that

$$
\begin{aligned}
& n^{-1}\sum_{i=1}^{n}(\hat{\phi}_i-\phi_i)g_1(\omega_{1,i})\hat{P}_{2,i}(\beta_{k,0})'(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'\varphi_n \\
&= n^{-1}\sum_{i=1}^{n}(\hat{\phi}_i-\phi_i)g_1(\omega_{1,i})\hat{P}_{2,i}(\beta_{k,0})'\tilde{\beta}_{\varphi,m_2}(\beta_{k,0})+o_p(n^{-1/2}) \\
&= n^{-1}\sum_{i=1}^{n}(\hat{\phi}_i-\phi_i)g_1(\omega_{1,i})\varphi(\hat{\omega}_{1,i}(\beta_{k,0}))+o_p(n^{-1/2}) \\
&= n^{-1}\sum_{i=1}^{n}(\hat{\phi}_i-\phi_i)g_1(\omega_{1,i})\varphi(\omega_{1,i})+o_p(n^{-1/2}). \qquad \text{(SB.118)}
\end{aligned}
$$

The claim in (SB.115) follows from (SB.117) and (SB.118).

Step 2. In this step, we show that

$$
\frac{\varphi_n'\hat{\mathbf{P}}_2(\beta_{k,0})(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}}{n}\hat{\mathbf{P}}_2(\beta_{k,0})'\mathbf{L}_2(\hat{\beta}_l-\beta_{l,0}) = \mathbb{E}[l_{2,i}\varphi(\omega_{1,i})](\hat{\beta}_l-\beta_{l,0})+o_p(n^{-1/2}). \quad \text{(SB.119)}
$$

By Assumptions SB1(i) and SB2(i), and (SB.53),

$$
n^{-1}\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_{k,0})l_{2,i} = O_p(1). \qquad \text{(SB.120)}
$$

Using the similar arguments for showing (SB.57), we get

$$
n^{-1}\sum_{i=1}^{n}\left(\hat{P}_{2,i}(\beta_{k,0})'\tilde{\beta}_{\varphi,m_2}(\beta_{k,0})-\varphi(\omega_{1,i})\right)^2 = O_p(m_1 n^{-1})
$$

which together with (SB.53) and (SB.120) implies that

$$
n^{-1}(\varphi_n-\hat{\varphi}_{m_2,n})'\hat{\mathbf{P}}_2(\beta_{k,0})(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'\mathbf{L}_2 = O_p(m_1^{1/2}n^{-1/2}), \qquad \text{(SB.121)}
$$

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

$$
n^{-1}\varphi_n'\hat{\mathbf{P}}_2(\beta_{k,0})(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'\mathbf{L}_2 = n^{-1}\sum_{i=1}^{n}l_{2,i}\hat{P}_{2,i}(\beta_{k,0})'\tilde{\beta}_{\varphi,m_2}(\beta_{k,0})+O_p(m_1^{1/2}n^{-1/2}). \qquad \text{(SB.122)}
$$

By the first order expansion,

$$
\begin{aligned}
& n^{-1}\sum_{i=1}^{n} l_{2,i}\hat{P}_{2,i}(\beta_{k,0})'\tilde{\beta}_{\varphi,m_2}(\beta_{k,0}) \\
&= n^{-1}\sum_{i=1}^{n} l_{2,i}\varphi(\omega_{1,i}) + n^{-1}\sum_{i=1}^{n} l_{2,i}(\varphi_{m_2}(\omega_{1,i}) - \varphi(\omega_{1,i})) \\
&+ n^{-1}\sum_{i=1}^{n} l_{2,i}(\hat{\phi}_i - \phi_i)\partial^1\tilde{P}_2(\tilde{\omega}_{1,i}(\beta_{k,0});\beta_{k,0})'\tilde{\beta}_{\varphi,m_2}(\beta_{k,0}) \\
&= n^{-1}\sum_{i=1}^{n} l_{2,i}\varphi(\omega_{1,i}) + O_p(\xi_{1,m_2} m_1^{1/2} n^{-1/2}), \qquad \text{(SB.123)}
\end{aligned}
$$

where the second equality is by Assumptions SB1(i), SB2(i, v, vi) and SB3(ii, iv), Lemma SB1, (SB.46) and (SB.56). Collecting the results in (SB.122) and (SB.123), we deduce that

$$
\begin{aligned}
& n^{-1}\varphi_n'\hat{\mathbf{P}}_2(\beta_{k,0})(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'\mathbf{L}_2 \\
&= n^{-1}\sum_{i=1}^{n} l_{2,i}\varphi(\omega_{1,i}) + O_p(\xi_{1,m_2} m_1^{1/2} n^{-1/2}) = \mathbb{E}[l_{2,i}\varphi(\omega_{1,i})] + o_p(1), \qquad \text{(SB.124)}
\end{aligned}
$$

where the second equality is by the Markov inequality, Assumptions SB1(i), SB2(i) and SB3(i, iv). The claim in (SB.119) follows from Assumption SB1(iii) and (SB.124).

Step 3. In this step, we show that

$$
\frac{\varphi_n'\hat{\mathbf{P}}_2(\beta_{k,0})(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}}{n}\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_{k,0})u_{2,i} = n^{-1}\sum_{i=1}^{n} u_{2,i}\varphi(\omega_{1,i}) + o_p(n^{-1/2}). \qquad \text{(SB.125)}
$$

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

$$
\begin{aligned}
\sum_{i=1}^{n}\upsilon_2'\hat{P}_{2,i}(\beta_{k,0})u_{2,i} &= \sum_{i=1}^{n}\upsilon_2'\tilde{P}_{2,i}(\beta_{k,0})u_{2,i} \\
&+ \sum_{i=1}^{n}\upsilon_2'\partial^1\tilde{P}_{2,i}(\beta_{k,0})(\hat{\phi}_i - \phi_i)u_{2,i} \\
&+ \sum_{i=1}^{n}\upsilon_2'\partial^2\tilde{P}_2(\tilde{\omega}_{1,i}(\beta_{k,0});\beta_{k,0})(\hat{\phi}_i - \phi_i)^2 u_{2,i}. \qquad \text{(SB.126)}
\end{aligned}
$$

By the Markov inequality, Assumptions SB1(i, iv, v), SB2(v, vi) and SB3(iv), Lemma SB1 and (SB.75), we can show that

$$
n^{-1}\sum_{i=1}^{n}\upsilon_2'\partial^1\tilde{P}_{2,i}(\beta_{k,0})(\hat{\phi}_i - \phi_i)u_{2,i} = \|\upsilon_2\|\, o_p(n^{-1/2}). \qquad \text{(SB.127)}
$$

By the Cauchy-Schwarz inequality, Assumptions SB2(v) and SB3(iv), and (SB.46)

$$n^{-1}\sum_{i=1}^{n}\upsilon_2'\partial^2\tilde{P}_2(\tilde{\omega}_{1,i}(\beta_{k,0});\beta_{k,0})(\hat{\phi}_i-\phi_i)^2u_{2,i}=\|\upsilon_2\|\,o_p(n^{-1/2}). \tag{SB.128}$$

By Assumptions SB1(i) and SB3(i), and (SB.53),

$$n^{-1}\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_{k,0})\varphi(\omega_{1,i})=O_p(1). \tag{SB.129}$$

Combining the results in (SB.53), (SB.126), (SB.127), (SB.128) and (SB.129), we get

$$\begin{aligned}&\frac{\varphi_n'\hat{\mathbf{P}}_2(\beta_{k,0})(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}}{n}\sum_{i=1}^{n}\hat{P}_{2,i}(\beta_{k,0})u_{2,i}\\&=\frac{\varphi_n'\hat{\mathbf{P}}_2(\beta_{k,0})(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}}{n}\sum_{i=1}^{n}\tilde{P}_{2,i}(\beta_{k,0})u_{2,i}+o_p(n^{-1/2}).\end{aligned} \tag{SB.130}$$

Since $n^{-1}\sum_{i=1}^{n}\tilde{P}_{2,i}(\beta_{k,0})u_{2,i}=O_p(m_2^{1/2}n^{-1/2})$ by the Markov inequality, Assumptions SB1(i) and SB2(iv), and (SB.75), we can use similar arguments for showing (SB.121) to get

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

where the second equality is by Assumption SB3(iv). Therefore

$$\begin{aligned}&\frac{\varphi_n'\hat{\mathbf{P}}_2(\beta_{k,0})(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}}{n}\sum_{i=1}^{n}\tilde{P}_{2,i}(\beta_{k,0})u_{2,i}\\&=n^{-1}\sum_{i=1}^{n}\tilde{P}_{2,i}(\beta_{k,0})'\tilde{\beta}_{\varphi,m_2}(\beta_{k,0})u_{2,i}+o_p(n^{-1/2})\\&=n^{-1}\sum_{i=1}^{n}u_{2,i}\varphi(\omega_{1,i})+o_p(n^{-1/2}),\end{aligned} \tag{SB.131}$$

where the second equality is by the Markov inequality, Assumptions SB1(i) and SB3(ii, iv), and (SB.75). The claim in (SB.125) follows from (SB.130) and (SB.131). *Q.E.D.*

Lemma SB10 *Under Assumptions SB1, SB2 and SB3, we have*

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

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

Proof of Lemma SB10. Since $\hat{\phi}(x_{1,i})-\phi(x_{1,i})=(\hat{\beta}_\phi-\beta_{\phi,m_1})'P_1(x_{1,i})+\phi_{m_1}(x_{1,i})-\phi(x_{1,i})$, we can write

$$
\begin{aligned}
&n^{-1}\sum_{i=1}^{n} g_{1,i}(\hat{\phi}(x_{1,i})-\phi(x_{1,i}))(v_{2,i}-v_{1,i}g_{1,i})\\
&=(\hat{\beta}_\phi-\beta_{\phi,m_1})'n^{-1}\sum_{i=1}^{n}P_1(x_{1,i})g_{1,i}(v_{2,i}-v_{1,i}g_{1,i})\\
&\quad+n^{-1}\sum_{i=1}^{n}g_{1,i}(\phi_{m_1}(x_{1,i})-\phi(x_{1,i}))(v_{2,i}-v_{1,i}g_{1,i}),
\end{aligned}
\tag{SB.132}
$$

where $g_{1,i}\equiv g_1(\omega_{1,i})$. By Assumptions SB1(i, iv, vi) and SB2(ii), and the Markov inequality

$$
n^{-1}\sum_{i=1}^{n}g_{1,i}(\phi_{m_1}(x_{1,i})-\phi(x_{1,i})))(v_{2,i}-v_{1,i}g_{1,i})=o_p(n^{-1/2}).
\tag{SB.133}
$$

By Assumptions SB1(i, v, vi) and SB2(i, ii), and the Markov inequality

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

which together with the LIE, Assumption SB3(iv) and (SB.42) implies that

$$
\begin{aligned}
&(\hat{\beta}_\phi-\beta_{\phi,m_1})'n^{-1}\sum_{i=1}^{n}P_1(x_{1,i})g_{1,i}(v_{2,i}-v_{1,i}g_{1,i})\\
&=n^{-1}\sum_{i=1}^{n}\eta_{1,i}P_1(x_{1,i})'Q_{m_1}^{-1}\mathbb{E}\left[P_1(x_{1,i})g_{1,i}(v_{2,i}^{*}-v_{1,i}g_{1,i})\right]\\
&-(\hat{\beta}_l-\beta_{l,0})\mathbb{E}[h_1(x_{1,i})P_1(x_{1,i})']Q_{m_1}^{-1}\mathbb{E}\left[P_1(x_{1,i})g_{1,i}(v_{2,i}-v_{1,i}g_{1,i})\right]+o_p(n^{-1/2}).
\end{aligned}
\tag{SB.134}
$$

By Assumptions SB1(i, ii, v), SB2(i, ii) and SB3(iii)

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

which together with the Markov inequality implies that

$$n^{-1}\sum_{i=1}^{n}\eta_{1,i}P_1(x_{1,i})'Q_{m_1}^{-1}\mathbb{E}[P_{1,i}g_{1,i}(v_{2,i}^* - v_{1,i}g_{1,i})] = n^{-1}\sum_{i=1}^{n}\eta_{1,i}g_{1,i}(v_{2,i}^* - v_{1,i}g_{1,i}) + o_p(n^{-1/2}). \quad \text{(SB.135)}$$

By Hölder's inequality, Assumptions SB1(ii, v), SB2(ii) and SB3(iii)

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

which combined with Assumption SB1(iii) implies that

$$(\hat{\beta}_l - \beta_{l,0})\mathbb{E}[l_{1,i}P_1(x_{1,i})']Q_{m_1}^{-1}\mathbb{E}\left[P_1(x_{1,i})g_{1,i}(v_{2,i} - v_{1,i}g_{1,i})\right]$$
$$= (\hat{\beta}_l - \beta_{l,0})\mathbb{E}[l_{1,i}P_1(x_{1,i})']Q_{m_1}^{-1}\mathbb{E}\left[P_1(x_{1,i})g_{1,i}(v_{2,i}^* - v_{1,i}g_{1,i})\right]$$
$$= (\hat{\beta}_l - \beta_{l,0})\mathbb{E}[l_{1,i}g_{1,i}(v_{2,i}^* - v_{1,i}g_{1,i})] + o_p(n^{-1/2})$$
$$= (\hat{\beta}_l - \beta_{l,0})\mathbb{E}[h_1(x_{1,i})g_{1,i}(v_{2,i} - v_{1,i}g_{1,i})] + o_p(n^{-1/2}). \quad \text{(SB.136)}$$

The claim of the lemma follows from (SB.132), (SB.133), (SB.134), (SB.135) and (SB.136). *Q.E.D.*

Lemma SB11 *Under Assumptions SB1, SB2 and SB3, we have*

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

Proof of Lemma SB11. Using the decomposition in (SB.52), and applying the results in (SB.58), (SB.60) and (SB.64), we have

$$\hat{\beta}_g(\hat{\beta}_k) - \tilde{\beta}_{g,m_2}(\hat{\beta}_k) = (\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k))^{-1} \sum_{i=1}^{n} (\hat{P}_{2,i}(\hat{\beta}_k) - \tilde{P}_{2,i}(\hat{\beta}_k)) u_{2,i}(\hat{\beta}_k) + O_p((m_1^{1/2} + m_2) n^{-1/2}). \tag{SB.137}$$

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

$$\begin{aligned} & n^{-1} \sum_{i=1}^{n} \upsilon_2' (\hat{P}_{2,i}(\hat{\beta}_k) - \tilde{P}_{2,i}(\hat{\beta}_k)) u_{2,i}(\hat{\beta}_k) \\ &= n^{-1} \sum_{i=1}^{n} \upsilon_2' \partial^1 \tilde{P}_{2,i}(\hat{\beta}_k)(\hat{\phi}_i - \phi_i) u_{2,i}(\hat{\beta}_k) \\ &+ n^{-1} \sum_{i=1}^{n} \upsilon_2' \partial^2 \hat{P}_2(\tilde{\omega}_{1,i}(\hat{\beta}_k); \hat{\beta}_k)(\hat{\phi}_i - \phi_i)^2 u_{2,i}(\hat{\beta}_k), \end{aligned} \tag{SB.138}$$

where $\tilde{\omega}_{1,i}(\hat{\beta}_k)$ lies between $\hat{\omega}_{1,i}(\hat{\beta}_k)$ and $\omega_{1,i}(\hat{\beta}_k)$. By (SB.55) and the compactness of Θ_k,

$$\sup_{\beta_k \in \Theta_k} |u_{2,i}(\beta_k)| \leq C + |y_{2,i}^*| + |k_{2,i}|. \tag{SB.139}$$

Using similar arguments in showing (SB.79), we have

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

which together with the Cauchy-Schwarz inequality, the triangle inequality, Assumptions SB2(vi) and SB3(iv), and (SB.139) implies that

$$\begin{aligned} & \left| n^{-1} \sum_{i=1}^{n} \upsilon_2' \partial^2 \hat{P}_{2,i}(\tilde{\omega}_{1,i}(\hat{\beta}_k); \hat{\beta}_k)(\hat{\phi}_i - \phi_i)^2 u_{2,i}(\hat{\beta}_k) \right| \\ & \leq \|\upsilon_2\| \, \xi_{2,m_2} n^{-1} \sum_{i=1}^{n} (\hat{\phi}_i - \phi_i)^2 (C + |y_{2,i}^*|) = \|\upsilon_2\| \, o_p(m_1^{1/2} n^{-1/2}). \end{aligned} \tag{SB.140}$$

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

$$\begin{aligned} & n^{-1}\sum_{i=1}^{n} \upsilon_2' \partial^1 \tilde{P}_{2,i}(\hat{\beta}_k)(\hat{\phi}_i - \phi_i) u_{2,i}(\hat{\beta}_k) \\ & = n^{-1}\sum_{i=1}^{n} \upsilon_2' \partial^1 \tilde{P}_{2,i}(\hat{\beta}_k)(\hat{\phi}_i - \phi_i) u_{2,i} \\ & - (\hat{\beta}_k - \beta_{k,0}) n^{-1}\sum_{i=1}^{n} \upsilon_2' \partial^1 \tilde{P}_{2,i}(\hat{\beta}_k) k_{2,i}(\hat{\phi}_i - \phi_i) \\ & - n^{-1}\sum_{i=1}^{n} \upsilon_2' \partial^1 \tilde{P}_{2,i}(\hat{\beta}_k)(\hat{\phi}_i - \phi_i)(g(\omega_{1,i}(\hat{\beta}_k), \hat{\beta}_k) - g(\omega_{1,i})). \end{aligned} \tag{SB.141}$$

By the Cauchy-Schwarz inequality, the triangle inequality, Assumption SB2(v) and Lemma SB1

$$\left| n^{-1}\sum_{i=1}^{n} \upsilon_2' \partial^1 \tilde{P}_{2,i}(\hat{\beta}_k) k_{2,i}(\hat{\phi}_i - \phi_i) \right| \leq \|\upsilon_2\| \, O_p(\xi_{1,m_2} m_1^{1/2} n^{-1/2}). \tag{SB.142}$$

Similarly we can show that

$$\begin{aligned} & \left| n^{-1}\sum_{i=1}^{n} \upsilon_2' \partial^1 \tilde{P}_{2,i}(\hat{\beta}_k)(\hat{\phi}_i - \phi_i)(g(\omega_{1,i}(\hat{\beta}_k), \hat{\beta}_k) - g(\omega_{1,i})) \right| \\ & \leq \|\upsilon_2\| \, |\hat{\beta}_k - \beta_{k,0}| O_p(\xi_{1,m_2} m_1^{1/2} n^{-1/2}). \end{aligned} \tag{SB.143}$$

By the Cauchy-Schwarz inequality, the triangle inequality, Assumption SB3(iv), Lemma SB27 and (SB.42),

$$\left| n^{-1}\sum_{i=1}^{n} \upsilon_2' \partial^1 \tilde{P}_{2,i}(\hat{\beta}_k)(\hat{\phi}_i - \phi_{m_1,i}) u_{2,i} \right| \leq \|\upsilon_2\| \, O_p((m_1^{1/2} + m_2) n^{-1/2}). \tag{SB.144}$$

Using similar arguments in the proof of Lemma SB27, we can show that

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

which together with the Cauchy-Schwarz inequality and Assumption SB3(iv) implies that

$$\left| n^{-1}\sum_{i=1}^{n} \upsilon_2' \partial^1 \tilde{P}_{2,i}(\hat{\beta}_k)(\phi_{m_1,i} - \phi_i) u_{2,i} \right| \leq \|\upsilon_2\| \, O_p((m_1^{1/2} + m_2) n^{-1/2}). \tag{SB.145}$$

Collecting the results in (SB.138), (SB.140), (SB.141), (SB.142), (SB.143), (SB.144) and (SB.145), we have

$$n^{-1}\sum_{i=1}^{n}(\hat{P}_{2,i}(\hat{\beta}_k)-\tilde{P}_{2,i}(\hat{\beta}_k))u_{2,i}(\hat{\beta}_k)=(\hat{\beta}_k-\beta_{k,0})O_p(\xi_{1,m_2}m_1^{1/2}n^{-1/2})+O_p((m_1^{1/2}+m_2)n^{-1/2}). \quad \text{(SB.146)}$$

The claim of the lemma follows from (SB.137) and (SB.146). *Q.E.D.*

Lemma SB12 *Under Assumptions SB1, SB2 and SB3, we have*

$$n^{-1}\sum_{i=1}^{n}\left|\hat{g}_i(\hat{\beta}_k)-g(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k)\right|^2=(\hat{\beta}_k-\beta_{k,0})^2O_p(\xi_{1,m_2}^2m_1n^{-1})+O_p\left((m_1+m_2^2)n^{-1}\right).$$

Proof of Lemma SB12. For any $\beta_k\in\Theta_k$ we deduce by the Cauchy-Schwarz inequality, Assumption SB2(iii) and (SB.49) that

$$\begin{aligned}
&n^{-1}\sum_{i=1}^{n}\left|\tilde{P}_{2,i}(\hat{\beta}_k)'\hat{\beta}_g(\hat{\beta}_k)-g(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k)\right|^2\\
&\leq 2n^{-1}\sum_{i=1}^{n}\left|\tilde{P}_{2,i}(\hat{\beta}_k)'\hat{\beta}_g(\hat{\beta}_k)-g_{m_2}(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k)\right|^2\\
&\qquad+2n^{-1}\sum_{i=1}^{n}\left|g_{m_2}(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k)-g(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k)\right|^2\\
&\leq C||\hat{\beta}_g(\hat{\beta}_k)-\tilde{\beta}_{g,m_2}(\hat{\beta}_k)||^2+Cm_2^{-2r_g} \qquad \text{(SB.147)}
\end{aligned}$$

wpa1, which together with Assumption SB2(vi) and Lemma SB11 implies that

$$\begin{aligned}
&n^{-1}\sum_{i=1}^{n}\left|\tilde{P}_{2,i}(\hat{\beta}_k)'\hat{\beta}_g(\hat{\beta}_k)-g(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k)\right|^2\\
&=(\hat{\beta}_k-\beta_{k,0})^2O_p(\xi_{1,m_2}^2m_1n^{-1})+O_p\left((m_1+m_2^2)n^{-1}\right). \qquad \text{(SB.148)}
\end{aligned}$$

By the mean value expansion, the Cauchy-Schwarz inequality, Assumptions SB2(v) and SB3(iv), Lemma

SB1, Lemma SB11 and (SB.51)

$$
\begin{aligned}
&n^{-1}\sum_{i=1}^{n}\left|(\hat{P}_{2,i}(\hat{\beta}_k)-\tilde{P}_{2,i}(\hat{\beta}_k))'(\hat{\beta}_g(\hat{\beta}_k)-\tilde{\beta}_{g,m_2}(\hat{\beta}_k))\right|^2 \\
&= n^{-1}\sum_{i=1}^{n}(\hat{\phi}_i-\phi_i)^2\left(\partial^1\tilde{P}_2(\tilde{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k)'(\hat{\beta}_g(\hat{\beta}_k)-\tilde{\beta}_{g,m_2}(\hat{\beta}_k))\right)^2 \\
&\leq \xi_{1,m_2}^2||\hat{\beta}_g(\hat{\beta}_k)-\tilde{\beta}_{g,m_2}(\hat{\beta}_k)||^2 n^{-1}\sum_{i=1}^{n}(\hat{\phi}_i-\phi_i)^2 \\
&= (\hat{\beta}_k-\beta_{k,0})^2 O_p(\xi_{1,m_2}^2 m_1 n^{-1}) + O_p\left((m_1+m_2^2)n^{-1}\right). \qquad \text{(SB.149)}
\end{aligned}
$$

By Assumptions SB2(ii, iii, vi), and Lemma SB1, we have

$$
\begin{aligned}
&n^{-1}\sum_{i=1}^{n}\left|(\hat{P}_{2,i}(\hat{\beta}_k)-\tilde{P}_{2,i}(\hat{\beta}_k))'\tilde{\beta}_{g,m_2}(\hat{\beta}_k)\right|^2 \\
&\leq 2n^{-1}\sum_{i=1}^{n}\left|(g_1(\hat{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k)-g_1(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k))\right|^2 + O_p(n^{-1}) \\
&\leq Cn^{-1}\sum_{i=1}^{n}(\hat{\phi}_i-\phi_i)^2 + O_p(n^{-1}) = O_p(m_1 n^{-1}) \qquad \text{(SB.150)}
\end{aligned}
$$

which together with (SB.149) implies that

$$
\begin{aligned}
&n^{-1}\sum_{i=1}^{n}\left|(\hat{P}_{2,i}(\hat{\beta}_k)-\tilde{P}_{2,i}(\hat{\beta}_k))'\hat{\beta}_g(\hat{\beta}_k)\right|^2 \\
&= (\hat{\beta}_k-\beta_{k,0})^2 O_p(\xi_{1,m_2}^2 m_1 n^{-1}) + O_p\left((m_1+m_2^2)n^{-1}\right). \qquad \text{(SB.151)}
\end{aligned}
$$

The claim of the lemma follows from (SB.148) and (SB.151). *Q.E.D.*

Lemma SB13 *Under Assumptions SB1, SB2 and SB3, we have*

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

Proof of Lemma SB13. Since $\hat{g}_{1,i}(\hat{\beta}_k)=\partial^1\hat{P}_{2,i}(\hat{\beta}_k)'\hat{\beta}_g(\hat{\beta}_k)$, we can use similar arguments in showing

(SB.148) to get

$$n^{-1}\sum_{i=1}^{n}\left|\partial^1 \tilde{P}_{2,i}(\hat{\beta}_k)'\hat{\beta}_g(\hat{\beta}_k) - g_1(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k)\right|^2$$
$$= (\hat{\beta}_k - \beta_{k,0})^2 O_p(\xi_{1,m_2}^4 m_1 n^{-1}) + O_p\left(\xi_{1,m_2}^2 (m_1 + m_2^2) n^{-1}\right). \quad \text{(SB.152)}$$

Using similar arguments in showing (SB.149) and (SB.150), we can show that

$$n^{-1}\sum_{i=1}^{n}\left|(\partial^1 \hat{P}_{2,i}(\hat{\beta}_k) - \partial^1 \tilde{P}_{2,i}(\hat{\beta}_k))'(\hat{\beta}_g(\hat{\beta}_k) - \tilde{\beta}_{g,m_2}(\hat{\beta}_k))\right|^2$$
$$\leq (\hat{\beta}_k - \beta_{k,0})^2 O_p(\xi_{2,m_2}^2 \xi_{1,m_2}^2 m_1^2 n^{-2}) + O_p\left(\xi_{2,m_2}^2 m_1 (m_1 + m_2^2) n^{-2}\right) \quad \text{(SB.153)}$$

and

$$n^{-1}\sum_{i=1}^{n}\left|(\partial^1 \hat{P}_{2,i}(\hat{\beta}_k) - \partial^1 \tilde{P}_{2,i}(\hat{\beta}_k))'\tilde{\beta}_{g,m_2}(\hat{\beta}_k)\right|^2 = O_p(m_1 n^{-1}), \quad \text{(SB.154)}$$

which implies that

$$n^{-1}\sum_{i=1}^{n}\left|(\partial^1 \hat{P}_{2,i}(\hat{\beta}_k) - \partial^1 \tilde{P}_{2,i}(\hat{\beta}_k))'\hat{\beta}_g(\hat{\beta}_k)\right|^2$$
$$\leq (\hat{\beta}_k - \beta_{k,0})^2 O_p(\xi_{2,m_2}^2 \xi_{1,m_2}^2 m_1^2 n^{-2}) + O_p\left(\xi_{2,m_2}^2 (m_1 + m_2^2) m_1 n^{-2}\right). \quad \text{(SB.155)}$$

The claim of the lemma follows from Assumption SB3(iv), (SB.152) and (SB.155). *Q.E.D.*

Lemma SB14 *Under Assumptions SB1, SB2 and SB3, we have*

$$n^{-1}\sum_{i=1}^{n} k_{2,i}(k_{2,i} - k_{1,i}\hat{g}_{1,i}(\hat{\beta}_k)) = \mathbb{E}[k_{2,i}(k_{2,i} - k_{1,i}g_1(\omega_{1,i}))] + o_p(1) \quad \text{(SB.156)}$$

and

$$n^{-1}\sum_{i=1}^{n} l_{2,i}k_{1,i}(\hat{g}_{1,i}(\hat{\beta}_k) - \hat{g}_{1,i}(\beta_{k,0}) = o_p(1). \quad \text{(SB.157)}$$

Proof of Lemma SB14. By the Cauchy-Schwarz inequality, Assumptions SB2(ii, vi) and SB3(iv), Lemma SB13 and the consistency of $\hat{\beta}_k$, we have

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

where the third equality is by the Markov inequality. This proves the claim in (SB.156). Similarly, by Assumptions SB2(ii, vi) and SB3(iv), Lemma SB13 and the consistency of $\hat{\beta}_k$, we have

$$\begin{aligned} & n^{-1}\sum_{i=1}^{n}\left|\hat{g}_{1,i}(\hat{\beta}_k)-\hat{g}_{1,i}(\beta_{k,0})\right|^2 \\ & \leq 2n^{-1}\sum_{i=1}^{n}\left|g_1(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k)-g_1(\omega_{1,i}(\beta_{k,0});\beta_{k,0})\right|^2+o_p(1) \\ & \leq C(\hat{\beta}_k-\beta_{k,0})^2+o_p(1)=o_p(1). \end{aligned} \tag{SB.158}$$

By the Markov inequality and Assumption SB2(i), $n^{-1}\sum_{i=1}^{n} l_{2,i}^2 k_{1,i}^2 = O_p(1)$ which together with (SB.158) proves the claim in (SB.157). *Q.E.D.*

Lemma SB15 *Let $a_{2,i}=a_2(\omega_{1,i})$. Then under Assumptions SB1, SB2 and SB3, we have*

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

Proof of Lemma SB15. First note that

$$\begin{aligned} & n^{-1}\sum_{i=1}^{n}(\hat{g}_i(\hat{\beta}_k)-\hat{g}_i(\beta_{k,0}))(k_{2,i}-k_{1,i}\hat{g}_{1,i}(\beta_{k,0})) \\ & = -n^{-1}\sum_{i=1}^{n} k_{1,i}(\hat{g}_i(\hat{\beta}_k)-\hat{g}_i(\beta_{k,0}))(\hat{g}_{1,i}(\beta_{k,0})-g_{1,i}) \\ & + n^{-1}\sum_{i=1}^{n}(\hat{g}_i(\hat{\beta}_k)-g(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k)-\hat{g}_i(\beta_{k,0})+g(\omega_{1,i}))(k_{2,i}-k_{1,i}g_{1,i}) \\ & + n^{-1}\sum_{i=1}^{n}(g(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k)-g(\omega_{1,i}))(k_{2,i}-k_{1,i}g_{1,i}). \end{aligned} \tag{SB.159}$$

By the Cauchy-Schwarz inequality, Assumption SB3(iv), Lemma SB12 and (SB.88),

$$n^{-1}\sum_{i=1}^{n}k_{1,i}(\hat{g}_i(\hat{\beta}_k)-\hat{g}_i(\beta_{k,0}))(\hat{g}_{1,i}(\beta_{k,0})-g_{1,i})=(\hat{\beta}_k-\beta_{k,0})o_p(1)+o_p(n^{-1/2}). \tag{SB.160}$$

Similarly, we can use Lemma SB12 to get

$$\begin{aligned}&n^{-1}\sum_{i=1}^{n}(\hat{g}_i(\hat{\beta}_k)-g(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k)-\hat{g}_i(\beta_{k,0})+g(\omega_{1,i}))(k_{2,i}-k_{1,i}g_{1,i})\\&=(\hat{\beta}_k-\beta_{k,0})o_p(1)+O_p((m_2+m_1^{1/2})n^{-1/2}).\end{aligned} \tag{SB.161}$$

Moreover, by Assumptions SB2(ii) and the consistency of $\hat{\beta}_k$

$$\begin{aligned}&n^{-1}\sum_{i=1}^{n}(g(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k)-g(\omega_{1,i}(\beta_{k,0});\beta_{k,0}))(k_{2,i}-k_{1,i}g_{1,i})\\&=(\hat{\beta}_k-\beta_{k,0})n^{-1}\sum_{i=1}^{n}\frac{\partial g(\omega_{1,i}(\beta_{k,0});\beta_{k,0})}{\partial\beta_k}(k_{2,i}-k_{1,i}g_{1,i})+(\hat{\beta}_k-\beta_{k,0})o_p(1).\end{aligned} \tag{SB.162}$$

Since

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

by Assumptions SB1(i) and SB2(ii), and the Markov inequality,

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

which together with (SB.162) implies that

$$\begin{aligned}&n^{-1}\sum_{i=1}^{n}(g(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k)-g(\omega_{1,i}(\beta_{k,0});\beta_{k,0}))(k_{2,i}-k_{1,i}g_{1,i})\\&=-(\hat{\beta}_k-\beta_{k,0})\left(\mathbb{E}[(a_{2\,i}+v_{1,i}g_{1,i})(k_{2,i}-k_{1,i}g_{1,i})]+o_p(1)\right)+o_p(n^{-1/2}).\end{aligned} \tag{SB.163}$$

The claim of the lemma follows from (SB.159), (SB.160), (SB.161) and (SB.163). *Q.E.D.*

Lemma SB16 *Under Assumptions SB1, SB2 and SB3, we have*

$$\hat{\beta}_k - \beta_{k,0} = O_p((m_1^{1/2} + m_2)n^{-1/2}). \tag{SB.164}$$

Proof of Lemma SB16. Using Assumption SB1(iii), Lemma SB14 and Lemma SB15, we can use the decomposition in (SA.25) to deduce that

$$\begin{aligned} & n^{-1}\sum_{i=1}^{n}(\hat{J}_i(\hat{\beta}_k) - \hat{J}_i(\beta_{k,0})) \\ & = -(\hat{\beta}_k - \beta_{k,0})\left(\mathbb{E}[(v_{2,i} - v_{1,i}g_{1,i})^2] + o_p(1)\right) \\ & - n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}(\hat{g}_{1,i}(\hat{\beta}_k) - \hat{g}_{1,i}(\beta_{k,0})) + O_p((m_1^{1/2} + m_2)n^{-1/2}). \end{aligned} \tag{SB.165}$$

In view of the first order condition of $\hat{\beta}_k$, (SA.5), (SA.28), (SB.165) and Lemma SA2, the claim of the lemma follows if one can show that

$$n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}(\hat{g}_{1,i}(\hat{\beta}_k) - \hat{g}_{1,i}(\beta_{k,0})) = (\hat{\beta}_k - \beta_{k,0})o_p(1) + O_p((m_1^{1/2} + m_2)n^{-1/2}). \tag{SB.166}$$

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

$$\begin{aligned} n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}\hat{g}_{1,i}(\hat{\beta}_k) &= n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}\partial^1\hat{P}_{2,i}(\hat{\beta}_k)'\hat{\beta}_g(\hat{\beta}_k) \\ &= n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}\partial^1\tilde{P}_{2,i}(\hat{\beta}_k)'\hat{\beta}_g(\hat{\beta}_k) \\ &+ n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}(\hat{\phi}_i - \phi_i)\partial^2\tilde{P}_2(\tilde{\omega}_{1,i};\hat{\beta}_k)'\hat{\beta}_g(\hat{\beta}_k), \end{aligned} \tag{SB.167}$$

where $\tilde{\omega}_{1,i}$ lies between $\hat{\omega}_{1,i}(\hat{\beta}_k)$ and $\omega_{1,i}(\hat{\beta}_k)$. By Assumption SB1(i), (SB.75) and the Markov inequality,

$$n^{-1}\sum_{i=1}^{n} u_{2,i}^2 k_{1,i}^2 = O_p(1). \tag{SB.168}$$

By the triangle inequality

$$\begin{aligned}
&\left|n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(\partial^1\tilde{P}_{2,i}(\hat{\beta}_k)'\hat{\beta}_g(\hat{\beta}_k)-g_1(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k))\right| \\
&\leq\left|n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(g_1(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k)-\partial^1\tilde{P}_{2,i}(\hat{\beta}_k)'\tilde{\beta}_{g,m_2}(\hat{\beta}_k))\right| \\
&+\left|n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}\partial^1\tilde{P}_{2,i}(\hat{\beta}_k)'(\hat{\beta}_g(\hat{\beta}_k)-\tilde{\beta}_{g,m_2}(\hat{\beta}_k))\right|. \qquad \text{(SB.169)}
\end{aligned}$$

By Assumptions SB2(iii, vi) and (SB.168)

$$n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(g_1(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k)-\partial^1\tilde{P}_{2,i}(\hat{\beta}_k)'\tilde{\beta}_{g,m_2}(\hat{\beta}_k))=O_p((m_1^{1/2}+m_2)n^{-1/2}). \qquad \text{(SB.170)}$$

By the Cauchy-Schwarz inequality, Assumption SB2(v), Lemma SB11 and Lemma SB26

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

which together with (SB.169) and (SB.170) implies that

$$\begin{aligned}
&n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(\partial^1\tilde{P}_{2,i}(\hat{\beta}_k)'\hat{\beta}_g(\hat{\beta}_k)-g_1(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k)) \\
&=|\hat{\beta}_k-\beta_{k,0}|o_p(1)+O_p((m_1^{1/2}+m_2)n^{-1/2}). \qquad \text{(SB.171)}
\end{aligned}$$

By the Cauchy-Schwarz inequality and the triangle inequality, Assumptions SB2(ii, iii, v, vi) and SB3(iv), Lemma SB1, Lemma SB11 and (SB.168)

$$
\begin{aligned}
&\left|n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(\hat{\phi}_i-\phi_i)\partial^2\tilde{P}_2(\tilde{\omega}_{1,i};\hat{\beta}_k)'\hat{\beta}_g(\hat{\beta}_k)\right| \\
&\leq n^{-1}\sum_{i=1}^{n}\left|u_{2,i}k_{1,i}(\hat{\phi}_i-\phi_i)\partial^2\tilde{P}_2(\tilde{\omega}_{1,i};\hat{\beta}_k)'\tilde{\beta}_{g,m_2}(\hat{\beta}_k)\right| \\
&+n^{-1}\sum_{i=1}^{n}\left|u_{2,i}k_{1,i}(\hat{\phi}_i-\phi_i)\partial^2\tilde{P}_2(\tilde{\omega}_{1,i};\hat{\beta}_k)'(\hat{\beta}_g(\hat{\beta}_k)-\tilde{\beta}_{g,m_2}(\hat{\beta}_k))\right| \\
&\leq \left(\sup_{\omega\in\Omega_{C_\omega}}\left|\partial^2\tilde{P}_2(\omega;\hat{\beta}_k)'\tilde{\beta}_{g,m_2}(\hat{\beta}_k)\right|+\xi_{2,m_2}\left\|\hat{\beta}_g(\hat{\beta}_k)-\tilde{\beta}_{g,m_2}(\hat{\beta}_k)\right\|\right)n^{-1}\sum_{i=1}^{n}\left|u_{2,i}k_{1,i}(\hat{\phi}_i-\phi_i)\right| \\
&=|\hat{\beta}_k-\beta_{k,0}|o_p(1)+O_p((m_1^{1/2}+m_2)n^{-1/2}) \qquad \text{(SB.172)}
\end{aligned}
$$

which together with (SB.167), (SB.171) and (SB.172) proves

$$
n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(\hat{g}_{1,i}(\hat{\beta}_k)-g_1(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k))=(\hat{\beta}_k-\beta_{k,0})o_p(1)+O_p((m_1^{1/2}+m_2)n^{-1/2}). \qquad \text{(SB.173)}
$$

Similarly, we can show that

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

which together with (SB.173) implies that

$$
\begin{aligned}
&n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(\hat{g}_{1,i}(\hat{\beta}_k)-\hat{g}_{1,i}(\beta_{k,0})) \\
&=n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(g_1(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k)-g_1(\omega_{1,i}(\beta_{k,0});\beta_{k,0})) \\
&+(\hat{\beta}_k-\beta_{k,0})o_p(1)+O_p((m_1^{1/2}+m_2)n^{-1/2}). \qquad \text{(SB.174)}
\end{aligned}
$$

By Assumption SB2(ii), the Markov inequality and the consistency of $\hat{\beta}_k$,

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

which together with (SB.174) proves (SB.166). *Q.E.D.*

Lemma SB17 *Under Assumptions SB1, SB2 and SB3, we have*

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

Proof of Lemma SB17. We define $\hat{\mathbf{P}}_2^*(\beta_k) = (\hat{P}_{2,1}^*(\beta_k), \ldots, \hat{P}_{2,n}^*(\beta_k))'$ where $\hat{P}_{2,i}^*(\beta_k) = B(\beta_{k,0})P_2(\hat{\omega}_{1,i}(\beta_k))$. Then we can write

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

and therefore,

$$\begin{aligned}
&(B(\beta_{k,0})')^{-1}B(\hat{\beta}_k)'\hat{\beta}_g(\hat{\beta}_k) - \hat{\beta}_g(\beta_{k,0}) \\
&= \left[(\hat{\mathbf{P}}_2^*(\hat{\beta}_k)'\hat{\mathbf{P}}_2^*(\hat{\beta}_k))^{-1} - (\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}\right]\hat{\mathbf{P}}_2^*(\hat{\beta}_k)'\hat{\mathbf{Y}}_2^*(\hat{\beta}_k) \\
&+ (\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}(\hat{\mathbf{P}}_2^*(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\beta_{k,0}))'\hat{\mathbf{Y}}_2^*(\beta_k) \\
&+ (\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'(\hat{\mathbf{Y}}_2^*(\hat{\beta}_k) - \hat{\mathbf{Y}}_2^*(\beta_{k,0})). \qquad \text{(SB.175)}
\end{aligned}$$

By (SB.53), Assumption SB1(i) and the Markov inequality,

$$\begin{aligned}
&\left\|(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'\mathbf{K}_2\right\|^2 \\
&\leq (\lambda_{\min}(n^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0})))^{-1}n^{-1}\sum_{i=1}^n k_{2,i}^2 = O_p(1). \qquad \text{(SB.176)}
\end{aligned}$$

Since $\hat{\mathbf{Y}}_2^*(\hat{\beta}_k) - \hat{\mathbf{Y}}_2^*(\beta_{k,0}) = -(\hat{\beta}_k - \beta_{k,0})\mathbf{K}_2$, by (SB.176) we get

$$(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'(\hat{\mathbf{Y}}_2^*(\hat{\beta}_k) - \hat{\mathbf{Y}}_2^*(\beta_{k,0})) = (\hat{\beta}_k - \beta_{k,0})O_p(1). \qquad \text{(SB.177)}$$

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

$$\upsilon_2'(\tilde{P}_2(\hat{\omega}_{1,i}(\hat{\beta}_k);\beta_{k,0}) - \tilde{P}_2(\hat{\omega}_{1,i}(\beta_{k,0});\beta_{k,0})) = -\upsilon_2'\partial^1\tilde{P}_2(\hat{\omega}_{1,i}(\tilde{\beta}_k);\beta_{k,0})k_{1,i}(\hat{\beta}_k - \beta_{k,0}), \qquad \text{(SB.178)}$$

where $\tilde{\beta}_k$ lies between $\hat{\beta}_k$ and $\beta_{k,0}$. By Assumption SB3(iv) and Lemma SB16, $\hat{\omega}_{1,i}(\tilde{\beta}_k) \in \Omega_{\varepsilon_n}(\beta_{k,0})$ for any $i = 1, \ldots, n$ wpa1. By the Cauchy-Schwarz inequality and (SB.178)

$$\left|\upsilon_2'(\hat{P}_{2,i}^*(\hat{\beta}_k) - \hat{P}_{2,i}(\beta_{k,0}))\right| \leq \|\upsilon_2\|\,\xi_{1,m_2}\left|k_{1,i}(\hat{\beta}_k - \beta_{k,0})\right| \qquad \text{(SB.179)}$$

wpa1. Therefore we have wpa1,

$$
\begin{aligned}
&\upsilon_2'(\hat{\mathbf{P}}_2^*(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\beta_{k,0}))'(\hat{\mathbf{P}}_2^*(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\beta_{k,0}))\upsilon_2 \\
&= \sum_{i=1}^n (\upsilon_2'(\hat{P}_{2,i}^*(\hat{\beta}_k) - \hat{P}_{2,i}(\beta_{k,0}))^2 \leq \|\upsilon_2\|^2 \, \xi_{1,m_2}^2 (\hat{\beta}_k - \beta_{k,0})^2 \sum_{i=1}^n k_{1,i}^2 \qquad \text{(SB.180)}
\end{aligned}
$$

which implies that

$$
||\hat{\mathbf{P}}_2^*(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\beta_{k,0})||_S = |\hat{\beta}_k - \beta_{k,0}| O_p(\xi_{1,m_2} n^{1/2}). \qquad \text{(SB.181)}
$$

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, Assumption SB2(i) and the compactness of Θ_k implies that

$$
\sup_{\beta_k \in \Theta_k} n^{-1} \sum_{i=1}^n (y_{2,i}^*(\beta_k))^2 = O_p(1). \qquad \text{(SB.182)}
$$

By the Cauchy-Schwarz inequality,(SB.53), (SB.181) and (SB.182),

$$
\begin{aligned}
&\left\| (\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1} (\hat{\mathbf{P}}_2^*(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\beta_{k,0}))' \hat{\mathbf{Y}}_2^*(\beta_k) \right\| \\
&\leq (\lambda_{\min}(n^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0})))^{-1} n^{-1} ||\hat{\mathbf{P}}_2^*(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\beta_{k,0})||_S \left\| \hat{\mathbf{Y}}_2^*(\hat{\beta}_k) \right\| \\
&= |\hat{\beta}_k - \beta_{k,0}| O_p(\xi_{1,m_2}). \qquad \text{(SB.183)}
\end{aligned}
$$

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

$$
\begin{aligned}
&[(\hat{\mathbf{P}}_2^*(\hat{\beta}_k)'\hat{\mathbf{P}}_2^*(\hat{\beta}_k))^{-1} - (\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}] \hat{\mathbf{P}}_2^*(\hat{\beta}_k)' \hat{\mathbf{Y}}_2^*(\hat{\beta}_k) \\
&= (\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1} (\hat{\mathbf{P}}_2(\beta_{k,0}) - \hat{\mathbf{P}}_2^*(\hat{\beta}_k))' \hat{\mathbf{P}}_2(\hat{\beta}_k) \hat{\beta}_g(\hat{\beta}_k) \qquad \text{(SB.184)} \\
&+ (\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1} \hat{\mathbf{P}}_2(\beta_{k,0})' (\hat{\mathbf{P}}_2(\beta_{k,0}) - \hat{\mathbf{P}}_2^*(\hat{\beta}_k)) (B(\beta_{k,0})')^{-1} B(\hat{\beta}_k)' \hat{\beta}_g(\hat{\beta}_k).
\end{aligned}
$$

By the Cauchy-Schwarz inequality, Assumption SB2(vi), (SB.53), (SB.65) and (SB.181),

$$\begin{aligned}
&\left\|(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}(\hat{\mathbf{P}}_2(\beta_{k,0})-\hat{\mathbf{P}}_2^*(\hat{\beta}_k))'\hat{\mathbf{P}}_2(\hat{\beta}_k)\hat{\beta}_g(\hat{\beta}_k)\right\| \\
&\leq (\lambda_{\min}(n^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0})))^{-1}n^{-1}||\hat{\mathbf{P}}_2^*(\hat{\beta}_k)-\hat{\mathbf{P}}_2(\beta_{k,0})||_S\left\|\hat{\mathbf{P}}_2(\hat{\beta}_k)\hat{\beta}_g(\hat{\beta}_k)\right\| \\
&= |\hat{\beta}_k-\beta_{k,0}|O_p(\xi_{1,m_2}). \qquad \text{(SB.185)}
\end{aligned}$$

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

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

which together with Assumptions SB2(ii, iii), Lemma SB11 and Lemma SB16 implies that

$$\left\|(\hat{\mathbf{P}}_2(\beta_{k,0})-\hat{\mathbf{P}}_2^*(\hat{\beta}_k))(B(\beta_{k,0})')^{-1}B(\hat{\beta}_k)'\hat{\beta}_g(\hat{\beta}_k)\right\| = |\hat{\beta}_k-\beta_{k,0}|O_p(n^{1/2}). \qquad \text{(SB.186)}$$

By (SB.53) and (SB.186)

$$(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'(\hat{\mathbf{P}}_2(\beta_{k,0})-\hat{\mathbf{P}}_2^*(\hat{\beta}_k))(B(\beta_{k,0})')^{-1}B(\hat{\beta}_k)'\hat{\beta}_g(\hat{\beta}_k) = (\hat{\beta}_k-\beta_{k,0})O_p(1). \qquad \text{(SB.187)}$$

Collecting the results in (SB.184), (SB.185) and (SB.187) we get

$$[(\hat{\mathbf{P}}_2^*(\hat{\beta}_k)\hat{\mathbf{P}}_2^*(\hat{\beta}_k)')^{-1}-(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}]\hat{\mathbf{P}}_2^*(\hat{\beta}_k)'\hat{\mathbf{Y}}_2^*(\hat{\beta}_k) = (\hat{\beta}_k-\beta_{k,0})O_p(\xi_{1,m_2}). \qquad \text{(SB.188)}$$

The claim of the lemma follows from (SB.175), (SB.177), (SB.183) and (SB.188). *Q.E.D.*

Lemma SB18 *Under Assumptions SB1, SB2 and SB3, we have*

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

PROOF OF LEMMA SB18. By the definition of $\hat{g}_{1,i}(\hat{\beta}_k)$, we can write

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

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

$$\begin{aligned}
& n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}(\hat{g}_{1,i}(\hat{\beta}_k) - \hat{g}_{1,i}(\beta_{k,0})) \\
&= n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}(\partial^1 \tilde{P}_2(\hat{\omega}_{1,i}(\hat{\beta}_k); \beta_{k,0}) - \partial^1 \tilde{P}_2(\omega_{1,i}(\beta_{k,0}); \beta_{k,0}))'\upsilon_{2,*} \\
&\quad - n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}(\partial^1 \tilde{P}_2(\hat{\omega}_{1,i}(\beta_{k,0}); \beta_{k,0}) - \partial^1 \tilde{P}_2(\omega_{1,i}(\beta_{k,0}); \beta_{k,0}))'\hat{\beta}_g(\beta_{k,0}) \\
&\quad + n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}\partial^1 \tilde{P}_2(\omega_{1,i}(\beta_{k,0}); \beta_{k,0})'(\upsilon_{2,*} - \hat{\beta}_g(\beta_{k,0})). \qquad \text{(SB.189)}
\end{aligned}$$

Since $\upsilon_{2,*} = (B(\beta_{k,0})')^{-1}B(\hat{\beta}_k)'\hat{\beta}_g(\hat{\beta}_k)$, by Assumption SB3(iv), Lemma SB16, Lemma SB17 and (SB.65)

$$\|\upsilon_{2,*}\| = O_p(1). \qquad \text{(SB.190)}$$

By the second order expansion,

$$\begin{aligned}
& n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}(\partial^1 \tilde{P}_2(\hat{\omega}_{1,i}(\hat{\beta}_k); \beta_{k,0}) - \partial^1 \tilde{P}_2(\omega_{1,i}(\beta_{k,0}); \beta_{k,0}))'\upsilon_{2,*} \\
&= n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}(\hat{\omega}_{1,i}(\hat{\beta}_k) - \omega_{1,i})\partial^2 \tilde{P}_{2,i}(\beta_{k,0})'\upsilon_{2,*} \\
&\quad + n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}(\hat{\omega}_{1,i}(\hat{\beta}_k) - \omega_{1,i})^2\partial^3 \tilde{P}_2(\tilde{\omega}_{1,i}; \beta_{k,0})'\upsilon_{2,*}, \qquad \text{(SB.191)}
\end{aligned}$$

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

$$\begin{aligned}
& n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}(\hat{\omega}_{1,i}(\hat{\beta}_k) - \omega_{1,i})\partial^2 \tilde{P}_2(\omega_{1,i}(\beta_{k,0}); \beta_{k,0})'\upsilon_{2,*} \\
&= n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}(\hat{\phi}_i - \phi_i)\partial^2 \tilde{P}_{2,i}(\beta_{k,0})'\upsilon_{2,*} \\
&\quad - (\hat{\beta}_k - \beta_{k,0})n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}^2\partial^2 \tilde{P}_{2,i}(\beta_{k,0})'\upsilon_{2,*}. \qquad \text{(SB.192)}
\end{aligned}$$

By the Cauchy-Schwarz inequality, Assumptions SB1(i) and SB2(v), (SB.75) and (SB.190),

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

which together with Assumption SB3(iv) implies that

$$(\hat{\beta}_k - \beta_{k,0}) n^{-1} \sum_{i=1}^{n} u_{2,i} k_{1,i}^2 \partial^2 \tilde{P}_{2,i}(\beta_{k,0})' \upsilon_{2,*} = (\hat{\beta}_k - \beta_{k,0}) o_p(1). \tag{SB.193}$$

By the Cauchy-Schwarz inequality, Assumptions SB1(i, iv, vi), SB2(v) and SB3(iv), (SB.75) and (SB.190),

$$n^{-1} \sum_{i=1}^{n} u_{2,i} k_{1,i} (\phi_{m_1,i} - \phi_i) \partial^2 \tilde{P}_2(\omega_{1,i}(\beta_{k,0}); \beta_{k,0})' \upsilon_{2,*} = O_p(\xi_{2,m_2} n^{-1}) = o_p(n^{-1/2}). \tag{SB.194}$$

By the Cauchy-Schwarz inequality, Assumptions SB1(i, iv, vi), SB2(v) and SB3(iv), (SB.42), (SB.75) and (SB.190),

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

which together with (SB.192), (SB.193) and (SB.194) implies that

$$n^{-1} \sum_{i=1}^{n} u_{2,i} k_{1,i} (\hat{\omega}_{1,i}(\hat{\beta}_k) - \omega_{1,i}) \partial^2 \tilde{P}_2(\omega_{1,i}(\beta_{k,0}); \beta_{k,0})' \upsilon_{2,*} = (\hat{\beta}_k - \beta_{k,0}) o_p(1) + o_p(n^{-1/2}). \tag{SB.195}$$

Using the similar arguments in showing (SB.79), we can show that

$$n^{-1} \sum_{i=1}^{n} \left| u_{2,i} k_{1,i} (\hat{\phi}_i - \phi_i)^2 \right| = O_p(m_1 n^{-1}). \tag{SB.196}$$

Moreover by the Markov inequality, Assumption SB1(i) and (SB.75)

$$n^{-1} \sum_{i=1}^{n} (\left| u_{2,i} k_{1,i}^2 \right| + |u_{2,i} k_{1,i}|) = O_p(1). \tag{SB.197}$$

By the Cauchy-Schwarz inequality, Assumption SB3(iv), Lemma SB16, (SB.196) and (SB.197)

$$
\begin{aligned}
& n^{-1}\sum_{i=1}^{n}|u_{2,i}k_{1,i}|\,(\hat{\omega}_{1,i}(\hat{\beta}_k)-\omega_{1,i})^2 \\
& \leq 2n^{-1}\sum_{i=1}^{n}\left|u_{2,i}k_{1,i}(\hat{\phi}_i-\phi_i)^2\right| + 2(\hat{\beta}_k-\beta_{k,0})^2 n^{-1}\sum_{i=1}^{n}\left|u_{2,i}k_{1,i}^2\right| = o_p(n^{-1/2}).
\end{aligned}
\tag{SB.198}
$$

By the Cauchy-Schwarz inequality, Assumptions SB2(ii, iii, v) and SB3(iv), Lemma SB11 and (SB.198)

$$
\begin{aligned}
& \left|n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(\hat{\omega}_{1,i}(\hat{\beta}_k)-\omega_{1,i})^2\partial^3\tilde{P}_2(\tilde{\omega}_{1,i};\beta_{k,0})'\upsilon_{2,*}\right| \\
& \leq \left|n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(\hat{\omega}_{1,i}(\hat{\beta}_k)-\omega_{1,i})^2\partial^3\tilde{P}_2(\tilde{\omega}_{1,i};\hat{\beta}_k)'\tilde{\beta}_{g,m_2}(\hat{\beta}_k)\right| \\
& + \left|n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(\hat{\omega}_{1,i}(\hat{\beta}_k)-\omega_{1,i})^2\partial^3\tilde{P}_2(\tilde{\omega}_{1,i};\hat{\beta}_k)'(\hat{\beta}_g(\hat{\beta}_k)-\tilde{\beta}_{g,m_2}(\hat{\beta}_k))\right| \\
& = (\hat{\beta}_k-\beta_{k,0})o_p(1)+o_p(n^{-1/2})
\end{aligned}
\tag{SB.199}
$$

which together with (SB.191) and (SB.195) implies that

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

Using similar arguments in proving (SB.200), we can show that

$$
n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}(\partial^1\tilde{P}_2(\hat{\omega}_{1,i}(\beta_{k,0});\beta_{k,0})-\partial^1\tilde{P}_{2,i}(\beta_{k,0}))'\hat{\beta}_g(\beta_{k,0}) = o_p(n^{-1/2}). \tag{SB.201}
$$

By the Cauchy-Schwarz inequality, Assumptions SB1(i), SB2(v) and SB3(iv), Lemma SB17 and (SB.75)

$$
\begin{aligned}
& \left\|n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}\partial^1\tilde{P}_{2,i}(\beta_{k,0})'(\upsilon_{2,*}-\hat{\beta}_g(\beta_{k,0}))\right\| \\
& \leq \left\|n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}\partial^1\tilde{P}_{2,i}(\beta_{k,0})\right\|\left\|\upsilon_{2,*}-\hat{\beta}_g(\beta_{k,0})\right\| = (\hat{\beta}_k-\beta_{k,0})o_p(1).
\end{aligned}
\tag{SB.202}
$$

The claim of the lemma follows from (SB.189), (SB.200), (SB.201) and (SB.202). *Q.E.D.*

Lemma SB19 *Let* $\hat{\mathbf{G}}_n = (\hat{g}(\hat{\omega}_{1,1}(\hat{\beta}_k);\hat{\beta}_k),\ldots,\hat{g}(\hat{\omega}_{1,n}(\hat{\beta}_k);\hat{\beta}_k))'$, $\mathbf{G}_n = (g(\omega_{1,1}),\ldots,g(\omega_{1,n}))'$ *and* $\mathbf{U}_2 = (u_{2,1},\ldots,u_{2,n})'$. *Then under Assumptions SB1, SB2 and SB3, we have* *(i)* $n^{-1}\mathbf{U}_2'(\hat{\mathbf{P}}_2^*(\hat{\beta}_k)-\hat{\mathbf{P}}_2(\beta_{k,0}))\hat{\beta}_\varphi(\beta_{k,0})=(\hat{\beta}_k-\beta_{k,0})o_p(1)+o_p(n^{-1/2})$;

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

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

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

Proof of Lemma SB19. (i) By the first order expansion,

$$\begin{aligned}&n^{-1}\mathbf{U}_2'(\hat{\mathbf{P}}_2^*(\hat{\beta}_k)-\hat{\mathbf{P}}_2(\beta_{k,0}))(\hat{\beta}_\varphi(\beta_{k,0})-\tilde{\beta}_{\varphi,m_2}(\beta_{k,0}))\\&=-(\hat{\beta}_k-\beta_{k,0})n^{-1}\sum_{i=1}^n u_{2,i}k_{1,i}\partial^1\tilde{P}_2(\tilde{\omega}_{1,i};\beta_{k,0})'(\hat{\beta}_\varphi(\beta_{k,0})-\tilde{\beta}_{\varphi,m_2}(\beta_{k,0})),\end{aligned}\tag{SB.203}$$

where $\tilde{\omega}_{1,i}$ lies between $\hat{\omega}_{1,i}(\hat{\beta}_k)$ and $\hat{\omega}_{1,i}(\beta_{k,0})$. By Assumptions SB2(v, vi) and SB3(iv), (SB.112) and (SB.197),

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

which together with (SB.203) implies that

$$n^{-1}\mathbf{U}_2'(\hat{\mathbf{P}}_2^*(\hat{\beta}_k)-\hat{\mathbf{P}}_2(\beta_{k,0}))(\hat{\beta}_\varphi(\beta_{k,0})-\tilde{\beta}_{\varphi,m_2}(\beta_{k,0}))=(\hat{\beta}_k-\beta_{k,0})o_p(1).\tag{SB.204}$$

By Assumptions SB3(i, ii, iv), Lemma SB1, Lemma SB16, (SB.75) and (SB.197)

$$\begin{aligned}&n^{-1}\mathbf{U}_2'(\hat{\mathbf{P}}_2^*(\hat{\beta}_k)-\hat{\mathbf{P}}_2(\beta_{k,0}))\tilde{\beta}_{\varphi,m_2}(\beta_{k,0})\\&=n^{-1}\sum_{i=1}^n u_{2,i}k_{1,i}(\varphi(\hat{\omega}_{1,i}(\hat{\beta}_k))-\varphi(\hat{\omega}_{1,i}(\beta_{k,0})))+o_p(n^{-1/2})\\&=n^{-1}\sum_{i=1}^n u_{2,i}k_{1,i}\varphi_1(\hat{\omega}_{1,i}(\beta_{k,0}))(\hat{\beta}_k-\beta_{k,0})+o_p(n^{-1/2})\\&=n^{-1}\sum_{i=1}^n u_{2,i}k_{1,i}\varphi_1(\omega_{1,i})(\hat{\beta}_k-\beta_{k,0})+(\hat{\beta}_k-\beta_{k,0})o_p(1)+o_p(n^{-1/2}).\end{aligned}\tag{SB.205}$$

By the Markov inequality, Assumptions SB1(i) and SB3(i), (SB.75).

$$n^{-1}\sum_{i=1}^n u_{2,i}k_{1,i}\varphi_1(\omega_{1,i})=O_p(n^{-1/2})\tag{SB.206}$$

which together with (SB.205) implies that

$$n^{-1}\mathbf{U}_2'(\hat{\mathbf{P}}_2^*(\hat{\beta}_k)-\hat{\mathbf{P}}_2(\beta_{k,0}))\tilde{\beta}_{\varphi,m_2}(\beta_{k,0})=(\hat{\beta}_k-\beta_{k,0})o_p(1)+o_p(n^{-1/2}). \tag{SB.207}$$

The first claim of the lemma follows by (SB.204) and (SB.207).

(ii) Using the similar arguments in showing (SB.191), we get

$$n^{-1}\mathbf{L}_2'(\hat{\mathbf{P}}_2^*(\hat{\beta}_k)-\hat{\mathbf{P}}_2(\beta_{k,0}))(\hat{\beta}_{\varphi}(\beta_{k,0})-\tilde{\beta}_{\varphi,m_2}(\beta_{k,0}))=(\hat{\beta}_k-\beta_{k,0})o_p(1). \tag{SB.208}$$

By the mean value expansion, Assumptions SB3(i, ii, iv), Lemma SB1, the consistency of $\hat{\beta}_k$ and the Markov inequality

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

which together with (SB.208) finishes the proof.

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

(iv) By the first-order expansion,

$$\begin{aligned}
&n^{-1}(\hat{\mathbf{G}}_n-\mathbf{G}_n)'(\hat{\mathbf{P}}_2^*(\hat{\beta}_k)-\hat{\mathbf{P}}_2(\beta_{k,0}))(\hat{\beta}_{\varphi}(\beta_{k,0})-\tilde{\beta}_{\varphi,m_2}(\beta_{k,0}))\\
&=n^{-1}\sum_{i=1}^n(\hat{g}_i(\hat{\beta}_k)-g(\omega_{1,i}))(\tilde{P}_2(\hat{\omega}_{1,i}(\hat{\beta}_k);\beta_{k,0})-\hat{P}_{2,i}(\beta_{k,0}))'(\hat{\beta}_{\varphi}(\beta_{k,0})-\tilde{\beta}_{\varphi,m_2}(\beta_{k,0}))\\
&=-(\hat{\beta}_k-\beta_{k,0})n^{-1}\sum_{i=1}^n(\hat{g}_i(\hat{\beta}_k)-g(\omega_{1,i}))\partial^1\tilde{P}_2(\tilde{\omega}_{1,i};\beta_{k,0})'(\hat{\beta}_{\varphi}(\beta_{k,0})-\tilde{\beta}_{\varphi,m_2}(\beta_{k,0})),
\end{aligned} \tag{SB.209}$$

where $\tilde{\omega}_{1,i}$ lies between $\hat{\omega}_{1,i}(\hat{\beta}_k)$ and $\hat{\omega}_{1,i}(\beta_{k,0})$. By Assumption SB2(v), Lemma SB12 and (SB.112), we get

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

which together with Assumption SB3(iv) and (SB.209) implies that

$$n^{-1}(\hat{\mathbf{G}}_n - \mathbf{G}_n)'(\hat{\mathbf{P}}_2^*(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\beta_{k,0}))(\hat{\beta}_\varphi(\beta_{k,0}) - \tilde{\beta}_{\varphi,m_2}(\beta_{k,0})) = (\hat{\beta}_k - \beta_{k,0})o_p(1). \tag{SB.210}$$

Using Assumptions SB3(i, ii, iv) and Lemma SB12, we get

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

which together with (SB.210) proves the claim. *Q.E.D.*

Lemma SB20 *Under Assumptions SB1, SB2 and SB3, we have*

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

Proof of Lemma SB20. By the definition of $\hat{g}(\hat{\omega}_{1,i}(\beta_k);\beta_k)$, we can write

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

and therefore

$$\begin{aligned}
&\hat{g}(\hat{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k) - \hat{g}(\hat{\omega}_{1,i}(\beta_{k,0});\beta_{k,0}) \\
&= (\hat{P}_{2,i}^*(\hat{\beta}_k) - \hat{P}_{2,i}(\beta_{k,0}))'(\hat{\mathbf{P}}_2^*(\hat{\beta}_k)'\hat{\mathbf{P}}_2^*(\hat{\beta}_k))^{-1}\hat{\mathbf{P}}_2^*(\hat{\beta}_k)'\hat{\mathbf{Y}}_2^*(\hat{\beta}_k) \\
&\quad + \hat{P}_{2,i}(\beta_{k,0})'\left[(\hat{\mathbf{P}}_2^*(\hat{\beta}_k)'\hat{\mathbf{P}}_2^*(\hat{\beta}_k))^{-1} - (\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}\right]\hat{\mathbf{P}}_2^*(\hat{\beta}_k)'\hat{\mathbf{Y}}_2^*(\hat{\beta}_k) \\
&\quad + \hat{P}_{2,i}(\beta_{k,0})'(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}(\hat{\mathbf{P}}_2^*(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\beta_{k,0}))'\hat{\mathbf{Y}}_2^*(\beta_k) \\
&\quad + \hat{P}_{2,i}(\beta_{k,0})'(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'(\hat{\mathbf{Y}}_2^*(\hat{\beta}_k) - \mathbf{Y}_2^*(\beta_{k,0})).
\end{aligned} \tag{SB.211}$$

The proof is divided into 4 steps. The claim of the lemma follows from the results in (SB.160), (SB.212), (SB.225), (SB.235) and (SB.237).

Step 1. In this step, we show that

$$n^{-1}\sum_{i=1}^{n}\upsilon_{2,*}'(\hat{P}_{2,i}^{*}(\hat{\beta}_k)-\hat{P}_{2,i}(\beta_{k,0}))(k_{2,i}-k_{1,i}g_{1,i})$$
$$=-(\hat{\beta}_k-\beta_{k,0})\left(\mathbb{E}[k_{1,i}g_{1,i}(k_{2,i}-k_{1,i}g_{1,i})]+o_p(1)\right)+o_p(n^{-1/2}), \tag{SB.212}$$

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

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

$$\upsilon_2'\left(\hat{P}_{2,i}^{*}(\hat{\beta}_k)-\tilde{P}_{2,i}(\beta_{k,0})-\partial^1\tilde{P}_{2,i}(\beta_{k,0})(\hat{\omega}_{1,i}(\hat{\beta}_k)-\omega_{1,i}(\beta_{k,0}))\right)$$
$$=\upsilon_2'\partial^2\tilde{P}_2(\tilde{\omega}_{1,i};\beta_{k,0})(\hat{\omega}_{1,i}(\hat{\beta}_k)-\omega_{1,i}(\beta_{k,0}))^2, \tag{SB.213}$$

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

$$\left|n^{-1}\sum_{i=1}^{n}\upsilon_2'\begin{pmatrix}\hat{P}_{2,i}^{*}(\hat{\beta}_k)-\tilde{P}_{2,i}(\beta_{k,0})\\ -\partial^1\tilde{P}_{2,i}(\beta_{k,0})(\hat{\omega}_{1,i}(\hat{\beta}_k)-\omega_{1,i}(\beta_{k,0}))\end{pmatrix}(k_{2,i}-k_{1,i}g_{1,i})\right|$$
$$\leq C\max_{i\leq n}\left\|\upsilon_2'\partial^2\tilde{P}_2(\tilde{\omega}_{1,i};\beta_{k,0})\right\|\left(n^{-1}\sum_{i=1}^{n}(\hat{\phi}_i-\phi_i)^2+(\hat{\beta}_k-\beta_{k,0})^2n^{-1}\sum_{i=1}^{n}k_{1,i}^2\right). \tag{SB.214}$$

By the definition of $\upsilon_{2,*}$, we can write $\upsilon_{2,*}'\partial^2\tilde{P}_2(\tilde{\omega}_{1,i};\beta_{k,0})=\hat{\beta}_g(\hat{\beta}_k)'\partial^2\tilde{P}_2(\tilde{\omega}_{1,i};\hat{\beta}_k)$ where $\tilde{\omega}_{1,i}\in\Omega_{\varepsilon_n}(\hat{\beta}_k)$ for any $i\leq n$ wpa1. By the triangle inequality, Assumptions SB2(iii, v, vi), Lemma SB11 and Lemma SB16

$$\max_{i\leq n}\left\|\upsilon_{2,*}'\partial^2\tilde{P}_2(\tilde{\omega}_{1,i};\beta_{k,0})\right\|$$
$$\leq\max_{i\leq n}\left\|\tilde{\beta}_{g,m_2}(\hat{\beta}_k)'\partial^2\tilde{P}_2(\tilde{\omega}_{1,i};\hat{\beta}_k)\right\|+\max_{i\leq n}\left\|(\hat{\beta}_g(\hat{\beta}_k)-\tilde{\beta}_{g,m_2}(\hat{\beta}_k))'\partial^2\tilde{P}_2(\tilde{\omega}_{1,i};\hat{\beta}_k)\right\|$$
$$=O_p(1)+O_p((m_2+m_1^{1/2})\xi_{2,m_2}n^{-1/2}). \tag{SB.215}$$

By the Markov inequality, Assumptions SB1(i) and SB3(iv), Lemma SB1, Lemma SB16, (SB.214) and

(SB.215), we get

$$n^{-1}\sum_{i=1}^{n}\upsilon_{2,*}'\begin{pmatrix}\hat{P}_{2,i}^{*}(\hat{\beta}_k)-\tilde{P}_{2,i}(\beta_{k,0})\\-\partial^1\tilde{P}_{2,i}(\beta_{k,0})(\hat{\omega}_{1,i}(\hat{\beta}_k)-\omega_{1,i}(\beta_{k,0}))\end{pmatrix}(k_{2,i}-k_{1,i}g_{1,i})$$
$$=(\hat{\beta}_k-\beta_{k,0})o_p(1)+o_p(n^{-1/2}). \tag{SB.216}$$

Similarly, we can show that

$$n^{-1}\sum_{i=1}^{n}\upsilon_{2,*}'\begin{pmatrix}\hat{P}_{2,i}(\beta_{k,0})-\tilde{P}_{2,i}(\beta_{k,0})\\-\partial^1\tilde{P}_{2,i}(\beta_{k,0})(\hat{\omega}_{1,i}(\beta_{k,0})-\omega_{1,i}(\beta_{k,0})\end{pmatrix}(k_{2,i}-k_{1,i}g_{1,i})=o_p(n^{-1/2}). \tag{SB.217}$$

Since $\hat{\omega}_{1,i}(\hat{\beta}_k)-\hat{\omega}_{1,i}(\beta_{k,0})=-k_{1,i}(\hat{\beta}_k-\beta_{k,0})$, using (SB.216) and (SB.217) we get

$$n^{-1}\sum_{i=1}^{n}\upsilon_{2,*}'(\hat{P}_{2,i}^{*}(\hat{\beta}_k)-\hat{P}_{2,i}(\beta_{k,0}))(k_{2,i}-k_{1,i}g_{1,i})$$
$$=-(\hat{\beta}_k-\beta_{k,0})n^{-1}\sum_{i=1}^{n}\upsilon_{2,*}'\partial^1\tilde{P}_{2,i}(\beta_{k,0})k_{1,i}(k_{2,i}-k_{1,i}g_{1,i})+(\hat{\beta}_k-\beta_{k,0})o_p(1)+o_p(n^{-1/2}). \tag{SB.218}$$

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

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

Therefore

$$n^{-1}\sum_{i=1}^{n}\upsilon_{2,*}'\partial^1\tilde{P}_{2,i}(\beta_{k,0})k_{1,i}(k_{2,i}-k_{1,i}g_{1,i})$$
$$=\mathbb{E}[k_{1,i}g_{1,i}(v_{2,i}-v_{1,i}g_{1,i})]+n^{-1}\sum_{i=1}^{n}(g_{1,i}k_{1,i}(k_{2,i}-k_{1,i}g_{1,i})-\mathbb{E}[k_{1,i}g_{1,i}(v_{2,i}-v_{1,i}g_{1,i})])$$
$$+n^{-1}\sum_{i=1}^{n}(\hat{\beta}_g(\hat{\beta}_k)'\partial^1\tilde{P}_2(\omega_{1,i}(\beta_{k,0});\hat{\beta}_k)-g_{1,i})k_{1,i}(k_{2,i}-k_{1,i}g_{1,i}). \tag{SB.219}$$

By Assumption SB3(iv) and Lemma SB16, $\omega_{1,i}\in\Omega_{\varepsilon_n}(\hat{\beta}_k)$ for any $i\leq n$ wpa1. Therefore by Assumption SB2(v),

$$\max_{i\leq n}\left\|\partial^1\tilde{P}_2(\omega_{1,i}(\beta_{k,0});\hat{\beta}_k)\right\|=O_p(\xi_{1,m_2}). \tag{SB.220}$$

By the triangle inequality and the Cauchy-Schwarz inequality, Assumptions SB2(ii, iii, iv) and SB3(iv), Lemma SB11 and Lemma SB16, and (SB.220),

$$
\begin{aligned}
&n^{-1}\sum_{i=1}^{n}\left|\hat{\beta}_g(\hat{\beta}_k)'\partial^1\tilde{P}_2(\omega_{1,i}(\beta_{k,0});\hat{\beta}_k)-g_1(\omega_{1,i}(\beta_{k,0});\beta_{k,0})\right| \\
&\leq n^{-1}\sum_{i=1}^{n}\left|(\hat{\beta}_g(\hat{\beta}_k)-\tilde{\beta}_{g,m_2}(\hat{\beta}_k))'\partial^1\tilde{P}_2(\omega_{1,i}(\beta_{k,0});\hat{\beta}_k)\right| \\
&+n^{-1}\sum_{i=1}^{n}\left|\tilde{\beta}_{g,m_2}(\hat{\beta}_k)'\partial^1\tilde{P}_2(\omega_{1,i}(\beta_{k,0});\hat{\beta}_k)-g_1(\omega_{1,i}(\beta_{k,0});\hat{\beta}_k)\right| \\
&+n^{-1}\sum_{i=1}^{n}\left|g_1(\omega_{1,i}(\beta_{k,0});\hat{\beta}_k)-g_1(\omega_{1,i}(\beta_{k,0});\beta_{k,0})\right| \\
&=O_p(\xi_{1,m_2}(m_2+m_1^{1/2})n^{-1/2})+O_p(m^{-r_g})+(\hat{\beta}_k-\beta_{k,0})O_p(1)=o_p(1) \qquad \text{(SB.221)}
\end{aligned}
$$

which together with Assumptions SB1(i) and SB2(ii) implies that

$$
n^{-1}\sum_{i=1}^{n}(\hat{\beta}_g(\hat{\beta}_k)'\partial^1\tilde{P}_2(\omega_{1,i}(\beta_{k,0});\hat{\beta}_k)-g_{1,i})k_{1,i}(k_{2,i}-k_{1,i}g_{1,i})=o_p(1). \qquad \text{(SB.222)}
$$

By Assumptions SB1(i), SB2(ii, iii, vi) and SB3(iv), Lemma SB1 and Lemma SB6

$$
n^{-1}\sum_{i=1}^{n}(\hat{\beta}_g(\beta_{k,0})'\partial^1\tilde{P}_2(\omega_{1,i}(\beta_{k,0});\beta_{k,0})-g_{1,i})k_{1,i}(k_{2,i}-k_{1,i}g_{1,i})=o_p(1). \qquad \text{(SB.223)}
$$

By Assumptions SB1(i) and SB2(ii), and the Markov inequality,

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

which together with (SB.219), (SB.222) and (SB.223) implies that

$$
n^{-1}\sum_{i=1}^{n}\upsilon_{2,*}'\partial^1\tilde{P}_{2,i}(\beta_{k,0})k_{1,i}(k_{2,i}-k_{1,i}g_{1,i})=\mathbb{E}[k_{1,i}g_{1,i}(v_{2,i}-v_{1,i}g_{1,i})]+o_p(1). \qquad \text{(SB.224)}
$$

The claim in (SB.212) follows from (SB.218) and (SB.224).

Step 2. In this step, we show that

$$
\begin{aligned}
&\upsilon_{2,*}'\left[\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0})-\hat{\mathbf{P}}_2^*(\hat{\beta}_k)'\hat{\mathbf{P}}_2^*(\hat{\beta}_k)\right]\hat{\beta}_\varphi(\beta_{k,0})\\
&=(\hat{\beta}_k-\beta_{k,0})\left(\mathbb{E}[g_{1,i}k_{1,i}(a_{2,i}-a_{1,i}g_{1,i})]+o_p(1)\right)\\
&\quad+n^{-1}\upsilon_{2,*}'\hat{\mathbf{P}}_2^*(\hat{\beta}_k)'(\hat{\mathbf{P}}_2(\beta_{k,0})-\hat{\mathbf{P}}_2^*(\hat{\beta}_k))\hat{\beta}_\varphi(\beta_{k,0})+o_p(n^{-1/2}),
\end{aligned}
\tag{SB.225}
$$

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

Since we can write

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

to prove (SB.225) it is sufficient to show that

$$
n^{-1}\upsilon_{2,*}'(\hat{\mathbf{P}}_2(\beta_{k,0})-\hat{\mathbf{P}}_2^*(\hat{\beta}_k))'\hat{\mathbf{P}}_2(\beta_{k,0})\hat{\beta}_\varphi(\beta_{k,0})=(\hat{\beta}_k-\beta_{k,0})\left(\mathbb{E}\left[g_{1,i}k_{1,i}\varphi_i\right]+o_p(1)\right)+o_p(n^{-1/2}),\tag{SB.226}
$$

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

By the Cauchy-Schwarz inequality, Assumption SB3(iv), (SB.53), (SB.112) and (SB.186),

$$
\begin{aligned}
&n^{-1}\left|\upsilon_{2,*}'(\hat{\mathbf{P}}_2(\beta_{k,0})-\hat{\mathbf{P}}_2^*(\hat{\beta}_k))'\hat{\mathbf{P}}_2(\beta_{k,0})(\hat{\beta}_\varphi(\beta_{k,0})-\tilde{\beta}_{\varphi,m_2}(\beta_{k,0}))\right|\\
&\leq n^{-1}\left\|\upsilon_{2,*}'(\hat{\mathbf{P}}_2(\beta_{k,0})-\hat{\mathbf{P}}_2^*(\hat{\beta}_k))\right\|\left\|\hat{\mathbf{P}}_2(\beta_{k,0})(\hat{\beta}_\varphi(\beta_{k,0})-\tilde{\beta}_{\varphi,m_2}(\beta_{k,0}))\right\|\\
&\leq\frac{\left\|\upsilon_{2,*}'(\hat{\mathbf{P}}_2(\beta_{k,0})-\hat{\mathbf{P}}_2^*(\hat{\beta}_k))\right\|\left\|\hat{\beta}_\varphi(\beta_{k,0})-\tilde{\beta}_{\varphi,m_2}(\beta_{k,0})\right\|}{(\lambda_{\max}(n^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0})))^{-1/2}n^{1/2}}\\
&=|\hat{\beta}_k-\beta_{k,0}|O_p((m_1^{1/2}+m_2^{1/2})n^{-1/2})=|\hat{\beta}_k-\beta_{k,0}|o_p(1).
\end{aligned}
\tag{SB.227}
$$

By the Cauchy-Schwarz inequality, Assumptions SB3(i, ii, iv), Lemma SB1, (SB.48) and (SB.186),

$$
\begin{aligned}
&n^{-1}\left|\upsilon_{2,*}'(\hat{\mathbf{P}}_2(\beta_{k,0})-\hat{\mathbf{P}}_2^*(\hat{\beta}_k))'(\hat{\mathbf{P}}_2(\beta_{k,0})-\tilde{\mathbf{P}}_2(\beta_{k,0}))'\tilde{\beta}_{\varphi,m_2}(\beta_{k,0})\right|\\
&\leq n^{-1}\left\|(\hat{\mathbf{P}}_2(\beta_{k,0})-\hat{\mathbf{P}}_2^*(\hat{\beta}_k))\upsilon_{2,*}\right\|\left\|(\hat{\mathbf{P}}_2(\beta_{k,0})-\tilde{\mathbf{P}}_2(\beta_{k,0}))\tilde{\beta}_{\varphi,m_2}(\beta_{k,0})\right\|_S\\
&=|\hat{\beta}_k-\beta_{k,0}|O_p(m_1^{1/2}n^{-1/2})=|\hat{\beta}_k-\beta_{k,0}|o_p(1).
\end{aligned}
\tag{SB.228}
$$

By (SB.186) and Assumption SB3(ii, iv),

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

where $\varphi_n=(\varphi_1,\ldots,\varphi_n)'$, which together with (SB.227) and (SB.228) implies that

$$\begin{aligned}
& n^{-1}\upsilon_{2,*}'(\hat{\mathbf{P}}_2(\beta_{k,0})-\hat{\mathbf{P}}_2^*(\hat{\beta}_k))'\hat{\mathbf{P}}_2(\beta_{k,0})\hat{\beta}_\varphi(\beta_{k,0}) \\
& =n^{-1}\upsilon_{2,*}'(\hat{\mathbf{P}}_2(\beta_{k,0})-\hat{\mathbf{P}}_2^*(\hat{\beta}_k))'\varphi_n+(\hat{\beta}_k-\beta_{k,0})o_p(1).
\end{aligned} \tag{SB.229}$$

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

$$\upsilon_{2,*}'(\hat{\mathbf{P}}_2(\beta_{k,0})-\hat{\mathbf{P}}_2^*(\hat{\beta}_k))'\varphi_n=\sum_{i=1}^n\hat{\beta}_g(\hat{\beta}_k)'(\tilde{P}_2(\hat{\omega}_{1,i}(\beta_{k,0});\hat{\beta}_k)-\tilde{P}_2(\hat{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k))\varphi_i. \tag{SB.230}$$

By the first-order expansion, the triangle inequality and the Cauchy-Schwarz inequality, Assumptions SB1(i) and SB3(i, iv), Lemma SB11 and Lemma SB16, we have

$$\begin{aligned}
& n^{-1}\sum_{i=1}^n(\hat{\beta}_g(\hat{\beta}_k)-\tilde{\beta}_{g,m_2}(\hat{\beta}_k))'(\tilde{P}_2(\hat{\omega}_{1,i}(\beta_{k,0});\hat{\beta}_k)-\tilde{P}_2(\hat{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k))\varphi_i \\
& =(\hat{\beta}_k-\beta_{k,0})n^{-1}\sum_{i=1}^n(\hat{\beta}_g(\hat{\beta}_k)-\tilde{\beta}_{g,m_2}(\hat{\beta}_k))'\partial^1\tilde{P}_2(\tilde{\omega}_{1,i};\hat{\beta}_k)k_{1,i}\varphi_i \\
& =(\hat{\beta}_k-\beta_{k,0})O_p((m_1^{1/2}+m_2)n^{-1/2})O_p(\xi_{1,m_2})=(\hat{\beta}_k-\beta_{k,0})o_p(1).
\end{aligned} \tag{SB.231}$$

By Assumptions SB1(i), SB2(iii) and SB3(i),

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

and

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

which together with (SB.230) and (SB.231) implies that

$$
\begin{aligned}
& n^{-1}\upsilon_{2,*}'(\hat{\mathbf{P}}_2(\beta_{k,0}) - \hat{\mathbf{P}}_2^*(\hat{\beta}_k))'\varphi_n \\
&= n^{-1}\sum_{i=1}^{n}(g(\hat{\omega}_{1,i}(\beta_{k,0});\hat{\beta}_k) - g(\hat{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k))k_{1,i}\varphi_i + (\hat{\beta}_k - \beta_{k,0})o_p(1) + o_p(n^{-1/2}). \qquad \text{(SB.232)}
\end{aligned}
$$

By Assumptions SB1(i), SB2(ii) and SB3(i), Lemma SB1 and Lemma SB16

$$
n^{-1}\sum_{i=1}^{n} k_{1,i}\varphi_i(g_1(\hat{\omega}_{1,i}(\beta_{k,0});\hat{\beta}_k) - g_1(\omega_{1,i}(\beta_{k,0});\beta_{k,0})) = o_p(1). \qquad \text{(SB.233)}
$$

By Assumptions SB1(i), SB2(ii) and SB3(i), and Lemma SB16

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

which together with (SB.233) implies that

$$
\begin{aligned}
& n^{-1}\sum_{i=1}^{n}(g(\hat{\omega}_{1,i}(\beta_{k,0});\hat{\beta}_k) - g(\hat{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k))k_{1,i}\varphi_i \\
&= (\hat{\beta}_k - \beta_{k,0})n^{-1}\sum_{i=1}^{n} g_{1,i}k_{1,i}\varphi_i + (\hat{\beta}_k - \beta_{k,0})o_p(1) \\
&= (\hat{\beta}_k - \beta_{k,0})\mathbb{E}[g_{1,i}k_{1,i}\varphi_i] + (\hat{\beta}_k - \beta_{k,0})o_p(1) + o_p(n^{-1/2}), \qquad \text{(SB.234)}
\end{aligned}
$$

where the second equality is by the Markov inequality. The claim in (SB.226) now follows from (SB.229), (SB.232) and (SB.234).

Step 3. In this step, we show that

$$
n^{-1}(\hat{\mathbf{Y}}_2^*(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\hat{\beta}_k)\hat{\beta}_g(\hat{\beta}_k))'(\hat{\mathbf{P}}_2^*(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\beta_{k,0}))\hat{\beta}_\varphi(\beta_{k,0}) = (\hat{\beta}_k - \beta_{k,0})o_p(1) + o_p(n^{-1/2}). \qquad \text{(SB.235)}
$$

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

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

Therefore,

$$
\begin{aligned}
& n^{-1}(\hat{\mathbf{Y}}_2^*(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\hat{\beta}_k)\hat{\beta}_g(\hat{\beta}_k))'(\hat{\mathbf{P}}_2^*(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\beta_{k,0}))\hat{\beta}_\varphi(\beta_{k,0}) \\
& = n^{-1}\mathbf{U}_2'(\hat{\mathbf{P}}_2^*(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\beta_{k,0}))\hat{\beta}_\varphi(\beta_{k,0}) \\
& - (\hat{\beta}_l - \beta_{l,o})n^{-1}\mathbf{L}_2'(\hat{\mathbf{P}}_2^*(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\beta_{k,0}))\hat{\beta}_\varphi(\beta_{k,0}) \\
& - n^{-1}(\hat{\beta}_k - \beta_{l,o})\mathbf{K}_2'(\hat{\mathbf{P}}_2^*(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\beta_{k,0}))\hat{\beta}_\varphi(\beta_{k,0}) \\
& - n^{-1}(\hat{\mathbf{G}}_2 - \mathbf{G}_2)'(\hat{\mathbf{P}}_2^*(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\beta_{k,0}))\hat{\beta}_\varphi(\beta_{k,0}) \qquad \text{(SB.236)}
\end{aligned}
$$

which combined with Lemma SB19 proves (SB.235).

Step 4. In this step, we show that

$$
n^{-1}(\hat{\mathbf{Y}}_2^*(\hat{\beta}_k) - \hat{\mathbf{Y}}_2^*(\beta_{k,0}))'\hat{\mathbf{P}}_2(\beta_{k,0})\hat{\beta}_\varphi(\beta_{k,0}) = -(\hat{\beta}_k - \beta_{k,0})\left(\mathbb{E}[k_{2,i}(a_{2,i} - a_{1,i}g_{1,i})] + o_p(1)\right). \qquad \text{(SB.237)}
$$

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

$$
n^{-1}(\hat{\mathbf{Y}}_2^*(\hat{\beta}_k) - \hat{\mathbf{Y}}_2^*(\beta_{k,0}))'\hat{\mathbf{P}}_2(\beta_{k,0})\hat{\beta}_\varphi(\beta_{k,0}) = -n^{-1}(\hat{\beta}_k - \beta_{k,0})'\mathbf{K}_2'\hat{\mathbf{P}}_2(\beta_{k,0})\hat{\beta}_\varphi(\beta_{k,0}) \qquad \text{(SB.238)}
$$

and

$$
\begin{aligned}
n^{-1}\mathbf{K}_2'\hat{\mathbf{P}}_2(\beta_{k,0})\hat{\beta}_\varphi(\beta_{k,0}) = {} & \mathbb{E}[k_{2,i}\varphi(\omega_{1,i})] + n^{-1}\sum_{i=1}^n(k_{2,i}\varphi(\omega_{1,i}) - \mathbb{E}[k_{2,i}\varphi(\omega_{1,i})]) \\
& + n^{-1}\sum_{i=1}^n k_{2,i}(\varphi(\omega_{1,i}) - \varphi(\hat{\omega}_{1,i}(\beta_{k,0}))) \\
& + n^{-1}\sum_{i=1}^n k_{2,i}(\varphi(\hat{\omega}_{1,i}(\beta_{k,0})) - \hat{P}_{2,i}(\beta_{k,0})'\tilde{\beta}_{\varphi,m_2}(\beta_{k,0})) \\
& + n^{-1}\sum_{i=1}^n k_{2,i}\hat{P}_{2,i}(\beta_{k,0})'(\hat{\beta}_\varphi(\beta_{k,0}) - \tilde{\beta}_{\varphi,m_2}(\beta_{k,0})). \qquad \text{(SB.239)}
\end{aligned}
$$

By Assumptions SB1(i) and SB3(i, ii), and the Markov inequality and Lemma SB1, we have

$$
n^{-1}\sum_{i=1}^n(k_{2,i}\varphi(\omega_{1,i}) - \mathbb{E}[k_{2,i}\varphi(\omega_{1,i})]) = o_p(1) \qquad \text{(SB.240)}
$$

and

$$
n^{-1}\sum_{i=1}^n k_{2,i}(\varphi(\hat{\omega}_{1,i}(\beta_{k,0})) - \hat{P}_{2,i}(\beta_{k,0})'\tilde{\beta}_{\varphi,m_2}(\beta_{k,0})) = o_p(1) \qquad \text{(SB.241)}
$$

and

$$n^{-1}\sum_{i=1}^{n}k_{2,i}(\varphi(\omega_{1,i})-\varphi(\hat{\omega}_{1,i}(\beta_{k,0})))=o_p(1). \tag{SB.242}$$

By the Cauchy-Schwarz inequality, (SB.53) and (SB.110)

$$\begin{aligned}&\left|n^{-1}\sum_{i=1}^{n}k_{2,i}\hat{P}_{2,i}(\beta_{k,0})'(\hat{\beta}_{\varphi}(\beta_{k,0})-\tilde{\beta}_{\varphi,m_2}(\beta_{k,0}))\right|\\&\leq C\lambda_{\max}(n^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))\left\|\hat{\beta}_{\varphi}(\beta_{k,0})-\tilde{\beta}_{\varphi,m_2}(\beta_{k,0})\right\|=o_p(1).\end{aligned} \tag{SB.243}$$

The claim in (SB.237) follows from (SB.238), (SB.239), (SB.240), (SB.241) and (SB.242). *Q.E.D.*

SB.4 Auxiliary Results for the Standard Error Estimation

Assumption SB4 *(i) There exists $\hat{\varepsilon}_{1,i}$ for $i=1,\ldots,n$ such that $n^{-1}\sum_{i=1}^{n}(\hat{\varepsilon}_{1,i}-\varepsilon_{1,i})^4=o_p(1)$; (ii) there exist $r_h>1$ and $\beta_{h,m}\in\mathbb{R}^m$ such that $\sup_{x\in\mathcal{X}}|h_m(x)-h(x)|=O(m^{-r_h})$ where $h_m(x)\equiv P_1(x)'\beta_{h,m}$ and $\xi_{0,m_1}m^{-r_h}=o(1)$; (iii) $\Omega>0$.*

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

Lemma SB21 *Under Assumptions SB1, SB2 and SB3, we have*

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

(ii) $\max_{i\leq n}|\hat{g}_{1,i}-g_{1,i}|=O_p(\xi_{1,m_2}(m_2+m_1^{1/2})n^{-1/2})$;

(iii) $n^{-1}\sum_{i=1}^{n}(\hat{\varsigma}_i-v_{2,i}+v_{1,i}g_{1,i})^4=O_p\left((m_2^4+m_1^2)\xi_{1,m_2}^4\xi_{0,m_2}^2n^{-2}\right)$;

(iv) $n^{-1}\sum_{i=1}^{n}(\hat{u}_{2,i}-u_{2,i})^4=O_p((m_1^2+m_2^4)\xi_{0,m_2}^2n^{-2})$;

(v) $\max_{i\leq n}|\hat{h}_i-h_i|=o_p(1)$.

Proof of Lemma SB21. (i) For any $\upsilon_2\in\mathbb{R}^{m_2}$, by Assumption SB2(v) and (SA.7) in Theorem SA1

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

which implies that

$$\left\|\hat{\mathbf{P}}_2(\hat{\beta}_k)-\hat{\mathbf{P}}_2(\beta_{k,0})\right\|_S=O_p(\xi_{1,m_2}n^{-1/2}). \tag{SB.244}$$

By the triangle inequality and the Cauchy-Schwarz inequality, Assumption SB3(iv), (SB.53) and (SB.244)

$$
\begin{aligned}
& n^{-1}\left\|\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k)-\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0})\right\|_S \\
& \leq n^{-1}\left\|(\hat{\mathbf{P}}_2(\hat{\beta}_k)-\hat{\mathbf{P}}_2(\beta_{k,0}))'\hat{\mathbf{P}}_2(\beta_{k,0})\right\|_S \\
& + n^{-1}\left\|\hat{\mathbf{P}}_2(\beta_{k,0})'(\hat{\mathbf{P}}_2(\hat{\beta}_k)-\hat{\mathbf{P}}_2(\beta_{k,0}))\right\|_S \\
& + n^{-1}\left\|(\hat{\mathbf{P}}_2(\hat{\beta}_k)-\hat{\mathbf{P}}_2(\beta_{k,0}))'(\hat{\mathbf{P}}_2(\hat{\beta}_k)-\hat{\mathbf{P}}_2(\beta_{k,0}))\right\|_S = O_p(\xi_{1,m_2}n^{-1/2}) \qquad \text{(SB.245)}
\end{aligned}
$$

which proves the claim.

(ii) Using the similar arguments in deriving (SB.152), we can show that

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

which together with (SA.7) and Lemma SB11 implies that

$$
\max_{i\leq n}\left|\partial^1\hat{P}_{2,i}(\hat{\beta}_k)'\hat{\beta}_g(\hat{\beta}_k)-g_1(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k)\right| = O_p(\xi_{1,m_2}(m_2+m_1^{1/2})n^{-1/2}). \qquad \text{(SB.246)}
$$

Using similar arguments in showing (SB.155), we get

$$
\begin{aligned}
& \max_{i\leq n}\left|(\partial^1\hat{P}_{2,i}(\hat{\beta}_k)-\partial^1\tilde{P}_2(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k))'\hat{\beta}_g(\hat{\beta}_k)\right| \\
& \leq \max_{i\leq n}\left|(\partial^1\hat{P}_{2,i}(\hat{\beta}_k)-\partial^1\tilde{P}_2(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k))'\tilde{\beta}_{g,m_2}(\hat{\beta}_k)\right| \\
& + \max_{i\leq n}\left|(\partial^1\hat{P}_{2,i}(\hat{\beta}_k)-\partial^1\tilde{P}_2(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k))'(\hat{\beta}_g(\hat{\beta}_k)-\tilde{\beta}_{g,m_2}(\hat{\beta}_k))\right| \\
& \leq \max_{i\leq n}\left|g_1(\hat{\omega}_{1,i}(\hat{\beta}_k);\hat{\beta}_k)-g_1(\omega_{1,i}(\hat{\beta}_k);\hat{\beta}_k)\right|+O_p(\xi_{2,m_2}\xi_{1,m_1}m_1^{1/2}(m_2+m_1^{1/2})n^{-1}) \\
& = O_p(\xi_{1,m_2}(m_2+m_1^{1/2})n^{-1/2}). \qquad \text{(SB.247)}
\end{aligned}
$$

By Assumption SB2(ii) and (SA.7), we have

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

which together with (SB.246) and (SB.247) proves the second claim of the lemma.

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

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

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

$$n^{-1}\sum_{i=1}^{n}(\hat{\varsigma}_i - v_{2,i} + v_{1,i}g_{1,i})^4 \leq Cn^{-1}\sum_{i=1}^{n}(\Delta\hat{k}_{2,i} - \Delta k_{2,i})^4 + Cn^{-1}\sum_{i=1}^{n}(\hat{\varphi}_i - \varphi_i)^4. \tag{SB.248}$$

By Assumption SB3(ii), Lemma SB13, Lemma SB21(ii) and (SA.7),

$$\begin{aligned} n^{-1}\sum_{i=1}^{n}(\Delta\hat{k}_{2,i} - \Delta k_{2,i})^4 &= n^{-1}\sum_{i=1}^{n}k_{1,i}^4(\hat{g}_{1,i}(\hat{\beta}_k) - g_1(\omega_{1,i}(\beta_{k,0});\beta_{k,0}))^4 \\ &\leq Cn^{-1}\sum_{i=1}^{n}(\hat{g}_{1,i}(\hat{\beta}_k) - g_1(\omega_{1,i}(\beta_{k,0});\beta_{k,0}))^4 \\ &= O_p((m_2^4 + m_1^2)\xi_{1,m_2}^4 n^{-2}). \end{aligned} \tag{SB.249}$$

Similarly we can show that

$$n^{-1}\sum_{i=1}^{n}(\Delta\hat{k}_{2,i} - \Delta k_{2,i})^2 = O_p(\xi_{1,m_2}^2(m_2^2 + m_1)n^{-1}). \tag{SB.250}$$

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

$$\begin{aligned} \hat{\varphi}_i - \varphi_i &= \hat{P}_{2,i}(\hat{\beta}_k)'(\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k))^{-1}\hat{\mathbf{P}}_2(\hat{\beta}_k)(\Delta\hat{\mathbf{K}}_2 - \Delta\mathbf{K}_2) \\ &\quad + \hat{P}_{2,i}(\hat{\beta}_k)'(\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k))^{-1}(\hat{\mathbf{P}}_2(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\beta_{k,0}))\Delta\mathbf{K}_2 \\ &\quad + \hat{P}_{2,i}(\hat{\beta}_k)'[(\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k))^{-1} - (\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}]\hat{\mathbf{P}}_2(\beta_{k,0})\Delta\mathbf{K}_2 \\ &\quad + (\hat{P}_{2,i}(\hat{\beta}_k) - \hat{P}_{2,i}(\beta_{k,0}))'(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})\Delta\mathbf{K}_2 \\ &\quad + \hat{P}_{2,i}(\beta_{k,0})'(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})\Delta\mathbf{K}_2 - \varphi_i, \end{aligned} \tag{SB.251}$$

where $\Delta\hat{\mathbf{K}}_2 = (\Delta\hat{k}_{2,1}, \ldots, \Delta\hat{k}_{2,n})'$ and $\Delta\mathbf{K}_2 = (\Delta k_{2,1}, \ldots, \Delta k_{2,n})'$. By Assumption SB2(v), (SB.53) and

(SB.250),

$$
\begin{aligned}
&n^{-1}\sum_{i=1}^{n}(\hat{P}_{2,i}(\hat{\beta}_k)'(\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k))^{-1}\hat{\mathbf{P}}_2(\hat{\beta}_k)(\Delta\hat{\mathbf{K}}_2-\Delta\mathbf{K}_2))^4\\
&\le \xi_{0,m_2}^2\left\|(\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k))^{-1}\hat{\mathbf{P}}_2(\hat{\beta}_k)(\Delta\hat{\mathbf{K}}_2-\Delta\mathbf{K}_2)\right\|^2\\
&\times n^{-1}\sum_{i=1}^{n}(\hat{P}_{2,i}(\hat{\beta}_k)'(\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k))^{-1}\hat{\mathbf{P}}_2(\hat{\beta}_k)(\Delta\hat{\mathbf{K}}_2-\Delta\mathbf{K}_2))^2\\
&\le (\lambda_{\min}(n^{-1}\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k)))^{-1}\xi_{0,m_2}^2\left(n^{-1}\sum_{i=1}^{n}(\Delta\hat{k}_{2,i}-\Delta k_{2,i})^2\right)^2\\
&= O_p\left((m_2^4+m_1^2)\xi_{1,m_2}^4\xi_{0,m_2}^2 n^{-2}\right). \qquad \text{(SB.252)}
\end{aligned}
$$

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

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

where $\tilde{\omega}_{1,i}$ lies between $\hat{\omega}_{1,i}(\hat{\beta}_k)$ and $\hat{\omega}_{1,i}(\beta_{k,0})$, which together with Assumptions SB1(i) and SB2(ii), (SA.7) and (SB.53) implies that

$$
n^{-1}\sum_{i=1}^{n}\left|\hat{P}_{2,i}(\hat{\beta}_k)'(\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k))^{-1}(\hat{\mathbf{P}}_2(\hat{\beta}_k)-\hat{\mathbf{P}}_2(\beta_{k,0}))\Delta\mathbf{K}_2\right|^4 = O_p(\xi_{1,m_2}^4\xi_{0,m_2}^2 n^{-2}). \qquad \text{(SB.253)}
$$

By Assumptions SB2(iv, vi) and SB3(i, ii) and (SB.112),

$$\begin{aligned}\left\|\hat{\beta}_{\varphi}(\beta_{k,0})\right\| &\leq \left\|\tilde{\beta}_{\varphi,m_2}(\beta_{k,0})\right\| + \left\|\hat{\beta}_{\varphi}(\beta_{k,0}) - \tilde{\beta}_{\varphi,m_2}(\beta_{k,0})\right\| \\ &\leq (\lambda_{\min}(Q_{m_2}(\beta_{k,0})))^{-1/2}\mathbb{E}\left\|\varphi_{m_2}\right\|_2 + O_p((m_1^{1/2}+m_2^{1/2})n^{-1/2}) \\ &\leq C\mathbb{E}\left\|\varphi_{m_2}\right\|_2 + C\mathbb{E}\left\|\varphi_{m_2}-\varphi(\omega_{1,i})\right\|_2 + o_p(1) = O_p(1). \end{aligned} \quad \text{(SB.254)}$$

By the Cauchy-Schwarz inequality, Lemma SB21(i), (SB.65) and (SB.254),

$$\begin{aligned}& n^{-1}\sum_{i=1}^{n}\left|\hat{P}_{2,i}(\hat{\beta}_k)'[(\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k))^{-1} - (\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}]\hat{\mathbf{P}}_2(\beta_{k,0})\Delta\mathbf{K}_2\right|^4 \\ &= n^{-1}\sum_{i=1}^{n}\left|\hat{P}_{2,i}(\hat{\beta}_k)'(\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k))^{-1}[\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0})]\hat{\beta}_{\varphi}(\beta_{k,0})\right|^4 \\ &\leq \xi_{0,m_2}^2\left\|(\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k))^{-1}[\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0})]\hat{\beta}_{\varphi}(\beta_{k,0})\right\|^2 \\ &\times n^{-1}\sum_{i=1}^{n}\left|\hat{P}_{2,i}(\hat{\beta}_k)'(\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k))^{-1}[\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k) - \hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0})]\hat{\beta}_{\varphi}(\beta_{k,0})\right|^2 \\ &\leq \frac{\xi_{0,m_2}^2\left\|\hat{\beta}_{\varphi}(\beta_{k,0})\right\|^4}{(\lambda_{\min}(n^{-1}\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k)))^3}\left\|n^{-1}\hat{\mathbf{P}}_2(\hat{\beta}_k)'\hat{\mathbf{P}}_2(\hat{\beta}_k) - n^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0})\right\|_S^4 \\ &= O_p(\xi_{0,m_2}^2\xi_{1,m_2}^4 n^{-2}). \end{aligned} \quad \text{(SB.255)}$$

By the first order expansion, (SA.7) in Theorem SA1, Assumption SB3(ii) and (SB.254),

$$\begin{aligned}& n^{-1}\sum_{i=1}^{n}((\hat{P}_{2,i}(\hat{\beta}_k) - \hat{P}_{2,i}(\beta_{k,0}))'\hat{\beta}_{\varphi}(\beta_{k,0}))^4 \\ &= (\hat{\beta}_k - \beta_{k,0})^4 n^{-1}\sum_{i=1}^{n}(\partial\tilde{P}_2(\hat{\omega}_{1,i}(\tilde{\beta}_k);\tilde{\beta}_k)/\partial\beta_k)'\hat{\beta}_{\varphi}(\beta_{k,0}))^4 \\ &\leq (\hat{\beta}_k - \beta_{k,0})^4\xi_{1,m_2}^4\left\|\hat{\beta}_{\varphi}(\beta_{k,0})\right\|^4 = O_p(\xi_{1,m_2}^4 n^{-2}). \end{aligned} \quad \text{(SB.256)}$$

By Assumptions SB2(v) and SB3(i, iv), Lemma SB1, (SB.49), (SB.112), and (SB.254)

$$
\begin{aligned}
&n^{-1}\sum_{i=1}^{n}(\hat{P}_{2,i}(\beta_{k,0})'(\hat{\mathbf{P}}_2(\beta_{k,0})'\hat{\mathbf{P}}_2(\beta_{k,0}))^{-1}\hat{\mathbf{P}}_2(\beta_{k,0})\boldsymbol{\Delta}\mathbf{K}_2-\varphi_i)^4\\
&\leq Cn^{-1}\sum_{i=1}^{n}((\hat{P}_{2,i}(\beta_{k,0})-\tilde{P}_{2,i}(\beta_{k,0}))'\hat{\beta}_{\varphi}(\beta_{k,0}))^4\\
&+Cn^{-1}\sum_{i=1}^{n}(\tilde{P}_{2,i}(\beta_{k,0})'(\hat{\beta}_{\varphi}(\beta_{k,0})-\tilde{\beta}_{\varphi,m_2}(\beta_{k,0})))^4\\
&+Cn^{-1}\sum_{i=1}^{n}(\tilde{P}_{2,i}(\beta_{k,0})'\tilde{\beta}_{\varphi,m_2}(\beta_{k,0})-\varphi_i)^4\\
&\leq C\left\|\hat{\beta}_{\varphi}(\beta_{k,0})\right\|^4\xi_{1,m_2}^4 n^{-1}\sum_{i=1}^{n}(\hat{\phi}_i-\phi_i)^4\\
&+\xi_{1,m_2}^2\lambda_{\max}(n^{-1}\tilde{\mathbf{P}}_2(\beta_{k,0})'\tilde{\mathbf{P}}_2(\beta_{k,0}))\left\|\hat{\beta}_{\varphi}(\beta_{k,0})-\tilde{\beta}_{\varphi,m_2}(\beta_{k,0})\right\|^4+O(m_2^{-4r_{\varphi}})\\
&=O_p((\xi_{1,m_2}^4\xi_{0,m_2}^2 m_1^2+\xi_{1,m_2}^2 m_2^2)n^{-2}). \qquad \text{(SB.257)}
\end{aligned}
$$

Collecting the results in (SB.251), (SB.252), (SB.253), (SB.255), (SB.256) and (SB.257), we get

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

which together with (SB.248), (SB.249) and (SB.251) proves the third claim of the lemma.

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

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

which implies that

$$
\begin{aligned}
n^{-1}\sum_{i=1}^{n}(\hat{u}_{2,i}-u_{2,i})^4 &\leq C(\hat{\beta}_l-\beta_{l,0})^4 n^{-1}\sum_{i=1}^{n}l_{2,i}^4+C(\hat{\beta}_k-\beta_{k,0})^4 n^{-1}\sum_{i=1}^{n}k_{2,i}^4\\
&+Cn^{-1}\sum_{i=1}^{n}(\hat{g}_i(\hat{\beta}_k)-g(\omega_{1,i}(\beta_{k,0});\beta_{k,0}))^4\\
&=Cn^{-1}\sum_{i=1}^{n}(\hat{g}_i(\hat{\beta}_k)-g(\omega_{1,i}(\beta_{k,0});\beta_{k,0}))^4+O_p(n^{-2}), \qquad \text{(SB.258)}
\end{aligned}
$$

where the equality is by Assumptions SB1(i, iii) and SB2(i, ii), and Lemma SB16. Using similar arguments in showing Lemma SB12, we can show that

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

which together with (SA.7) and Lemma SB12 shows that

$$n^{-1}\sum_{i=1}^{n}(\hat{g}_i(\hat{\beta}_k)-g(\omega_{1,i}(\beta_{k,0});\beta_{k,0}))^4=O_p((m_1^2+m_2^4)\xi_{0,m_2}^2n^{-2}). \tag{SB.259}$$

The claim of the lemma follows from (SB.258) and (SB.259).

(v) Let $\hat{\beta}_h\equiv(\mathbf{P}_1'\mathbf{P}_1)^{-1}\sum_{i=1}^n P_1(x_{1,i})l_{1,i}$. By Assumptions SB1 and SB4(ii), we can use similar arguments in showing (SB.42) to get

$$\hat{\beta}_h-\beta_{h,m}=O_p(m_1^{1/2}n_1^{-1/2}+m_1^{-r_{h_1}}). \tag{SB.260}$$

Therefore by the triangle inequality, Assumption SB1(vi) and (SB.260),

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

where the second equality is by Assumptions SB1(vi) and SB4(ii). *Q.E.D.*

Lemma SB22 *Under Assumptions SB1, SB2, SB3 and SB4, we have*

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

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

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

PROOF OF LEMMA SB24. (i) By Assumptions SB1(i) and SB2(ii), and the Markov inequality

$$n^{-1}\sum_{i=1}^{n}(v_{2,i}-v_{1,i}g_{1,i})^2=\Upsilon+O_p(n^{-1/2})=O_p(1) \tag{SB.261}$$

which together with Assumption SB4(iv) and Lemma SB21(iii) proves the first claim of the lemma.

(ii) Let $\tilde{\Gamma}_n = \sum_{i=1}^{n}(l_{2,i} - h_{1,i}g_{1,i})(v_{2,i} - v_{1,i}g_{1,i})$. Then by Assumptions SB1(i, ii, iii) and SB2(i, ii), and the Slutsky Theorem, we have

$$n^{-1}\sum_{i=1}^{n}\varepsilon_{1,i}^2 = \mathbb{E}\left[\varepsilon_{1,i}^2\right] + O_p(n^{-1/2}) \tag{SB.262}$$

and

$$n^{-1}\sum_{i=1}^{n}(l_{2,i} - h_{1,i}g_{1,i})(v_{2,i} - v_{1,i}g_{1,i}) = \mathbb{E}\left[(l_{2,i} - h_{1,i}g_{1,i})(v_{2,i} - v_{1,i}g_{1,i})\right] + O_p(n^{-1/2}), \tag{SB.263}$$

which implies that

$$\tilde{\Gamma}_n = \Gamma + O_p(n^{-1/2}). \tag{SB.264}$$

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

$$\begin{aligned}
\hat{\Gamma}_n - \tilde{\Gamma}_n &= n^{-1}\sum_{i=1}^{n}\Big[(l_{2,i} - \hat{h}_i\hat{g}_{1,i})(\hat{v}_{2,i} - \hat{v}_{1,i}\hat{g}_{1,i}) - (l_{2,i} - h_i g_{1,i})(v_{2,i} - v_{1,i}g_{1,i})\Big] \\
&= -n^{-1}\sum_{i=1}^{n}(\hat{h}_i\hat{g}_{1,i} - h_i g_{1,i})(\hat{v}_{2,i} - \hat{v}_{1,i}\hat{g}_{1,i} - v_{2,i} + v_{1,i}g_{1,i}) \\
&\quad - n^{-1}\sum_{i=1}^{n}(\hat{h}_i\hat{g}_{1,i} - h_i g_{1,i})(v_{2,i} - v_{1,i}g_{1,i}) \\
&\quad + n^{-1}\sum_{i=1}^{n}(l_{2,i} - h_i g_{1,i})(\hat{v}_{2,i} - \hat{v}_{1,i}\hat{g}_{1,i} - v_{2,i} + v_{1,i}g_{1,i}).
\end{aligned} \tag{SB.265}$$

The second claim of the lemma follows from (SB.265), Assumption SB3(iv) and Lemma SB21(ii, iii, v).

(iii) Since $\hat{\eta}_{1,i} = \eta_{1,i} - l_{1,i}(\hat{\beta}_l - \beta_{l,0}) - (\hat{\phi}_i - \phi_i)$, by Assumptions SB1(ii, iii) and SB3(iv), the Markov inequality and Lemma SB1

$$\begin{aligned}
n^{-1}\sum_{i=1}^{n}(\hat{\eta}_{1,i} - \eta_{1,i})^4 &\leq C(\hat{\beta}_l - \beta_{l,0})^4 n^{-1}\sum_{i=1}^{n} l_{1,i}^4 + \max_{i\leq n}(\hat{\phi}_i - \phi_i)^2 n^{-1}\sum_{i=1}^{n}(\hat{\phi}_i - \phi_i)^2 \\
&= O_p(n^{-2}) + O_p(\xi_{0,m_1}^2 m_1^2 n^{-2}) = O_p(\xi_{0,m_1}^2 m_1^2 n^{-2}) = o_p(1).
\end{aligned} \tag{SB.266}$$

By Assumptions SB2(ii) and SB3(iv), and Lemma SB21(ii)

$$\max_{i\leq n}\hat{g}_{1,i}^4 \leq C\max_{i\leq n}(\hat{g}_{1,i} - g_{1,i})^4 + C\max_{i\leq n} g_{1,i}^4 = O_p(1). \tag{SB.267}$$

By Assumptions SB1(i, ii) and SB3(iv), Lemma SB21(ii, iv), (SB.266) and (SB.267), we get

$$
\begin{aligned}
&n^{-1}\sum_{i=1}^{n}(\hat{u}_{2,i}-\hat{\eta}_{1,i}\hat{g}_{1,i}-u_{2,i}+\eta_{1,i}g_{1,i})^4 \\
&\leq Cn^{-1}\sum_{i=1}^{n}(\hat{u}_{2,i}-u_{2,i})^4+C\max_{i\leq n}\hat{g}_{1,i}^4 n^{-1}\sum_{i=1}^{n}(\hat{\eta}_{1,i}-\eta_{1,i})^4 \\
&+C\max_{i\leq n}(\hat{g}_{1,i}-g_{1,i})^4 n^{-1}\sum_{i=1}^{n}\eta_{1,i}^4=o_p(1), \qquad \text{(SB.268)}
\end{aligned}
$$

which together with Lemma SB21(iii), Assumptions SB1(i, ii), SB2(i, ii) and SB3(iv) implies that

$$
n^{-1}\sum_{i=1}^{n}\begin{pmatrix} (\hat{u}_{2,i}-\hat{\eta}_{1,i}\hat{g}_{1,i})(\hat{v}_{2,i}-\hat{v}_{1,i}\hat{g}_{1,i}) \\ -(u_{2,i}-\eta_{1,i}g_{1,i})(v_{2,i}-v_{1,i}g_{1,i}) \end{pmatrix}^2=o_p(1). \qquad \text{(SB.269)}
$$

By Assumptions SB1(i, ii, iii) and SB4, and (SB.266), we have

$$
n^{-1}\sum_{i=1}^{n}\hat{\varepsilon}_{1,i}^4+n^{-1}\sum_{i=1}^{n}\hat{\eta}_{1,i}^4=O_p(1), \qquad \text{(SB.270)}
$$

which combined with Lemma SB22(ii), (SB.266) and Assumption SB4 implies that

$$
\begin{aligned}
n^{-1}\sum_{i=1}^{n}(\hat{\Gamma}_n\hat{\varepsilon}_{1,i}\hat{\eta}_{1,i}-\Gamma\varepsilon_{1,i}\eta_{1,i})^2 &\leq C(\hat{\Gamma}_n-\Gamma)^2 n^{-1}\sum_{i=1}^{n}\hat{\varepsilon}_{1,i}^2\hat{\eta}_{1,i}^2 \\
&+C\Gamma^2 n^{-1}\sum_{i=1}^{n}(\hat{\varepsilon}_{1,i}-\varepsilon_{1,i})^2\hat{\eta}_{1,i}^2 \\
&+C\Gamma^2 n^{-1}\sum_{i=1}^{n}\varepsilon_{1,i}^2(\hat{\eta}_{1,i}-\eta_{1,i})^2=o_p(1). \qquad \text{(SB.271)}
\end{aligned}
$$

Let $\tilde{\Omega}_n=n^{-1}\sum_{i=1}^{n}\left((u_{2,i}-\eta_{1,i}g_{1,i})(v_{2,i}-v_{1,i}g_{1,i})-\Gamma\varepsilon_{1,i}\eta_{1,i}\right)^2$. Then by Assumptions SB1(i) and SB2(ii), and the Markov inequality

$$
\tilde{\Omega}_n=\Omega+O_p(n^{-1/2}). \qquad \text{(SB.272)}
$$

By the definition of $\tilde{\Omega}_n$ and $\hat{\Omega}_n$, the triangle inequality and the Cauchy-Schwarz inequality, (SB.269), (SB.271) and (SB.272), we get

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

which together with (SB.272) proves the third claim of the Lemma. *Q.E.D.*

SB.5 Preliminary Results

Lemma SB23 (Matrix Bernstein) *Consider a finite sequence $\{d_i\}$ of independent, random matrices with dimension $m_1 \times m_2$. Assume that*

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

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

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

Then for any $t \geq 0$

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

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

Lemma SB24 *Let $S_{2,i}(\beta_k) = \tilde{P}_{2,i}(\beta_k)\tilde{P}_{2,i}(\beta_k)'$ where $\tilde{P}_{2,i}(\beta_k) = \tilde{P}_2(\omega_{1,i}(\beta_k), \beta_k)$ for any $\beta_k \in \Theta_k$. Then under Assumptions SB1(i) and SB2(iv, v, vi), we have*

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

Proof of Lemma SB24. For any $\beta_k \in \Theta_k$, by the triangle inequality and Assumptions SB2(iv, v),

$$\|S_{2,i}(\beta_k) - \mathbb{E}[S_{2,i}(\beta_k)]\|_S \leq \|S_{2,i}(\beta_k)\|_S + \|\mathbb{E}[S_{2,i}(\beta_k)]\|_S \leq C\xi_{0,m_2}^2. \tag{SB.273}$$

By Assumptions SB1(i) and SB2(iv, v),

$$\left\|\sum_{i=1}^{n}\mathbb{E}\left[(S_{2,i}(\beta_k) - \mathbb{E}[S_{2,i}(\beta_k)])^2\right]\right\|_S \leq n\left(\left\|\mathbb{E}\left[(S_{2,i}(\beta_k))^2\right]\right\|_S + \left\|(\mathbb{E}[S_{2,i}(\beta_k)])^2\right\|_S\right) \leq Cn\xi_{0,m_2}^2. \tag{SB.274}$$

Therefore we can use Lemma SB23 to deduce that

$$\mathbb{P}\left(\left\|n^{-1}\sum_{i=1}^{n}S_{2,i}(\beta_k) - \mathbb{E}[S_{2,i}(\beta_k)]\right\|_S \geq t\right) \leq 2m_2\exp\left(-\frac{1}{C}\frac{nt^2/2}{\xi_{0,m_2}^2(1+t/3)}\right) \tag{SB.275}$$

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

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

$$\begin{aligned}\|S_{2,i}(\beta_{k,1}) - S_{2,i}(\beta_{k,2})\|_S \leq{}& \left\|S_{2,i}(\beta_{k,1}) - \tilde{P}_{2,i}(\beta_{k,2})\tilde{P}_{2,i}(\beta_{k,1})'\right\|_S \\ &+ \left\|\tilde{P}_{2,i}(\beta_{k,2})\tilde{P}_{2,i}(\beta_{k,1})' - S_{2,i}(\beta_{k,2})\right\|_S.\end{aligned} \tag{SB.276}$$

By the mean value expansion and the Cauchy-Schwarz inequality, and Assumption SB2(v)

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

where $\tilde{\beta}_{k,12}$ lies between $\beta_{k,1}$ and $\beta_{k,2}$, which together with Assumption SB2(vi) implies that

$$\left\|S_{2,i}(\beta_{k,1}) - \tilde{P}_{2,i}(\beta_{k,2})\tilde{P}_{2,i}(\beta_{k,1})'\right\|_S \leq Cm_2^3|\beta_{k,2} - \beta_{k,1}|. \tag{SB.277}$$

The same upper bound can be established for the second term in the right hand side of the inequality of

(SB.276). Therefore,

$$\|S_{2,i}(\beta_{k,1}) - S_{2,i}(\beta_{k,2})\|_S \leq Cm_2^3 \left|\beta_{k,2} - \beta_{k,1}\right|. \quad \text{(SB.278)}$$

Similarly, we can show that

$$\|\mathbb{E}\left[S_{2,i}(\beta_{k,1})\right] - \mathbb{E}\left[S_{2,i}(\beta_{k,2})\right]\|_S \leq Cm_2^3 \left|\beta_{k,2} - \beta_{k,1}\right|. \quad \text{(SB.279)}$$

Combining the results in (SB.278) and (SB.279), and applying the triangle inequality, we get

$$\left\| \begin{array}{c} n^{-1}\sum_{i=1}^{n}\left(S_{2,i}(\beta_{k,1}) - \mathbb{E}\left[S_{2,i}(\beta_{k,1})\right]\right) \\ -n^{-1}\sum_{i=1}^{n}\left(S_{2,i}(\beta_{k,2}) - \mathbb{E}\left[S_{2,i}(\beta_{k,2})\right]\right) \end{array} \right\|_S \leq C_S m_2^3 \left|\beta_{k,2} - \beta_{k,1}\right|, \quad \text{(SB.280)}$$

where C_S is a finite fixed constant. Since the parameter space Θ_k is compact, there exist $\{\beta_k(l)\}_{l=1,\ldots,K_n}$ such that for any $\beta_k \in \Theta_k$

$$\min_{l=1,\ldots,K_n} |\beta_k - \beta_k(l)| \leq (C_S m_2^3 n^{1/2})^{-1}, \quad \text{(SB.281)}$$

where $K_n \leq 2C_S m_2^3 n^{1/2}$. For any $\beta_k \in \Theta_k$, by (SB.280) and (SB.281)

$$\left\|n^{-1}\sum_{i=1}^{n} S_{2,i}(\beta_k) - \mathbb{E}\left[S_{2,i}(\beta_k)\right]\right\|_S \leq \max_{l=1,\ldots,K_n}\left\|n^{-1}\sum_{i=1}^{n} S_{2,i}(\beta_k(l)) - \mathbb{E}\left[S_{2,i}(\beta_k(l))\right]\right\|_S + n^{-1/2}. \quad \text{(SB.282)}$$

Therefore for any $B > 1$,

$$\begin{aligned} &\mathbb{P}\left(\sup_{\beta_k\in\Theta_k}\left\|n^{-1}\sum_{i=1}^{n} S_{2,i}(\beta_k) - \mathbb{E}\left[S_{2,i}(\beta_k)\right]\right\|_S \geq B(\xi_{0,m_2}^2 \log(n) n^{-1})^{1/2}\right) \\ &\leq \mathbb{P}\left(\max_{l=1,\ldots,K_n}\left\|n^{-1}\sum_{i=1}^{n} S_{2,i}(\beta_k(l)) - \mathbb{E}\left[S_{2,i}(\beta_k(l))\right]\right\|_S \geq (B-1)(\xi_{0,m_2}^2 \log(n) n^{-1})^{1/2}\right) \\ &\leq \sum_{l=1}^{K_n}\mathbb{P}\left(\left\|n^{-1}\sum_{i=1}^{n} S_{2,i}(\beta_k(l)) - \mathbb{E}\left[S_{2,i}(\beta_k(l))\right]\right\|_S \geq (B-1)(\xi_{0,m_2}^2 \log(n) n^{-1})^{1/2}\right) \\ &\leq 2K_n m_2 \exp\left(-\frac{B}{C}\frac{\log(n)}{1+(\xi_{0,m_2}^2 \log(n) n^{-1})^{1/2}}\right), \end{aligned} \quad \text{(SB.283)}$$

where the last inequality is by (SB.275). The claim of the theorem follows from (SB.283) and Assumption SB2(vi). *Q.E.D.*

Lemma SB25 *Let* $u_{2,i}(\beta_k) = y_{2,i}^* - k_{2,i}\beta_k - g(\omega_{1,i}(\beta_k), \beta_k)$. *Then under Assumptions SB1 and SB2(ii,* *iii, v, vi), we have*

$$\sup_{\beta_k\in\Theta_k}\left\|n^{-1}\sum_{i=1}^{n}\tilde{P}_2(\omega_{1,i}(\beta_k),\beta_k)u_{2,i}(\beta_k)\right\| = O_p(m_2 n^{-1/2}).$$

Proof of Lemma SB25. Define $\pi_n(\beta_k) = n^{-1/2}\sum_{i=1}^{n}\tilde{P}_2(\omega_{1,i}(\beta_k),\beta_k)u_{2,i}(\beta_k)$. For any $\beta_k \in \Theta_k$, by Assumption SB2(i) and (SB.55),

$$\mathbb{E}\left[(u_{2,i}(\beta_k))^4\big|\,\omega_{1,i}(\beta_k)\right] \leq C\mathbb{E}\left[(y^*_{2,i})^4 + k^4_{2,i}\big|\,\omega_{1,i}(\beta_k)\right] + C\left|g(\omega_{1,i}(\beta_k);\beta_k)\right|^4 \leq C. \tag{SB.284}$$

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

$$\begin{aligned}
&\mathbb{E}\left[\|\pi_n(\beta_{k,1})-\pi_n(\beta_{k,2})\|^2\right]\\
&=\mathbb{E}\left[\left\|\tilde{P}_2(\omega_{1,i}(\beta_{k,1}),\beta_{k,1})u_{2,i}(\beta_{k,1})-\tilde{P}_2(\omega_{1,i}(\beta_{k,2}),\beta_{k,2})u_{2,i}(\beta_{k,2})\right\|^2\right]\\
&\leq 2\mathbb{E}\left[(u_{2,i}(\beta_{k,2}))^2\left\|\tilde{P}_2(\omega_{1,i}(\beta_{k,1}),\beta_{k,1})-\tilde{P}_2(\omega_{1,i}(\beta_{k,2}),\beta_{k,2})\right\|^2\right]\\
&\quad+2\mathbb{E}\left[\left\|\tilde{P}_2(\omega_{1,i}(\beta_{k,1}),\beta_{k,1})\right\|^2(u_{2,i}(\beta_{k,2})-u_{2,i}(\beta_{k,1}))^2\right].
\end{aligned}\tag{SB.285}$$

Consider any $\gamma\in\mathbb{R}^{m_2}$. By the mean value expansion and Assumption SB2(v)

$$\begin{aligned}
&\left|\gamma'(\tilde{P}_2(\omega_1(\beta_{k,1}),\beta_{k,1})-\tilde{P}_2(\omega_1(\beta_{k,2}),\beta_{k,2}))\right|^2\\
&=\left|\gamma'\partial\tilde{P}_2\left(\omega_{1,i}(\tilde{\beta}_{k,12});\tilde{\beta}_{k,12}\right)/\partial\beta_k\right|^2(\beta_{k,1}-\beta_{k,2})^2\leq\|\gamma\|^2\,\xi^2_{1,m_2}(\beta_{k,1}-\beta_{k,2})^2,
\end{aligned}$$

where $\tilde{\beta}_{k,12}$ lies between $\beta_{k,1}$ and $\beta_{k,2}$, which implies that

$$\left\|\tilde{P}_2(\omega_{1,i}(\beta_{k,1}),\beta_{k,1})-\tilde{P}_2(\omega_{1,i}(\beta_{k,2}),\beta_{k,2})\right\|^2\leq\xi^2_{1,m_2}(\beta_{k,1}-\beta_{k,2})^2. \tag{SB.286}$$

Therefore, by (SB.284) and (SB.286),

$$\mathbb{E}\left[(u_{2,i}(\beta_{k,2}))^2\left\|\tilde{P}_2(\omega_{1,i}(\beta_{k,1}),\beta_{k,1})-\tilde{P}_2(\omega_{1,i}(\beta_{k,2}),\beta_{k,2})\right\|^2\right]\leq C\xi^2_{1,m_2}(\beta_{k,2}-\beta_{k,1})^2. \tag{SB.287}$$

By the definition of $u_{2,i}(\beta_k)$, we can write

$$u_{1,i}(\beta_{k,2})-u_{1,i}(\beta_{k,1})=g(\omega_{1,i}(\beta_{k,1}),\beta_{k,1})-g(\omega_{1,i}(\beta_{k,2}),\beta_{k,2})+k_{2,i}(\beta_{k,2}-\beta_{k,1}).$$

Therefore, by Assumptions SB2(ii, iv), and the assumption that $k_{2,i}$ has bounded support, we have

$$\mathbb{E}\left[\left\|\tilde{P}_2(\omega_{1,i}(\beta_{k,1}),\beta_{k,1})\right\|^2 (u_{1,i}(\beta_{k,2})-u_{1,i}(\beta_{k,1}))^2\right] \leq Cm_2(\beta_{k,2}-\beta_{k,1})^2 \tag{SB.288}$$

which together with Assumption SB2(vi), (SB.285) and (SB.287) implies that

$$\left\|\left\|\pi_n(\beta_{k,1})-\pi_n(\beta_{k,2})\right\|\right\|_2 \leq Cm_2^2\left|\beta_{k,2}-\beta_{k,1}\right| \tag{SB.289}$$

for any $\beta_{k,1},\beta_{k,2}\in\Theta_k$.

We next use the chaining technique to prove the theorem. The proof follows similar arguments of proving Theorem 2.2.4 in van der Vaart and Wellner (1996). Construct nested sets $\Theta_{k,1}\subset\Theta_{k,2}\subset\cdots\subset\Theta_k$ such that $\Theta_{k,j}$ is a maximal set of points in the sense that for every $\beta_{k,j}$, $\beta'_{k,j}\in\Theta_{k,j}$ there is $\left|\beta_{k,j}-\beta'_{k,j}\right| > 2^{-j}$. Since Θ_k is a compact set, the number of the points in $\Theta_{k,j}$ is less than $C2^j$. Link every point $\beta_{k,j+1}\in\Theta_{k,j+1}$ to a unique $\beta_{k,j}\in\Theta_{k,j}$ such that $|\beta_{k,j+1}-\beta_{k,j}|\leq 2^{-j}$. Let $J_n=\min\{j:2^{-j}\leq Cm_2^{-1}\}$. Consider any positive integer $J>J_n$. Obtain for every $\beta_{k,J+1}$ a chain $\beta_{k,J+1},\ldots,\beta_{k,J_n}$ that connects it to a point β_{k,J_n} in Θ_{k,J_n}. For arbitrary points $\beta_{k,J+1}$, $\beta'_{k,J+1}$ in $\Theta_{k,J+1}$, by the triangle inequality

$$\begin{aligned}
&\left\|\pi_n(\beta_{k,J+1})-\pi_n(\beta_{k,J_n})-\left[\pi_n(\beta'_{k,J+1})-\pi_n(\beta'_{k,J_n})\right]\right\| \\
&=\left\|\sum_{j=J_n}^{J}\left[\pi_n(\beta_{k,j+1})-\pi_n(\beta_{k,j})\right]-\sum_{j=J_n}^{J}\left[\pi_n(\beta'_{k,j+1})-\pi_n(\beta'_{k,j})\right]\right\| \\
&\leq 2\sum_{j=J_n}^{J}\max\left\|\pi_n(\beta_{k,j+1})-\pi_n(\beta_{k,j})\right\|,
\end{aligned} \tag{SB.290}$$

where for fixed j the maximum is taken over all links $(\beta_{k,j+1},\beta_{k,j})$ from $\Theta_{k,j+1}$ to $\Theta_{k,j}$. Thus the jth maximum is taken over at most $C2^{j+1}$ many links. By Assumption SB2(vi), (SB.289), (SB.290), the triangle inequality and the finite maximum inequality,

$$\begin{aligned}
&\left\|\max\left\|\pi_n(\beta_{k,J+1})-\pi_n(\beta_{k,J_n})-\left[\pi_n(\beta'_{k,J+1})-\pi_n(\beta'_{k,J_n})\right]\right\|\right\|_2 \\
&\leq 2\sum_{j=J_n}^{J}\left\|\max\left\|\pi_n(\beta_{k,j+1})-\pi_n(\beta_{k,j})\right\|\right\|_2 \\
&\leq C\sum_{j=J_n}^{J}2^{j/2}\max\left\|\left\|\pi_n(\beta_{k,j+1})-\pi_n(\beta_{k,j})\right\|\right\|_2 \leq Cm_2^2\sum_{j=J_n}^{\infty}2^{-j/2}\leq Cm_2,
\end{aligned} \tag{SB.291}$$

where β_{k,J_n} and β'_{k,J_n} are the endpoints of the chains starting at $\beta_{k,J+1}$ and $\beta'_{k,J+1}$ respectively. Since the set Θ_{k,J_n} has at most Cm_2 many elements, by the finite maximum inequality, the triangle inequality, (SB.284) and Assumption SB2(iv)

$$\left\|\max\left\|\pi_n(\beta_{k,J_n})-\pi_n(\beta'_{k,J_n})\right\|\right\|_2 \leq Cm_2^{1/2}\max\left\|\left\|\pi_n(\beta_{k,J_n})\right\|\right\|_2 \leq Cm_2. \tag{SB.292}$$

Therefore, by the triangle inequality, (SB.291) and (SB.292),

$$\begin{aligned}&\left\|\max\left\|\pi_n(\beta_{k,J+1})-\pi_n(\beta'_{k,J+1})\right\|\right\|_2\\ &\leq \left\|\max\left\|\pi_n(\beta_{k,J+1})-\pi_n(\beta_{k,J_n})-\left[\pi_n(\beta'_{k,J+1})-\pi_n(\beta'_{k,J_n})\right]\right\|\right\|_2\\ &\quad+\left\|\max\left\|\pi_n(\beta_{k,J_n})-\pi_n(\beta'_{k,J_n})\right\|\right\|_2 \leq Cm_2.\end{aligned} \tag{SB.293}$$

Let J go to infinity, by (SB.293) we deduce that

$$\left\|\sup_{\beta_k,\beta'_k\in\Theta_k}\left\|\pi_n(\beta_k)-\pi_n(\beta'_k)\right\|\right\|_2 \leq Cm_2. \tag{SB.294}$$

By (SB.292), (SB.294) and the triangle inequality,

$$\left\|\sup_{\beta_k\in\Theta_k}\left\|\pi_n(\beta_k)\right\|\right\|_2 \leq \left\|\sup_{\beta_k\in\Theta_k}\left\|\pi_n(\beta_k)-\pi_n(\beta_{k,0})\right\|\right\|_2 + \left\|\left\|\pi_n(\beta_{k,0})\right\|\right\|_2 \leq Cm_2 \tag{SB.295}$$

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

Lemma SB26 *Under Assumptions SB1 and SB2(ii, iii, v, vi), we have*

$$\sup_{\beta_k\in\Theta_k}\left\|n^{-1}\sum_{i=1}^{n}u_{2,i}k_{1,i}\partial^1\tilde{P}_2(\omega_{1,i}(\beta_k),\beta_k)\right\| = O_p(m_2^{5/2}n^{-1/2}).$$

Proof of Lemma SB26. For ease of notations, we define $\pi_n(\beta_k)=n^{-1/2}\sum_{i=1}^{n}u_{2,i}k_{1,i}\partial^1\tilde{P}_2(\omega_{1,i}(\beta_k),\beta_k)$ for any $\beta_k\in\Theta_k$. By Assumptions SB1(i) and (SB.75), $\mathbb{E}\left[u_{2,i}^2k_{1,i}^2|x_{1,i}\right]\leq C$. Therefore for any $\beta_{k,1}$ and $\beta_{k,2}$, we can use similar arguments in showing (SB.287) to obtain

$$\left\|\left\|\pi_n(\beta_{k,1})-\pi_n(\beta_{k,2})\right\|\right\|_2 \leq C\xi_{2,m_2}\left|\beta_{k,1}-\beta_{k,2}\right|. \tag{SB.296}$$

Construct nested sets $\Theta_{k,1}\subset\Theta_{k,2}\subset\cdots\subset\Theta_k$ such that $\Theta_{k,j}$ is a maximal set of points in the sense that for every $\beta_{k,j}$, $\beta'_{k,j}\in\Theta_{k,j}$ there is $\left|\beta_{k,j}-\beta'_{k,j}\right|>2^{-j}$. Since Θ_k is a compact set, the number of the points in $\Theta_{k,j}$ is less than $C2^j$. Link every point $\beta_{k,j+1} \in \Theta_{k,j+1}$ to a unique $\beta_{k,j} \in \Theta_{k,j}$ such that $|\beta_{k,j+1} - \beta_{k,j}| \leq 2^{-j}$. Let $J_n = \min\{j : 2^{-j} \leq Cm_2^{-1}\}$. Consider any positive integer $J > J_n$. Obtain for every $\beta_{k,J+1}$ a chain $\beta_{k,J+1}, \ldots, \beta_{k,J_n}$ that connects it to a point β_{k,J_n} in Θ_{k,J_n}. For arbitrary points $\beta_{k,J+1}$, $\beta'_{k,J+1}$ in $\Theta_{k,J+1}$, by the triangle inequality and (SB.296)

$$\begin{aligned}
&\left\|\max\left\|\pi_n(\beta_{k,J+1}) - \pi_n(\beta_{k,J_n}) - \left[\pi_n(\beta'_{k,J+1}) - \pi_n(\beta'_{k,J_n})\right]\right\|\right\|_2 \\
&\leq 2\sum_{j=J_n}^{J} \left\|\max\left\|\pi_n(\beta_{k,j+1}) - \pi_n(\beta_{k,j})\right\|\right\|_2 \\
&\leq C\sum_{j=J_n}^{J} 2^{j/2}\max\left\|\left\|\pi_n(\beta_{k,j+1}) - \pi_n(\beta_{k,j})\right\|\right\|_2 \\
&\leq \xi_{2,m_2}\sum_{j=J_n}^{\infty} 2^{-j/2} \leq C\xi_{2,m_2}m_2^{-1}, \qquad \text{(SB.297)}
\end{aligned}$$

where β_{k,J_n} and β'_{k,J_n} are the endpoints of the chains starting at $\beta_{k,J+1}$ and $\beta'_{k,J+1}$ respectively. Since the set Θ_{k,J_n} has at most Cm_2 many elements, by the finite maximum inequality, the triangle inequality, (SB.284) and Assumption SB2(iii)

$$\left\|\max\left\|\pi_n(\beta_{k,J_n}) - \pi_n(\beta'_{k,J_n})\right\|\right\|_2 \leq Cm_2^{1/2}\max\left\|\left\|\pi_n(\beta_{k,J_n})\right\|\right\|_2 \leq Cm_2^{5/2}. \qquad \text{(SB.298)}$$

Then the claim of the lemma follows by applying the chaining arguments in the proof of Lemma SB25. *Q.E.D.*

Lemma SB27 *Under Assumptions SB1 and SB2(ii, iii, v, vi), we have*

$$\sup_{\beta_k \in \Theta_k}\left\|n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}\partial^1\tilde{P}_{2,i}(\beta_k)P_1(x_{1,i})'\right\| = O_p(m_2^{5/2}m_1^{1/2}n^{-1/2}).$$

Proof of Lemma SB27. For ease of notations, we define $n^{-1}\sum_{i=1}^{n} u_{2,i}k_{1,i}\partial^1\tilde{P}_{2,i}(\beta_k)P_1(x_{1,i})'$ for any $\beta_k \in \Theta_k$. By Assumptions SB1(i) and (SB.75), $\mathbb{E}\left[u_{2,i}^2k_{1,i}^2|x_{1,i}\right] \leq C$. Therefore for any $\beta_{k,1}$ and $\beta_{k,2}$, we can use similar arguments in showing (SB.287) to obtain

$$\left\|\left\|\pi_n(\beta_{k,1}) - \pi_n(\beta_{k,2})\right\|\right\|_2 \leq Cm_1^{1/2}\xi_{2,m_2}\left|\beta_{k,1} - \beta_{k,2}\right|. \qquad \text{(SB.299)}$$

Construct nested sets $\Theta_{k,1} \subset \Theta_{k,2} \subset \cdots \subset \Theta_k$ such that $\Theta_{k,j}$ is a maximal set of points in the sense that for every $\beta_{k,j}$, $\beta'_{k,j} \in \Theta_{k,j}$ there is $\left|\beta_{k,j} - \beta'_{k,j}\right| > 2^{-j}$. Since Θ_k is a compact set, the number of the points in $\Theta_{k,j}$ is less than $C2^j$. Link every point $\beta_{k,j+1} \in \Theta_{k,j+1}$ to a unique $\beta_{k,j} \in \Theta_{k,j}$ such that $|\beta_{k,j+1} - \beta_{k,j}| \leq 2^{-j}$. Let $J_n = \min\{j : 2^{-j} \leq Cm_2^{-1}\}$. Consider any positive integer $J > J_n$. Obtain for every $\beta_{k,J+1}$ a chain $\beta_{k,J+1}, \ldots, \beta_{k,J_n}$ that connects it to a point β_{k,J_n} in Θ_{k,J_n}. For arbitrary points $\beta_{k,J+1}$, $\beta'_{k,J+1}$ in $\Theta_{k,J+1}$, by the triangle inequality and (SB.296)

$$\left\|\max\left\|\pi_n(\beta_{k,J+1}) - \pi_n(\beta_{k,J_n}) - \left[\pi_n(\beta'_{k,J+1}) - \pi_n(\beta'_{k,J_n})\right]\right\|\right\|_2 \leq C\xi_{2,m_2} m_1^{1/2} m_2^{-1}. \tag{SB.300}$$

Since the set Θ_{k,J_n} has at most Cm_2 many elements, by the finite maximum inequality, the triangle inequality, (SB.284) and Assumption SB2(iii)

$$\left\|\max\left\|\pi_n(\beta_{k,J_n}) - \pi_n(\beta'_{k,J_n})\right\|\right\|_2 \leq Cm_2^{1/2} \max\left\|\left\|\pi_n(\beta_{k,J_n})\right\|\right\|_2 \leq C\xi_{1,m_2} m_1^{1/2} m_2^{1/2}. \tag{SB.301}$$

Then the claim of the lemma follows by applying the chaining arguments in the proof of Lemma SB25. *Q.E.D.*

References

Ackerberg, D. A., K. Caves, and G. Frazer (2015): "Identification properties of recent production function estimators," *Econometrica*, 83(6), 2411–2451.

Andrews, D. W. K. (1991): "Asymptotic Normality of Series Estimators for Nonparametric and Semiparametric Regression Models," *Econometrica*, 59(2), 307–345.

Belloni, A., V. Chernozhukov, D. Chetverikov, and K. Kato (2015): "Some new asymptotic theory for least squares series: Pointwise and uniform results," *Journal of Econometrics*, 186(2), 345 – 366, High Dimensional Problems in Econometrics.

Chen, X. (2007): "Large Sample Sieve Estimation of Semi-Nonparametric Models," in *Handbook of Econometrics*, ed. by J. Heckman, and E. Leamer, vol. 6B, chap. 76. Elsevier, 1 edn.

Hahn, J., Z. Liao, and G. Ridder (2022): "Identification and the Influence Function of Olley and Pakes'(1996) Production Function Estimator," *Working Paper, UCLA and USC.*

Newey, W. K. (1997): "Convergence rates and asymptotic normality for series estimators," *Journal of Econometrics*, 79(1), 147 – 168.

Olley, G. S., and A. Pakes (1996): "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica: Journal of the Econometric Society*, pp. 1263–1297.

Tropp, J. A. (2012): "User-Friendly Tail Bounds for Sums of Random Matrices," *Foundations of Computational Mathematics*, 12(4), 389–434.

van der Vaart, A., and J. Wellner (1996): *Weak Convergence and Empirical Processes*. Springer-Verlag.