# Appendix for "Inference in models with partially identified control functions"

Andres Aradillas-Lopez[†]

**Abstract**

This document includes the step-by-step proofs of Result 1 and Lemma 1 in the paper, along with additional results and extensions referenced throughout the paper, such as the description of our estimator for the variance of our test-statistic and its asymptotic properties. Every section in this document has the format **AX.X** and every equation has the format **(A-XX)**. Any section or equation that we reference here which does not have this format refers to a section or an equation in the paper.

## A1 Proof of Result 1

The statements in Result 1 are a summary of the results described in equations (10), (11), (12), (13) and (14) in the text. We will show here that these equations follow from the restrictions (R1), (R2), (R3) and (R4). In what follows, $(V_i, V_j)$ represent to independent draws from $F$. We begin with equation (10). By restriction (R1), for any $\beta_1 \in \Theta$, there exists $\underline{d} < \overline{d}$ such that $[\underline{d}, \overline{d}] \subseteq Supp(X_{1L}'\beta_1) \cap Supp(X_{1U}'\beta_1)$. From here, it follows that $P_F(X_{1Uj}'\beta_1 \leq X_{1Li}'\beta_1) > 0$ for all $\beta_1 \in \Theta$. Next, also by (R2), for any pair $\beta_1 \neq \widetilde{\beta}_1$ in $\Theta$, there exist $\underline{c} < \overline{c}$ such that $[\underline{c}, \overline{c}] \subseteq Supp(X_{1L}'\beta_1 | X_{1L}'\widetilde{\beta}_1, X_{1U}'\widetilde{\beta}_1) \cap Supp(X_{1U}'\beta_1 | X_{1L}'\widetilde{\beta}_1, X_{1U}'\widetilde{\beta}_1)$. Thus, from (R1) for any $\beta_1 \neq \widetilde{\beta}_1$ in $\Theta$, there exist $\underline{c} < \overline{c}$ such that, for any $\varepsilon > 0$, if we let $0 < \varepsilon' \leq \varepsilon \wedge \overline{c} - \underline{c}$, then $P_F(X_{1Li}'\beta_1 < X_{1Uj}'\beta_1 < X_{1Li}'\beta_1 + \varepsilon' \,,\, X_{1Uj}'\widetilde{\beta}_1 \leq X_{1Li}'\widetilde{\beta}_1) > 0$. Since $\varepsilon' < \varepsilon$, the event $X_{1Li}'\beta_1 < X_{1Uj}'\beta_1 < X_{1Li}'\beta_1 + \varepsilon'$ implies $X_{1Li}'\beta_1 < X_{1Uj}'\beta_1 < X_{1Li}'\beta_1 + \varepsilon$. Since $\varepsilon > 0$ was arbitrary, The above yields $P_F\left(X_{1Li}'\beta_1 < X_{1Uj}'\beta_1 < X_{1Li}'\beta_1 + \varepsilon \,,\, X_{1Uj}'\widetilde{\beta}_1 \leq X_{1Li}'\widetilde{\beta}_1\right) > 0$ $\forall\, \beta_1, \widetilde{\beta}_1 \in \Theta$: $\beta_1 \neq \widetilde{\beta}_1$, $\forall\, \varepsilon > 0$. Therefore, equation (10) follows from the restrictions in (R1). We move on to proving equation (11). Take any $\beta_1 \in \Theta : \beta_1 \neq \beta_{10}$. By (R1) and part (i) of (R2), $\exists\, \delta > 0$ such that $P_F(X_{1Uj}'\beta_1 \leq X_{1Li}'\beta_1 \,,\, X_{1Li}'\beta_{10} < X_{1Uj}'\beta_{10} \,,\, H_F(X_{1Uj}'\beta_{10}) > H_F(X_{1Li}'\beta_{10}) + \delta) > 0$. Let $\varepsilon \equiv \delta/3$. By part (ii) of (R2), $P_F(X_{1Uj}'\beta_1 \leq X_{1Li}'\beta_1 \,,\, X_{1Li}'\beta_{10} < X_{1Uj}'\beta_{10} \,,\, H_F(X_{1Uj}'\beta_{10}) > H_F(X_{1Li}'\beta_{10} \,,\, \mu_{1F}(W_{1i} < H_F(X_{1Li}'\beta_{10}) + \varepsilon \,,\, \mu_{1F}(W_{1j}) > H_F(X_{1Uj}'\beta_{10} - \varepsilon) > 0$. Thus, $P_F\left(\mu_{1F}(W_{1i}) < \mu_{1F}(W_{1j}) \,,\, X_{1Uj}'\beta_1 \leq X_{1Li}'\beta_1\right) > 0$. Since $\beta_1 \neq \beta_{10}$ was an arbitrary element in $\Theta$, this immediately implies $P_F\left(\mu_{1F}(W_{1i}) < \mu_{1F}(W_{1j}) \,,\, X_{1Uj}'\beta_1 \leq X_{1Li}'\beta_1\right) > 0$ $\forall\, \beta_1 \in \Theta : \beta_1 \neq \beta_{10}$. Therefore, equation (11) follows from restrictions (R1) and (R2).

[†]Department of Economics, Pennsylvania State University, University Park, PA 16802, United States. Email: aaradill@psu.edu

We move on to equation (12). Fix $\varepsilon > 0$. By the Lipschitz restriction in part (i) of (R3), $\exists\ \delta > 0 : |u - u'| < \delta \ \Rightarrow\ |\lambda_F(u) - \lambda_F(u')| < \varepsilon/3$, and y restriction (R1) (and equation 10), we have $P_F(X'_{1Li}\beta_{10} - \delta < X'_{1Uj}\beta_{10} \leq X'_{1Li}\beta_{10}) > 0$. Next, by part (ii) of (R3), we also have $P_F(\lambda_F(X'_{1Li}\beta_{10}) - \varepsilon/3 < E_F[\lambda_F(X'_{1i}\beta_{10})|V_i] \leq \lambda_F(X'_{1Li}\beta_{10})\ ,\ \lambda_F(X'_{1Uj}\beta_{10}) \leq E_F[\lambda_F(X'_{1j}\beta_{10})|V_j] < \lambda_F(X'_{1Uj}\beta_{10}) + \varepsilon/3) > 0$. Since $\varepsilon > 0$ was arbitrary, it follows from here that, if restrictions (R1) and (R3) hold, we have $P_F\Big(\Big|E_F[\lambda_F(X'_{1i}\beta_0)|V_i] - E_F[\lambda_F(X'_{1j}\beta_0)|V_j]\Big| < \varepsilon\ ,\ X'_{1Uj}\beta_{10} \leq X'_{1Li}\beta_{10}\Big) > 0 \quad \forall\ \varepsilon > 0$. This proves the second part of equation (12). Now we prove the first part. By restriction (R1) and part (i) of (R3) (strict monotonicity), there exists $\varepsilon > 0$ such that $P_F(X'_{1Uj}\beta_{10} \leq X'_{1Li}\beta_1\ ,\ X'_{1Uj}\beta_{10} > X'_{1Li}\beta_{10}\ ,\ \lambda_F(X'_{1Uj}\beta_{10}) < \lambda_F(X'_{1Li}\beta_{10}) - \varepsilon) > 0$. Take any $\beta_1 \in \Theta : \beta_1 \neq \beta_{10}$. Combining the previous result with part (ii) of (R3), this means that there exists $\varepsilon > 0$ such that $P_F(X'_{1Uj}\beta_{10} \leq X'_{1Li}\beta_1\ ,\ X'_{1Uj}\beta_{10} > X'_{1Li}\beta_{10}\ ,\ \lambda_F(X'_{1Uj}\beta_{10}) < \lambda_F(X'_{1Li}\beta_{10}) - \varepsilon\ \ ,\ E_F[\lambda_F(X'_{1j}\beta_{10})|V_j] < \lambda_F(X'_{1Uj}\beta_{10}) + \varepsilon/3\ \ ,\ E_F[\lambda_F(X'_{1i}\beta_{10})|V_i] > \lambda_F(X'_{1Li}\beta_{10}) - \varepsilon/3) > 0$. Thus, $P_F(E_F[\lambda_F(X'_{1i}\beta_{10})|V_i] > E_F[\lambda_F(X'_{1j}\beta_{10})|V_j]\ ,\ X'_{1Uj}\beta_1 \leq X'_{1Li}\beta_1) > 0$. Since $\beta_1 \neq \beta_{10}$ was an arbitrary element in $\Theta$, it follows that if restrictions (R1) and (R3) hold, $P_F\Big(E_F[\lambda_F(X'_{1i}\beta_0)|V_i] > E_F[\lambda_F(X'_{1j}\beta_0)|V_j]\ ,\ X'_{1Uj}\beta_1 \leq X'_{1Li}\beta_1\Big) > 0\ \forall\ \beta_1 \in \Theta : \beta_1 \neq \beta_{10}$. This shows the first part of equation (12) and concludes the proof that both parts of this equation hold.

We move on to proving equation (13). Take any $\beta_1 \in \Theta$. From restriction (R1) and equation (10), $P_F(X'_{1Uj}\beta_1 \leq X'_{1Li}\beta_1) > 0$, and from restriction (R4), for this $\beta_1$ and any $\delta \neq 0$, $P_F(X'_{2j}\delta_2 > X'_{2i}\delta_2 | X'_{1Uj}\beta_1 \leq X'_{1Li}\beta_1) > 0$ and $P_F(X'_{2j}\delta_2 < X'_{2i}\delta_2 | X'_{1Uj}\beta_1 \leq X'_{1Li}\beta_1) > 0$. Combined, this implies that, if restrictions (R1) and (R4) hold, $P_F\Big(X'_{2j}\delta_2 > X'_{2i}\delta_2\ ,\ X'_{1Uj}\beta_1 \leq X'_{1Li}\beta_1\Big) > 0$ and $P_F\Big(X'_{2j}\delta_2 < X'_{2i}\delta_2\ ,\ X'_{1Uj}\beta_1 \leq X'_{1Li}\beta_1\Big) > 0\ \forall\ \beta_1 \in \Theta,\ \forall\ \delta_2 \neq 0$. This proves equation (13). Finally, we move on to proving equation (14). Recall first that, for any $\beta_2$, we have $\mu_{2F}(V) - X'_2\beta_2 = E_F[\lambda_F(X'_1\beta_{10})|V] + X'_2(\beta_{20} - \beta_2)$. For any $(\beta_1, \beta_2) \neq (\beta_{10}, \beta_{20})$, there are two possible cases: (i) $\beta_2 \neq \beta_{20}$ or (ii) $\beta_2 = \beta_{20}$ and $\beta_1 \neq \beta_{10}$. Let us begin with case (i). Take any $(\beta_1, \beta_2) \in \Theta$ such that $\beta_2 \neq \beta_{20}$. Combining restrictions (R1) (equation (10)), and (R3) (equation (12)) with restriction (R4), there exists $\varepsilon > 0$ such that $P_F(|E_F[\lambda_F(X'_{1i}\beta_{10})|V_i] - E_F[\lambda_F(X'_{1j}\beta_{10})|V_j]| < \varepsilon\ ,\ X'_{2i}(\beta_{20} - \beta_2) > X'_{2j}(\beta_{20} - \beta_2) + \varepsilon\ ,\ X'_{1Uj}\beta_1 \leq X'_{1Li}\beta_1) > 0$ and, therefore, $P_F(E_F[\lambda_F(X'_{1i}\beta_{10})|V_i] + X'_{2i}(\beta_{20} - \beta_2) > E_F[\lambda_F(X'_{1j}\beta_{10})|V_j] + X'_{2j}(\beta_{20} - \beta_2)\ ,\ X'_{1Uj}\beta_1 \leq X'_{1Li}\beta_1) > 0$. Now, consider case (ii) and take any $(\beta_1, \beta_{20})$ where $\beta_1 \in \Theta$ and $\beta_1 \neq \beta_{10}$. From the first part of equation (12), we immediately have $P_F\Big(E_F[\lambda_F(X'_{1i}\beta_0)|V_i] > E_F[\lambda_F(X'_{1j}\beta_0)|V_j]\ ,\ X'_{1Uj}\beta_1 \leq X'_{1Li}\beta_1\Big) > 0$. Combined, cases (i) and (ii) yield that, if restrictions (R1), (R3) and (R4) hold, then $P_F(E_F[\lambda_F(X'_{1i}\beta_{10})|V_i] + X'_{2i}(\beta_{20} - \beta_2) > E_F[\lambda_F(X'_{1j}\beta_{10})|V_j] + X'_{2j}(\beta_{20} - \beta_2)\ ,\ X'_{1Uj}\beta_1 \leq X'_{1Li}\beta_1) > 0$ for any $(\beta_1, \beta_2) \neq (\beta_{10}, \beta_{20})$. From here, since $\mu_{2F}(V) - X'_2\beta_2 = E_F[\lambda_F(X'_1\beta_{10})|V] + X'_2(\beta_{20} - \beta_2)$, we have that, if restrictions (R1), (R3) and (R4) hold, then $P_F(\mu_{2F}(V_i) - X'_{2i}\beta_2 > \mu_{2F}(V_j) - X'_{2j}\beta_2\ ,\ X'_{1Uj}\beta_1 \leq X'_{1Li}\beta_1) > 0$ for any $(\beta_1, \beta_2) \in \Theta : (\beta_1, \beta_2) \neq (\beta_{10}, \beta_{20})$. This is exactly the claim in equation (14). Therefore, we have shown that this equation follows from restrictions (R1), (R3) and (R4). This concludes the proof that the results described in equations (10), (11), (12), (13) and (14) in the text follow from restrictions (R1), (R2), (R3) and (R4). Since the statements in Result 1 are a summary of the results in these equations, this concludes the proof of Result 1 ■

### A1.1 Without an exclusion restriction between $X_2$ and $W_1$, the result in (13) cannot hold

The exclusion restriction in (R4) is a *necessary* condition for (13) to hold. The key is the following claim.

**Claim 2** *Suppose $X_2 = W_1$. Then, for any $\beta_1 \in \Theta$, there exists a $\delta_2$ such that $\delta_2' X_2 = -\beta_1' X_{1L} - \beta_1' X_{1U}$.*

**Proof:** Split our regressors in $X_1$ as $X_1 = (X_1^1, \ldots, X_1^{r_1}, X_1^{r_1+1}, \ldots, X_1^{d_1})$, where $(X_1^1, \ldots, X_1^{r_1})$ are interval-data observed, and $(X_1^{r_1+1}, \ldots, X_1^{d_1})$ are exactly observed (we can have $r_1 = d_1$, so all regressors are interval-data observed). Recall that $W_1 \equiv \underline{X}_1 \cup \overline{X}_1$, so we can express,

$$W_1 = (\underline{X}_1^1, \ldots, \underline{X}_1^{r_1}, \overline{X}_1^1, \ldots, \overline{X}_1^{r_1}, X_1^{r_1+1}, \ldots, X_1^{d_1}).$$

Take any $\beta_1 \in \Theta$ and express it accordingly as $\beta_1 = (\beta_1^1, \ldots \beta_1^{r_1}, \beta_1^{r_1+1}, \ldots, \beta_1^{d_1})$. Let

$$\delta_2 \equiv -(\beta_1^1, \ldots, \beta_1^{r_1}, \beta_1^1, \ldots, \beta_1^{r_1}, 2 \cdot \beta_1^{r_1+1}, \ldots, 2 \cdot \beta_1^{d_1}).$$

Suppose $X_2 = W_1$. Then,

$$\delta_2' X_2 = -\left( \sum_{\ell=1}^{r_1} \beta_1^\ell \underline{X}_1^\ell + \sum_{\ell=1}^{r_1} \beta_1^\ell \overline{X}_1^\ell + \sum_{\ell=r_1+1}^{d_1} 2\beta_1^\ell X_1^\ell \right) = -\beta_1' X_{1L} - \beta_1' X_{1U}. \quad ■$$

Thus, if $X_2 = W_1$, for any $\beta_1 \in \Theta$, there exists a $\delta_2$ such that $\delta_2' X_{2i} = -\beta_1' X_{1Ui} - \beta_1' X_{1Li} \leq -2\beta_1' X_{1Li}$, and $\delta_2' X_{2j} = -\beta_1' X_{1Uj} - \beta_1' X_{1Lj} \geq -2\beta_1' X_{1Uj}$. Thus, having $\beta_1' X_{1Uj} \leq \beta_1' X_{1Li}$ implies $\delta_2' X_{2j} \geq \delta_2' X_{2i}$ (since $\delta_2' X_{2j} \geq -2\beta_1' X_{1Uj} \geq -2\beta_1' X_{1Li} \geq \delta_2' X_{2i}$), so $P_F\left(X_{2j}'\delta_2 < X_{2i}'\delta_2 \, , \, X_{1Uj}'\beta_1 \leq X_{1Li}'\beta_1\right) = 0$ for this particular $\delta_2$. Also, letting $\widetilde{\delta_2} \equiv -\delta_2$, we have $P_F\left(X_{2j}'\widetilde{\delta_2} > X_{2i}'\widetilde{\delta_2} \, , \, X_{1Uj}'\beta_1 \leq X_{1Li}'\beta_1\right) = 0$. Thus, the condition in (13) cannot hold if $X_2 = W_1$. This explains the exclusion restriction in part (i) of (13).

## A2 Some alternative versions of our bivariate sample selection model

The bivariate sample selection model described in Section 3.1, which served as the foundation of the results in the paper, can be modified in various ways. Here we discuss two modifications/extensions. The first one describes the case where we have unobserved covariates in both the selection and outcome equations, with bounds that depend on observables (as in the main case we studied in the paper). The second modification discusses the truncated-data case, where our data consists only of observations where $Y_{1i}^* > 0$. In each case we discuss the pairwise functional inequalities that result, which are the equivalent versions of the inequalities in (17) in the general model we studied in Section 3.1 of the main text. Once we describe these pairwise inequalities, inference would be carried out by modifying the procedure we proposed in Section 3 accordingly.

### A2.1 A bivariate sample selection model with unobserved covariates in the selection and outcome equations

Suppose now that at least a subset of components of $X_2$ in the outcome equation are also unobserved, but that we have interval data for these covariates, so that

$$X_{2L}'\beta_{20} \leq X_2'\beta_{20} \leq X_{2U}'\beta_{20} \quad \text{w.p.1.} \tag{A-1}$$

where $(X_{2L}, X_{2U})$ are observable. We assume that the bounds in equation (19) remain valid for the selection-equation control function. Group $W_2 \equiv (X_{2L} \cup X_{2U})$, and $V \equiv (W_1, W_2)$. Suppose we have a random sample $(Y_{1i}, Y_{2i}, V_i)_{i=1}^n$ generated by $F$. Maintain the restrictions Assumption 1, modifying part (i) to the restriction, $(\varepsilon_1, \varepsilon_2) \perp (X_1, X_2, V)$. As before, let $\mu_{2F}(V) \equiv E_F[Y_2|V, Y_1 = 1]$. We now have,

$$\mu_{2F}(V) = E_F[X_2'\beta_{20}|V] + E_F[\lambda_F(g_1(X_1, \beta_{10})|V].$$

Since $\lambda_F(\cdot)$ is nonincreasing and $H_F(\cdot)$ is nondecreasing, we now have

$$X_{2L}'\beta_{20} + \lambda_F(g_{1U}(W_1, \beta_{10})) \leq \mu_{2F}(V) \leq X_{2U}'\beta_{20} + \lambda_F(g_{1L}(W_1, \beta_{10})),$$
$$H_F(g_{1L}(W_1, \beta_{10})) \leq \mu_{1F}(W_1) \leq H_F(g_{1U}(W_1, \beta_{10})).$$

Again, without further restrictions, the above bounds are sharp for the functionals involved. For a given $\beta \equiv (\beta_1, \beta_2)$, let

$$m_1(V, \beta) \equiv \begin{pmatrix} -g_{2L}(W_2, \beta_2) \\ g_{1U}(W_1, \beta_1) \end{pmatrix} \qquad m_2(V, \beta) \equiv \begin{pmatrix} -g_{2U}(W_2, \beta_2) \\ g_{1L}(W_1, \beta_1) \end{pmatrix}$$

Let $(V_i, V_j)$ be independent draws from $F$. Since $\lambda_F(\cdot)$ is nonincreasing and $H_F(\cdot)$ is nondecreasing, the model produces the following two functional inequalities,

$$\begin{aligned} \Big(\mu_{2F}(V_i) - \mu_{2F}(V_j)\Big) \mathbb{1}\Big\{m_1(V_j, \beta_0) \leq m_2(V_i, \beta_0)\Big\} &\leq 0 \quad \text{w.p.1.} \\ \Big(\mu_{1F}(W_{1j}) - \mu_{1F}(W_{1i})\Big) \cdot \mathbb{1}\Big\{g_{1U}(W_{1j}, \beta_{10}) \leq g_{1L}(W_{1i}, \beta_{10})\Big\} &\leq 0 \quad \text{w.p.1.} \end{aligned} \tag{A-2}$$

(A-2) is a modified version of the pairwise inequalities in (17). While the second inequality (corresponding to the selection equation) is identical, the outcome-equation inequality is modified. Inference would then take place by replacing (17) with (A-2) in the construction of the statistic described in Section 3.

### A2.2 A bivariate sample selection model with truncated data

Suppose we have a truncated sample generated by the bivariate sample selection model described in Section 3.1. As we did there, group $V \equiv (X_2, W_1) \in \mathbb{R}^{L_v}$. Suppose our truncated sample is $(Y_{2i}, V_i)_{i=1}^n$, where $Y_{2i} = Y_{2i}^*$ and $Y_{1i}^* > 0$ for all $i$. By the truncated nature of our data, the second inequality (corresponding to the selection equation) in (17) is no longer useful, since $Y_{1i} = 1$ for all $i$. However, the first inequality in (17) is still valid and can be used for inference. The modification of the inferential procedure described in Section 3 is straightforward, as it would simply require dropping the selection-equation inequality from the construction of our statistic.

## A3 Proof of Lemma 1

We will focus for brevity on proving part (A) of Lemma 1. The proof of part (B) follows parallel and analogous steps, so we will just summarize it towards the end. Part (C) follows immediately from (A) and (B). We begin by presenting a maximal inequality result that will be useful throughout various steps of our proofs.

### A3.1 A useful maximal inequality result

Let us begin by presenting once again the definition of Euclidean classes of functions. What follows is taken from Nolan and Pollard (1987, Definition 8), Pakes and Pollard (1989, Definition 2.7), and Sherman (1994, Definition 3).

**Definition: Euclidean classes of functions**

Let $\mathcal{T}$ be a space and $d$ be a pseudometric defined on $\mathcal{T}$. For each $\varepsilon > 0$, define the *packing number* $D(\varepsilon, d, \mathcal{T})$ to be the largest number $D$ for which there exist points $m_1, \ldots, m_D$ in $\mathcal{T}$ such that $d(m_i, m_j) > \varepsilon$ for each $i \neq j$. Packing numbers are a measure of how big $\mathcal{T}$ is with respect to $d$. Let $\mathscr{G}$ be a class of functions defined on a set $\mathcal{S}_Z^k$. We say that $G$ is an *envelope* for $\mathscr{G}$ is $\sup_{\mathscr{G}} |g(\cdot)| \leq G(\cdot)$. Let $\mu$ be a measure on $\mathcal{S}_Z^k$ and denote $\mu h \equiv \int h(z_1, \ldots, z_k) d\mu(z_1, \ldots, z_k)$. We say that the class of functions $\mathscr{G}$ is Euclidean $(A, V)$ for the envelope $G$ if, for any measure $\mu$ such that $\mu G^2 < \infty$, we have $D(\varepsilon, d_\mu, \mathscr{G}) \leq A\varepsilon^{-V}$ $\forall$ $0 < \varepsilon \leq 1$, where, for $g_1, g_2 \in \mathscr{G}$, $d_\mu(g_1, g_2) = \left(\mu|g_1 - g_2|^2/\mu G^2\right)^{1/2}$. The constants $A$ and $V$ must not depend on $\mu$. ■

The name "Euclidean" is owed to the fact that $A\varepsilon^{-V}$ is the generic expression of packing numbers for any bounded subset of the *Euclidean* space $\mathbb{R}^V$. Examples of Euclidean classes of functions can be found, in Pollard (1984), Nolan and Pollard (1987), Pakes and Pollard (1989), Pollard (1990), Sherman (1994) and Andrews (1994). Notable examples found in many econometric models include the following.

(i) (Pakes and Pollard (1989, Lemma 2.13)) Let $\mathscr{G} = \{g(\cdot,t) : t \in T\}$ be a class of functions on $\mathscr{X}$ indexed by a bounded subset $T$ of $\mathbb{R}^d$. If there exists an $\alpha > 0$ and a $\phi(\cdot) \geq 0$ such that $|g(x,t) - g(x,t')| \leq \phi(x) \cdot \|t - t'\|^{\alpha}$ for $x \in \mathscr{X}$ and $t, t' \in T$. Then $\mathscr{G}$ is Euclidean for the envelope $G \equiv |g(\cdot,t_0)| + M\phi(\cdot)$, where $t_0 \in T$ is an arbitrary point and $M \equiv (2\sqrt{d}\sup_T \|t - t_0\|)^{\alpha}$.

(ii) (Nolan and Pollard (1987, Lemma 22), Pakes and Pollard (1989, Example 10)) Let $\lambda(\cdot)$ be a real-valued function of bounded variation on $\mathbb{R}$. The class $\mathscr{G}$ of all functions on $\mathbb{R}^d$ of the form $x \to \lambda(\alpha' x + \beta)$, with $\alpha$ ranging over $\mathbb{R}^d$ and $\beta$ ranging over $\mathbb{R}$ is Euclidean for the constant envelope $G \equiv \sup|\lambda|$.

(iii) (Pakes and Pollard (1989, p. 1033)) Classes of indicator functions over VC classes of sets are Euclidean for the constant envelope 1.

(iv) Type I, II and III classes of functions described in Andrews (1994) are special cases of Euclidean classes.

From the above examples, it follows from Assumptions 2 and 3 (compactness of $\Theta$ and the restriction that $E_F[\|X_4\|] \leq \overline{C}_4$ for all $F \in \mathcal{F}$), that the class of functions

$$\mathscr{G}_2 \equiv \{m(x_2) = x_2'\beta_2 \text{ for some } \beta_2 \in \Theta\}$$

is Euclidean. Pointwise algebraic operations such as products, linear combinations, minima and maxima allow us to combine Euclidean classes and preserve the Euclidean property (see Pakes and Pollard (1989, Lemma 2.14)). Empirical processes and U-processes produced by Euclidean classes of functions satisfy the *Pollard's entropy condition* (see Andrews (1994, Definition 4.2)) and *manageability* (see Pollard (1990, Definition 7.9), Andrews (1994, Equation A.1)).

### A3.1.1 A maximal inequality for degenerate U-processes

The following result is taken from Sherman (1994), who obtained maximal inequalities for degenerate U-Processes. Let $Z_1, \ldots, Z_n$ be i.i.d observations from a distribution $F$ on a set $\mathcal{S}_Z$. Let $k$ be a positive integer and $\mathscr{G}$ a class of real-valued functions on $\mathcal{S}_Z^k = \mathcal{S}_Z \otimes \cdots \otimes \mathcal{S}_Z$ (k factors). For each $g \in \mathscr{G}$, define

$$U_n^k g = (n)_k^{-1} \sum_{\boldsymbol{i}_k} g(Z_{i_1}, \ldots, Z_{i_k}),$$

where $(n)_k = n(n-1)\cdots(n-k+1)$ and $\sum_{\boldsymbol{i}_k}$ denotes the sum over the $(n)_k$ distinct integers $\{i_1, \ldots, i_k\}$ from the set $\{1, \ldots, n\}$. $U_n^k g$ is a U-statistic of order $k$ and the collection $\{U_n^k g \colon g \in \mathscr{G}\}$ is called a

U-process of order $k$, indexed by $\mathscr{G}$. If every $g \in \mathscr{G}$ is such that

$$\underbrace{E_F\left[g(s_1,\ldots,s_{i-1},Z,s_{i+1},\ldots,s_k)\right] \equiv 0}_{E_F[g(Z_1,\ldots,Z_k)|Z_1=s_1,\ldots,Z_{i-1}=s_{i-1},Z_{i+1}=s_{i+1},\ldots,Z_k=s_k]\equiv 0} \quad , \quad i = 1,\ldots,k,$$

then $\mathscr{G}$ is called an $F-$*degenerate* class of functions on $\mathcal{S}_Z^k$ and $\{U_n^k g \colon g \in \mathscr{G}\}$ is a *degenerate U-process* of order $k$.

**Result A1** ***(Sherman (1994, Corollary 4A))*** *Let $\mathscr{G}$ be a class of $F-$degenerate functions on $\mathcal{S}_Z^k$, $k \geq 1$. Suppose $\mathscr{G}$ is Euclidean $(A,V)$ for an envelope $G$ such that $E_F\left[G(Z_1,\ldots,Z_k)^{4p}\right] < \infty$ for a positive integer $p$. Then,*

$$E_F\left[\left(\sup_{\mathscr{G}}\left|n^{k/2}U_n^k g\right|\right)^p\right] \leq \Upsilon \cdot \left(E_F\left[G(Z_1,\ldots,Z_k)^{4p}\right]\right)^{1/2} \equiv \overline{M},$$

*where $\Upsilon$ is a constant that depends only on $p$, $A$, $V$ and $E_F\left[G(Z_1,\ldots,Z_k)^2\right]$. By a Chebyshev inequality, this implies that for each $\varepsilon > 0$,*

$$P_F\left(\sup_{\mathscr{G}}\left|n^{k/2}U_n^k g\right| > \varepsilon\right) \leq \frac{\overline{M}}{\varepsilon^p} \quad \text{and therefore} \quad P_F\left(\sup_{\mathscr{G}}\left|U_n^k g\right| > \varepsilon\right) \leq \frac{\overline{M}}{\left(n^{k/2}\cdot\varepsilon\right)^p}.$$

*From the last result, we also have*

$$\sup_{\mathscr{G}}\left|U_n^k g\right| = O_p\left(\frac{1}{n^{k/2}}\right). \quad \blacksquare$$

We will invoke Result A1 at various points throughout our proofs.

#### A3.1.2 VC classes of sets and Assumption 3

VC classes of sets are defined, e.g, in Pakes and Pollard (1989, Definition 2.2) and Kosorok (2008, Section 9.1.1). Verifiable criteria that suffice for a class of sets to have the VC property can be found, e.g, in Pollard (1984, Section II.4), Dudley (1984, Section 9), or Kosorok (2008, Section 9.1.1). An example commonly encountered in econometric models (Pakes and Pollard (1989, Lemma 2.4) is the class of sets of the form $\{g \geq t\}$ or $\{g > t\}$, with $g \in \mathscr{G}$ and $t \in \mathbb{R}$, where $\mathscr{G}$ is a finite dimensional vector space of real-valued functions. This class encompasses econometric models where the parameters of interest enter through linear indices. Combining VC classes of sets through a finite number of Boolean operations (e.g, unions, intersections and/or complements) preserves the VC property (Pakes and Pollard (1989, Lemma 2.5)). Assumption 3 implies that the following is a VC class of sets for each $F$, with VC dimension uniformly bounded over $\mathcal{F}$

by a finite constant $\overline{V}_D$,

$$\mathscr{D}^{\tau_2}_{1,F} \equiv \Big\{(v_1,v_2)\in\mathbb{R}^{L_v}\times\mathbb{R}^{L_v} : \tau_{2F}(v_1,v_2,\beta)\geq 0 \;\text{ for some } \beta\in\Theta\Big\}$$

And, by VC-preserving properties of Boolean operations described, e.g, in Pakes and Pollard (1989, Lemma 2.5), Assumption 3 implies that, for each $F\in\mathcal{F}$, the following class of sets is also a VC class, with VC dimension uniformly bounded over $\mathcal{F}$ by a finite constant,

$$\mathscr{D}^{\tau_2}_{2,F} \equiv \Big\{(v_1,v_2)\in\mathbb{R}^{L_v}\times\mathbb{R}^{L_v} : -c\leq\tau_{2F}(v_1,v_2,\beta)<0 \;\text{ for some } 0<c\leq c_0 \text{ and } \beta\in\Theta\Big\}.$$

Indicator functions for these classes of sets are relevant in our problem. The VC properties in Assumption 3 will lead us to invoke the maximal inequality properties in Result A1 since indicator functions over VC classes of sets are Euclidean classes of functions (Pakes and Pollard (1989, p. 1033)).

### A3.2 Asymptotic properties of $\widehat{Q}_2$ and $\widehat{R}_2$

**Note:** In all the results that follow, $\epsilon > 0$ denotes the constant described in Assumption 4 of the paper.

Recall that, as described in equation (27) in the paper, for a given $v\equiv(v^c,v^d)$, we defined,

$$\mathcal{K}\left(\frac{V_i^c-v^c}{h_n}\right)\equiv\prod_{m=1}^{r}\kappa\left(\frac{V_{mi}^c-v_m^c}{h_n}\right),\quad \Gamma(V_i,v,h_n)\equiv\mathcal{K}\left(\frac{V_i^c-v^c}{h_n}\right)\cdot\mathbb{1}\Big\{V_i^d=v^d\Big\},$$

and, from here,

$$\widehat{R}_2(v)\equiv\frac{1}{n\cdot h_n^r}\sum_{i=1}^{n}Y_{2i}Y_{1i}\phi_2(V_i)\Gamma(V_i,v,h_n),\quad \widehat{Q}_2(v)\equiv\frac{1}{n\cdot h_n^r}\sum_{i=1}^{n}Y_{1i}\phi_2(V_i)\Gamma(V_i,v,h_n).$$

We proceed next to characterize the asymptotic properties of $\widehat{R}_2(v)$ and $\widehat{Q}_2(v)$ under our assumptions. Let $\lambda(\cdot)$ be a real-valued function of bounded variation on $\mathbb{R}$. By Nolan and Pollard (1987, Lemma 22) (or Pakes and Pollard (1989, Example 10)), the class $\mathscr{G}$ of all functions on $\mathbb{R}^d$ of the form $x\to\lambda(\alpha'x+\beta)$, with $\alpha$ ranging over $\mathbb{R}^d$ and $\beta$ ranging over $\mathbb{R}$ is Euclidean for the constant envelope $G\equiv\sup|\lambda|$. Therefore, since our kernel is a function of bounded variation, the class of functions $\Big\{m(v)=k\left(\frac{v-u}{h}\right)$ for some $u\in\mathbb{R}, h>0\Big\}$ is Euclidean $(A_k,V_k)$ for the constant envelope $\overline{k}$ (neither $(A_k,V_k)$, nor $\overline{k}$ depend on $F$). From here and Sherman (1994, Lemma 5), the following empirical processes $\nu_n^{Q_2}(\cdot)$ and $\nu_n^{R_2}(\cdot)$ defined as follows, satisfy the conditions of Result A1,

$$\begin{aligned}&\left\{\nu_n^{Q_2}(v,h)=\frac{1}{n}\sum_{i=1}^{n}\left(Y_{1i}\phi_2(V_i)\Gamma(V_i,v,h)-E_F\left[Y_1\phi_2(V)\Gamma(V,v,h)\right]\right)\,:\,v\in\mathbb{R}^{L_V}\ ,\,h>0\right\},\\&\left\{\nu_n^{R_2}(v,h)=\frac{1}{n}\sum_{i=1}^{n}\left(Y_{2i}Y_{1i}\phi_2(V_i)\Gamma(V_i,v,h)-E_F\left[Y_2Y_1\phi_2(V)\Gamma(V,v,h)\right]\right)\,:\,v\in\mathbb{R}^{L_V}\ ,\,h>0\right\}\end{aligned} \tag{A-3}$$

for the constant envelope $\overline{\phi}\cdot\overline{K}$, and the envelope $|Y_2|\cdot\overline{\phi}\cdot\overline{K}$, respectively. From here, Result A1 and the condition that $E_F[|Y_2|^4]\leq\overline{D}_2$ for all $F\in\mathcal{F}$ (Assumption 3) imply that there exists a finite $\overline{M}$ such that, for each $\varepsilon>0$,

$$P_F\left(\sup_{\substack{v\in\mathbb{R}^{L_v}\\h>0}}\left|\nu_n^{Q_2}(v,h)\right|>\varepsilon\right)\leq\frac{\overline{M}}{n^{1/2}\cdot\varepsilon},\quad\text{and}\quad P_F\left(\sup_{\substack{v\in\mathbb{R}^{L_v}\\h>0}}\left|\nu_n^{R_2}(v,h)\right|>\varepsilon\right)\leq\frac{\overline{M}}{n^{1/2}\cdot\varepsilon}\quad\forall\,F\in\mathcal{F} \tag{A-4}$$

and therefore,

$$\sup_{\substack{v\in\mathbb{R}^{L_v}\\h>0}}\left|\nu_n^{Q_2}(v,h)\right|=O_p\left(\frac{1}{n^{1/2}}\right),\quad\text{and}\quad\sup_{\substack{v\in\mathbb{R}^{L_v}\\h>0}}\left|\nu_n^{R_2}(v,h)\right|=O_p\left(\frac{1}{n^{1/2}}\right),\quad\text{uniformly over }\mathcal{F} \tag{A-5}$$

We have,

$$\begin{aligned}\widehat{Q}_2(v)-Q_{2F}(v)&=\frac{1}{h_n^{r_v}}\cdot\nu_n^{Q_2}(v,h_n)+B_{n,F}^{Q_2}(v),\quad\text{where}\quad B_{n,F}^{Q_2}(v)\equiv\frac{1}{h_n^{r_v}}\cdot\Big(Q_{2F}(v)-E_F\left[Y_1\phi_2(V)\Gamma(V,v,h_n)\right]\Big),\\\widehat{R}_2(v)-R_{2F}(v)&=\frac{1}{h_n^{r_v}}\cdot\nu_n^{R_2}(v,h_n)+B_{n,F}^{R_2}(v),\quad\text{where}\quad B_{n,F}^{R_2}(v)\equiv\frac{1}{h_n^{r_v}}\cdot\Big(R_{2F}(v)-E_F\left[Y_2Y_1\phi_2(V)\Gamma(V,v,h_n)\right]\Big),\end{aligned} \tag{A-6}$$

The smoothness conditions in Assumption 2 and the kernel properties in Assumption 4 an $M^{th}$–order approximation implies that there exists a finite $\overline{B}$ such that

$$\sup_{v\in\mathcal{V}}\left|B_{n,F}^{Q_2}(v)\right|\leq\overline{B}\cdot h_n^M,\quad\text{and}\quad\sup_{v\in\mathcal{V}}\left|B_{n,F}^{R_2}(v)\right|\leq\overline{B}\cdot h_n^M\quad\forall\,F\in\mathcal{F} \tag{A-7}$$

From (A-6) and (A-7) we have,

$$\left.\begin{aligned}\sup_{v\in\mathcal{V}}\left|\widehat{Q}_2(v)-Q_{2F}(v)\right|&\leq\tfrac{1}{h_n^{r_v}}\cdot\sup_{v\in\mathcal{V}}\left|\nu_n^{Q_2}(v,h_n)\right|+\overline{B}\cdot h_n^M\\\sup_{v\in\mathcal{V}}\left|\widehat{R}_2(v)-R_{2F}(v)\right|&\leq\tfrac{1}{h_n^{r_v}}\cdot\sup_{v\in\mathcal{V}}\left|\nu_n^{R_2}(v,h_n)\right|+\overline{B}\cdot h_n^M\end{aligned}\right\}\ \forall\,F\in\mathcal{F} \tag{A-8}$$

From (A-5) and (A-8), and the bandwidth convergence restrictions in Assumption 4, we have

$$\left.\begin{aligned}\sup_{v\in\mathcal{V}}\left|\widehat{Q}_2(v)-Q_{2F}(v)\right| &= O_p\left(\tfrac{1}{h_n^{r_v}\cdot n^{1/2}}\right)+\overline{B}\cdot h_n^M = o_p\left(\tfrac{1}{n^{1/4+\epsilon/2}}\right)\\ \sup_{v\in\mathcal{V}}\left|\widehat{R}_2(v)-R_{2F}(v)\right| &= O_p\left(\tfrac{1}{h_n^{r_v}\cdot n^{1/2}}\right)+\overline{B}\cdot h_n^M = o_p\left(\tfrac{1}{n^{1/4+\epsilon/2}}\right)\end{aligned}\right\}\text{ uniformly over }\mathcal{F} \tag{A-9}$$

Where $\epsilon > 0$ denotes the constant described in Assumption 4. Take any sequence $\varepsilon_n > 0$ such that $n^{1/2}\cdot h_n^{r_v}\cdot\varepsilon_n \longrightarrow \infty$. Given the bandwidth convergence restrictions in Assumption 4, there exists $n_0 > 0$ such that $n^{1/2}\cdot h_n^{r_v}\cdot\varepsilon - \overline{B}\cdot n^{1/2}\cdot h_n^{r_v+M} > 0$, for all $n > n_0$, and from the results in (A-4) and (A-8),

$$\left.\begin{aligned}P_F\left(\sup_{v\in\mathcal{V}}\left|\widehat{Q}_2(v)-Q_{2F}(v)\right|>\varepsilon_n\right) &\le \tfrac{\overline{M}}{n^{1/2}\cdot h_n^{r_v}\cdot\varepsilon_n-\overline{B}\cdot n^{1/2}\cdot h_n^{r_v+M}}\\ P_F\left(\sup_{v\in\mathcal{V}}\left|\widehat{R}_2(v)-R_{2F}(v)\right|>\varepsilon_n\right) &\le \tfrac{\overline{M}}{n^{1/2}\cdot h_n^{r_v}\cdot\varepsilon_n-\overline{B}\cdot n^{1/2}\cdot h_n^{r_v+M}}\end{aligned}\right\}\ \forall\ F\in\mathcal{F},\ \forall\ n>n_0 \tag{A-10}$$

Therefore, under the conditions in Assumptions 2 and 4, we have

$$\left.\begin{aligned}\sup_{F\in\mathcal{F}} P_F\left(\sup_{v\in\mathcal{V}}\left|\widehat{Q}_2(v)-Q_{2F}(v)\right|>\varepsilon_n\right)\longrightarrow 0\\ \sup_{F\in\mathcal{F}} P_F\left(\sup_{v\in\mathcal{V}}\left|\widehat{R}_2(v)-R_{2F}(v)\right|>\varepsilon_n\right)\longrightarrow 0\end{aligned}\right\}\ \forall\ \varepsilon_n>0:\ n^{1/2}\cdot h_n^{r_v}\cdot\varepsilon_n\longrightarrow\infty$$

### A3.3 Asymptotic properties of $\widehat{\tau}_2(v,\widetilde{v},\beta)$

We have defined,

$$\begin{aligned}&\widehat{\tau}_2(v,\widetilde{v},\beta) =\\ &\left(\left(\widehat{R}_2(v)\widehat{Q}_2(\widetilde{v})-\widehat{R}_2(\widetilde{v})\widehat{Q}_2(v)\right)-(x_2'\beta_2-\widetilde{x}_2'\beta_2)\,\widehat{Q}_2(v)\widehat{Q}_2(\widetilde{v})\right)\cdot\mathbb{1}\{g_{1U}(\widetilde{w}_1,\beta_1)\le g_{1L}(w_1,\beta_1)\}\\ &\cdot\phi_2(v)\phi_2(\widetilde{v}),\\ &\tau_{2F}(v,\widetilde{v},\beta) =\\ &\left(\left(R_{2F}(v)Q_{2F}(\widetilde{v})-R_{2F}(\widetilde{v})Q_{2F}(v)\right)-(x_2'\beta_2-\widetilde{x}_2'\beta_2)\,Q_{2F}(v)Q_{2F}(\widetilde{v})\right)\cdot\mathbb{1}\{g_{1U}(\widetilde{w}_1,\beta_1)\le g_{1L}(w_1,\beta_1)\}\\ &\cdot\phi_2(v)\phi_2(\widetilde{v}),\end{aligned}$$

Therefore,

$$\widehat{\tau}_2(v,\widetilde{v},\beta)-\tau_{2F}(v,\widetilde{v},\beta)=$$
$$\Big(\Big(R_{2F}(v)-(x_2'\beta_2-\widetilde{x}_2'\beta_2)\,Q_{2F}(v)\Big)\cdot\Big(\widehat{Q}_2(\widetilde{v})-Q_{2F}(\widetilde{v})\Big)-\Big(R_{2F}(\widetilde{v})+(x_2'\beta_2-\widetilde{x}_2'\beta_2)\,Q_{2F}(\widetilde{v})\Big)$$
$$\cdot\Big(\widehat{Q}_2(v)-Q_{2F}(v)\Big)$$
$$+Q_{2F}(\widetilde{v})\cdot\Big(\widehat{R}_2(v)-R_{2F}(v)\Big)-Q_{2F}(v)\cdot\Big(\widehat{R}_2(\widetilde{v})-R_{2F}(\widetilde{v})\Big)\Big)\cdot\mathbb{1}\{g_{1U}(\widetilde{w}_1,\beta_1)\le g_{1L}(w_1,\beta_1)\}\,\phi_2(v)\phi_2(\widetilde{v})$$
$$+\xi_{a,n}^{\tau_2}(v,\widetilde{v},\beta),\tag{A-11}$$

where

$$\xi_{a,n}^{\tau_2}(v,\widetilde{v},\beta)\equiv\Big(\Big(\widehat{R}_2(v)-R_{2F}(v)\Big)\cdot\Big(\widehat{Q}_2(\widetilde{v})-Q_{2F}(\widetilde{v})\Big)-\Big(\widehat{R}_2(\widetilde{v})-R_{2F}(\widetilde{v})\Big)\cdot\Big(\widehat{Q}_2(v)-Q_{2F}(v)\Big)$$
$$-(x_2'\beta_2-\widetilde{x}_2'\beta_2)\cdot\Big(\widehat{Q}_2(v)-Q_{2F}(v)\Big)\cdot\Big(\widehat{Q}_2(\widetilde{v})-Q_{2F}(\widetilde{v})\Big)\Big)\cdot\mathbb{1}\{g_{1U}(\widetilde{w}_1,\beta_1)\le g_{1L}(w_1,\beta_1)\}\,\phi_2(v)\phi_2(\widetilde{v})$$

From the conditions in Assumption 2, there exists a finite constant $\overline{D}$ such that

$$\sup_{\substack{v,\widetilde{v}\in\mathbb{R}^{L_V}\times\mathbb{R}^{L_V}\\ \beta\in\Theta}}\max\Big\{|R_{2F}(v)|,\ |Q_{2F}(v)|,\ \Big|x_2'\beta_2-\widetilde{x}_2'\beta_2\Big|\Big\}\le 2\overline{D}\quad\forall\,F\in\mathcal{F}.$$

Therefore, there exists $\overline{D}_2$ such that, for each $F\in\mathcal{F}$,

$$\sup_{\substack{v,\widetilde{v}\in\mathbb{R}^{L_V}\times\mathbb{R}^{L_V}\\ \beta\in\Theta}}\Big|\widehat{\tau}_2(v,\widetilde{v},\beta)-\tau_{2F}(v,\widetilde{v},\beta)\Big|\le\overline{D}_2\cdot\left(\sup_{v\in\mathcal{V}}\Big|\widehat{Q}_2(v)-Q_{2F}(v)\Big|+\sup_{v\in\mathcal{V}}\Big|\widehat{R}_2(v)-R_{2F}(v)\Big|\right.$$
$$\left.+\sup_{v\in\mathcal{V}}\Big|\widehat{Q}_2(v)-Q_{2F}(v)\Big|\times\sup_{v\in\mathcal{V}}\Big|\widehat{R}_2(v)-R_{2F}(v)\Big|+\left(\sup_{v\in\mathcal{V}}\Big|\widehat{Q}_2(v)-Q_{2F}(v)\Big|\right)^2\right)$$

Therefore, there exists a finite constant $\overline{C}_2$ such that, for any $b>0$

$$P_F\left(\sup_{\substack{v,\widetilde{v}\in\mathbb{R}^{L_V}\times\mathbb{R}^{L_V}\\ \beta\in\Theta}}\Big|\widehat{\tau}_2(v,\widetilde{v},\beta)-\tau_{2F}(v,\widetilde{v},\beta)\Big|>b\right)\le P_F\left(\sup_{v\in\mathcal{V}}\Big|\widehat{Q}_2(v)-Q_{2F}(v)\Big|>\overline{C}_2\cdot\Big(b\wedge b^{1/2}\Big)\right)$$
$$+P_F\left(\sup_{v\in\mathcal{V}}\Big|\widehat{R}_2(v)-R_{2F}(v)\Big|>\overline{C}_2\cdot\Big(b\wedge b^{1/2}\Big)\right)\quad\forall\,F\in\mathcal{F}$$

Fix $b > 0$. From the previous result, equation (A-10) implies that, under Assumptions 2, 3 and 4, there exist constants $\overline{M}$, $\overline{B}$ and $\overline{C}_2$ and $n_0$ such that, for $n > n_0$,

$$P_F\left(\sup_{\substack{v,\widetilde{v}\in\mathbb{R}^{L_V}\times\mathbb{R}^{L_V}\\ \beta\in\Theta}}\left|\widehat{\tau}_2(v,\widetilde{v},\beta)-\tau_{2F}(v,\widetilde{v},\beta)\right|>b\right)\leq\frac{2\overline{M}}{n^{1/2}\cdot h_n^{r_v}\cdot\overline{C}_2\cdot\left(b\wedge b^{1/2}\right)-\overline{B}\cdot n^{1/2}\cdot h_n^{r_v+M}}\quad\forall\ F\in\mathcal{F}$$

In particular, take any sequence $b_n > 0$ such that $b_n \longrightarrow 0$ and $n^{1/2}\cdot h_n^r\cdot b_n\longrightarrow\infty$. The previous result implies that, under Assumptions 2, 3 and 4, for any such sequence, we have,

$$\sup_{F\in\mathcal{F}}P_F\left(\sup_{\substack{v,\widetilde{v}\in\mathbb{R}^{L_V}\times\mathbb{R}^{L_V}\\ \beta\in\Theta}}\left|\widehat{\tau}_2(v,\widetilde{v},\beta)-\tau_{2F}(v,\widetilde{v},\beta)\right|>b_n\right)\longrightarrow 0. \tag{A-12}$$

Note that (A-12) immediately implies,

$$\sup_{\substack{v,\widetilde{v}\in\mathbb{R}^{L_V}\times\mathbb{R}^{L_V}\\ \beta\in\Theta}}\left|\widehat{\tau}_2(v,\widetilde{v},\beta)-\tau_{2F}(v,\widetilde{v},\beta)\right|=o_p(1),\quad\text{uniformly over }\mathcal{F}, \tag{A-13}$$

and

$$\sup_{\substack{v,\widetilde{v}\in\mathbb{R}^{L_V}\times\mathbb{R}^{L_V}\\ \beta\in\Theta}}\mathbb{1}\left\{\left|\widehat{\tau}_2(v,\widetilde{v},\beta)-\tau_{2F}(v,\widetilde{v},\beta)\right|>b_n\right\}=o_p(1),\quad\text{uniformly over }\mathcal{F}. \tag{A-14}$$

Next, note that

$$\begin{aligned}
&\left|\mathbb{1}\left\{\widehat{\tau}_2(v,\widetilde{v},\beta)\geq -b_n\right\}-\mathbb{1}\left\{\tau_{2F}(v,\widetilde{v},\beta)\geq 0\right\}\right|\\
=&\mathbb{1}\left\{\widehat{\tau}_2(v,\widetilde{v},\beta)\geq -b_n\ ,\ -2b_n\leq\tau_{2F}(v,\widetilde{v},\beta)<0\right\}+\mathbb{1}\left\{\widehat{\tau}_2(v,\widetilde{v},\beta)\geq -b_n\ ,\ \tau_{2F}(v,\widetilde{v},\beta)<-2b_n\right\}\\
+&\mathbb{1}\left\{\widehat{\tau}_2(v,\widetilde{v},\beta)<-b_n\ ,\ \tau_{2F}(v,\widetilde{v},\beta)\geq 0\right\}\\
\leq&\mathbb{1}\left\{-2b_n\leq\tau_{2F}(v,\widetilde{v},\beta)<0\right\}+\mathbb{1}\left\{\left|\widehat{\tau}_2(v,\widetilde{v},\beta)-\tau_{2F}(v,\widetilde{v},\beta)\right|\geq b_n\right\}
\end{aligned} \tag{A-15}$$

And, by the conditions in Assumption 2, there exists a finite constant $\overline{\tau}_2$ such that

$$\sup_{\substack{v,\widetilde{v}\in\mathbb{R}^{L_V}\times\mathbb{R}^{L_V}\\ \beta\in\Theta}}\left|\tau_{2F}(v,\widetilde{v},\beta)\right|\leq\overline{\tau}_2\quad\forall\ F\in\mathcal{F}. \tag{A-16}$$

We have

$$\widehat{\tau}_2(v,\widetilde{v},\beta)\cdot\mathbb{1}\{\widehat{\tau}_2(v,\widetilde{v},\beta)\ge -b_n\} = (\tau_{2F}(v,\widetilde{v},\beta))_+ + \tau_{2F}(v,\widetilde{v},\beta)\cdot\Big(\mathbb{1}\{\widehat{\tau}_2(v,\widetilde{v},\beta)\ge -b_n\} - \mathbb{1}\{\tau_{2F}(v,\widetilde{v},\beta)\ge 0\}\Big) + (\widehat{\tau}_2(v,\widetilde{v},\beta)-\tau_{2F}(v,\widetilde{v},\beta))\cdot\mathbb{1}\{\widehat{\tau}_2(v,\widetilde{v},\beta)\ge -b_n\}.$$

From here, using the results in (A-15) and (A-16),

$$\begin{aligned}
&\sup_{\substack{v,\widetilde{v}\in\mathbb{R}^{L_V}\times\mathbb{R}^{L_V}\\ \beta\in\Theta}} \left|\widehat{\tau}_2(v,\widetilde{v},\beta)\cdot\mathbb{1}\{\widehat{\tau}_2(v,\widetilde{v},\beta)\ge -b_n\} - (\tau_{2F}(v,\widetilde{v},\beta))_+\right| \\
&\le \overline{\tau}_2\cdot \underbrace{\sup_{\substack{v,\widetilde{v}\in\mathbb{R}^{L_V}\times\mathbb{R}^{L_V}\\ \beta\in\Theta}} \mathbb{1}\left\{\left|\widehat{\tau}_2(v,\widetilde{v},\beta)-\tau_{2F}(v,\widetilde{v},\beta)\right| > b_n\right\}}_{o_p(1)\text{ uniformly over }\mathcal{F}\text{, by (A-14)}} \\
&+ \underbrace{\sup_{\substack{v,\widetilde{v}\in\mathbb{R}^{L_V}\times\mathbb{R}^{L_V}\\ \beta\in\Theta}} \Big(|\tau_{2F}(v,\widetilde{v},\beta)|\cdot\mathbb{1}\{-2b_n\le\tau_{2F}(v,\widetilde{v},\beta)<0\}\Big)}_{\le 2b_n\to 0\text{ for all }F\text{, by construction}} + \underbrace{\sup_{\substack{v,\widetilde{v}\in\mathbb{R}^{L_V}\times\mathbb{R}^{L_V}\\ \beta\in\Theta}} \left|\widehat{\tau}_2(v,\widetilde{v},\beta)-\tau_{2F}(v,\widetilde{v},\beta)\right|}_{o_p(1)\text{ uniformly over }\mathcal{F}\text{, by (A-13)}}
\end{aligned}$$

Therefore,

$$\sup_{\substack{v,\widetilde{v}\in\mathbb{R}^{L_V}\times\mathbb{R}^{L_V}\\ \beta\in\Theta}} \left|\widehat{\tau}_2(v,\widetilde{v},\beta)\cdot\mathbb{1}\{\widehat{\tau}_2(v,\widetilde{v},\beta)\ge -b_n\} - (\tau_{2F}(v,\widetilde{v},\beta))_+\right| = o_p(1), \quad \text{uniformly over } \mathcal{F}. \tag{A-17}$$

And from the definition of $\widehat{\mathcal{T}_2}(\beta)$ in equation (28), the result in (A-17) immediately implies,

$$\sup_{\beta\in\Theta}\left|\widehat{\mathcal{T}_2}(\beta)-\mathcal{T}_{2F}(\beta)\right| = o_p(1), \quad \text{uniformly over } \mathcal{F}. \tag{A-18}$$

Let us go back to (A-11). Plugging in (A-6) into (A-11), we have,

$$\begin{aligned}
&\widehat{\tau}_2(v,\widetilde{v},\beta)-\tau_{2F}(v,\widetilde{v},\beta) = \\
&\Big[\Big(R_{2F}(v)-(x_2'\beta_2-\widetilde{x}_2'\beta_2)\,Q_{2F}(v)\Big)\cdot\frac{1}{h_n^r}\cdot\nu_n^{Q_2}(v',h_n) \\
&-\Big(R_{2F}(\widetilde{v})+(x_2'\beta_2-\widetilde{x}_2'\beta_2)\,Q_{2F}(\widetilde{v})\Big)\cdot\frac{1}{h_n^r}\cdot\nu_n^{Q_2}(v,h_n) \\
&+Q_{2F}(\widetilde{v})\cdot\frac{1}{h_n^r}\cdot\nu_n^{R_2}(v,h_n)-Q_{2F}(v)\cdot\frac{1}{h_n^r}\cdot\nu_n^{R_2}(v',h_n)\Big]\cdot\mathbb{1}\{g_{1U}(\widetilde{w}_1,\beta_1)\le g_{1L}(w_1,\beta_1)\}\,\phi_2(v)\phi_2(\widetilde{v}) \\
&+\xi_{a,n}^{\tau_2}(v,\widetilde{v},\beta)+\xi_{b,n}^{\tau_2}(v,\widetilde{v},\beta),
\end{aligned} \tag{A-19}$$

where

$$\xi^{\tau_2}_{b,n}(v,\widetilde{v},\beta) \equiv \Big[\Big(R_{2F}(v) - (x_2'\beta_2 - \widetilde{x}_2'\beta_2)\,Q_{2F}(v)\Big)\cdot B^{Q_2}_{n,F}(\widetilde{v})$$
$$-\Big(R_{2F}(\widetilde{v}) + (x_2'\beta_2 - \widetilde{x}_2'\beta_2)\,Q_{2F}(\widetilde{v})\Big)\cdot B^{Q_2}_{n,F}(v)$$
$$+\,Q_{2F}(\widetilde{v})\cdot B^{R_2}_{n,F}(v) - Q_{2F}(v)\cdot B^{R_2}_{n,F}(\widetilde{v})\Big]\cdot \mathbb{1}\{g_{1U}(\widetilde{w}_1,\beta_1) \le g_{1L}(w_1,\beta_1)\}\,\phi_2(v)\phi_2(\widetilde{v})$$

By the conditions in Assumption 2 and the result in (A-7), there exist finite constants $\overline{D}_2$ and $\overline{B}$ such that,

$$\sup_{\substack{v,\widetilde{v}\in\mathbb{R}^{L_V}\times\mathbb{R}^{L_V}\\ \beta\in\Theta}} \left|\xi^{\tau_2}_{b,n}(v,\widetilde{v},\beta)\right| \le \overline{D}_2\cdot\left(\sup_{v\in\mathcal{V}}\left|B^{Q_2}_{n,F}(v)\right| + \sup_{v\in\mathcal{V}}\left|B^{Q_2}_{n,F}(v)\right|\right) \le 2\cdot\overline{D}_2\cdot\overline{B}\cdot h_n^M$$
$$\equiv \overline{B}_3\cdot h_n^M = o\left(\frac{1}{n^{1/2+\epsilon}}\right) \quad \forall\ F\in\mathcal{F} \tag{A-20}$$

where the last equality follows from Assumption 4, and $\epsilon > 0$ is the constant described there. Next we turn our attention to $\xi^{\tau_2}_{a,n}(v,\widetilde{v},\beta)$. By the conditions in Assumption 2, there exists a finite constant $\overline{D}$ such that,

$$\sup_{\substack{v,\widetilde{v}\in\mathbb{R}^{L_V}\times\mathbb{R}^{L_V}\\ \beta\in\Theta}} \left|\xi^{\tau_2}_{a,n}(v,\widetilde{v},\beta)\right| \le 2\cdot\sup_{v\in\mathcal{V}}\left|\widehat{Q}_2(v) - Q_{2F}(v)\right| \times \sup_{v\in\mathcal{V}}\left|\widehat{R}_2(v) - R_{2F}(v)\right| + \overline{D}\cdot\left(\sup_{v\in\mathcal{V}}\left|\widehat{Q}_2(v) - Q_{2F}(v)\right|\right)^2,$$

for all $F\in\mathcal{F}$. And from here, the result in (A-9) yields,

$$\sup_{\substack{v,\widetilde{v}\in\mathbb{R}^{L_V}\times\mathbb{R}^{L_V}\\ \beta\in\Theta}} \left|\xi^{\tau_2}_{a,n}(v,\widetilde{v},\beta)\right| = \left[o_p\left(\frac{1}{n^{1/4+\epsilon/2}}\right)\right]^2 = o_p\left(\frac{1}{n^{1/2+\epsilon}}\right), \quad \text{uniformly over } \mathcal{F}. \tag{A-21}$$

Where $\epsilon > 0$ denotes the constant described in Assumption 4. For a given $y_2\in\mathbb{R}$, $y_1\in\{0,1\}$ and $u,\widetilde{u}\in\mathbb{R}^{L_V}\times\mathbb{R}^{L_V}$ and $h>0$, let

$$\begin{aligned} \varphi^{Q_2}_F(y_1,u,\widetilde{u},h) &\equiv y_1\cdot\phi_2(u)\cdot\Gamma(u,\widetilde{u},h) - E_F\left[Y_1\phi_2(V)\cdot\Gamma(V,\widetilde{u},h)\right],\\ \varphi^{R_2}_F(y_2,y_1,u,\widetilde{u},h) &\equiv y_2\cdot y_1\cdot\phi_2(u)\cdot\Gamma(u,\widetilde{u},h) - E_F\left[Y_2Y_1\phi_2(V)\cdot\Gamma(V,\widetilde{u},h)\right], \end{aligned} \tag{A-22}$$

and for a given $v, \widetilde{v} \in \mathbb{R}^{L_V} \times \mathbb{R}^{L_V}$ and $\beta \in \Theta$, let

$$\zeta^{\tau_2}_{a,F}(v,\widetilde{v},\beta) \equiv (R_{2F}(v) - (x_2'\beta_2 - \widetilde{x}_2'\beta_2)\, Q_{2F}(v)) \cdot \mathbb{1}\{g_{1U}(\widetilde{w}_1,\beta_1) \le g_{1L}(w_1,\beta_1)\}\, \phi_2(v)\phi_2(\widetilde{v}),$$
$$\zeta^{\tau_2}_{b,F}(v,\widetilde{v},\beta) \equiv (R_{2F}(\widetilde{v}) + (x_2'\beta_2 - \widetilde{x}_2'\beta_2)\, Q_{2F}(\widetilde{v})) \cdot \mathbb{1}\{g_{1U}(\widetilde{w}_1,\beta_1) \le g_{1L}(w_1,\beta_1)\}\, \phi_2(v)\phi_2(\widetilde{v}),$$
$$\zeta^{\tau_2}_{c,F}(v,\widetilde{v},\beta) \equiv Q_{2F}(\widetilde{v}) \cdot \mathbb{1}\{g_{1U}(\widetilde{w}_1,\beta_1) \le g_{1L}(w_1,\beta_1)\}\, \phi_2(v)\phi_2(\widetilde{v}),$$
$$\zeta^{\tau_2}_{d,F}(v,\widetilde{v},\beta) \equiv Q_{2F}(v) \cdot \mathbb{1}\{g_{1U}(\widetilde{w}_1,\beta_1) \le g_{1L}(w_1,\beta_1)\}\, \phi_2(v)\phi_2(\widetilde{v}),$$

and

$$\begin{aligned}\zeta^{\tau_2}_{F}(Y_2,Y_1,V,v,\widetilde{v},\beta,h) \equiv\, & \zeta^{\tau_2}_{a,F}(v,\widetilde{v},\beta)\varphi^{Q_2}_F(Y_1,V,\widetilde{v},h) - \zeta^{\tau_2}_{b,F}(v,\widetilde{v},\beta)\varphi^{Q_2}_F(Y_1,V,v,h)\\ &+ \zeta^{\tau_2}_{c,F}(v,\widetilde{v},\beta)\varphi^{R_2}_F(Y_2,Y_1,V,v,h) - \zeta_{d,F}(v,\widetilde{v},\beta)^{\tau_2}\varphi^{R_2}_F(Y_2,Y_1,V,\widetilde{v},h)\end{aligned}$$

Note that $E_F[\zeta^{\tau_2}_F(Y_2,Y_1,V,v,\widetilde{v},\beta,h)] = 0$ for all $(v,\widetilde{v},\beta,h)$. Plugging in (A-20) and (A-21) into (A-19), and using the definitions of $\nu^{Q_2}_n(\cdot)$ and $\nu^{R_2}_n(\cdot)$ given in (A-3), we have

$$\begin{aligned}\widehat{\tau}_2(v,\widetilde{v},\beta) - \tau_{2F}(v,\widetilde{v},\beta) &= \frac{1}{h_n^{r_v}} \cdot \frac{1}{n}\sum_{k=1}^{n} \zeta^{\tau_2}_F(Y_{2k},Y_{1k},V_k,v,\widetilde{v},\beta,h_n) + \xi^{\tau_2}_n(v,\widetilde{v},\beta),\\ \text{where} \quad &\sup_{\substack{v,\widetilde{v}\in\mathbb{R}^{L_V}\times\mathbb{R}^{L_V}\\ \beta\in\Theta}} \left|\xi^{\tau_2}_n(v,\widetilde{v},\beta)\right| = o_p\left(\frac{1}{n^{1/2+\epsilon}}\right), \quad \text{uniformly over } \mathcal{F}\end{aligned} \tag{A-23}$$

Where $\epsilon > 0$ denotes the constant described in Assumption 4. Let

$$\alpha^{\tau_2}_F(v,\widetilde{v},\beta) \equiv \left(\zeta^{\tau_2}_{a,F}(v,\widetilde{v},\beta),\, \zeta^{\tau_2}_{b,F}(v,\widetilde{v},\beta),\, \zeta^{\tau_2}_{c,F}(v,\widetilde{v},\beta),\, \zeta^{\tau_2}_{d,F}(v,\widetilde{v},\beta)\right).$$

By the conditions in Assumption 2, there exists a finite constant $\overline{M}_2$ such that

$$\sup_{\substack{v,\widetilde{v}\in\mathbb{R}^{L_V}\times\mathbb{R}^{L_V}\\ \beta\in\Theta}} \left\|\alpha^{\tau_2}_F(v,\widetilde{v},\beta)\right\| \le \overline{M}_2 \quad \forall\, F \in \mathcal{F}.$$

Consider the class of functions,

$$\begin{aligned}\mathscr{H}_{1,F} \equiv \Big\{ & m(y_2,y_1,v) = \alpha_1\varphi^{Q_2}_F(y_1,v,\widetilde{u},h) + \alpha_2\varphi^{Q_2}_F(y_1,v,u,h) + \alpha_3\varphi^{R_2}_F(y_2,y_1,v,u,h) + \alpha_4\varphi^{R_2}_F(y_2,y_1,v,\widetilde{u},h) :\\ & u,\widetilde{u} \in \mathbb{R}^{L_V}\times\mathbb{R}^{L_V},\ \beta\in\Theta,\ h>0,\ \|(\alpha_1,\alpha_2,\alpha_3,\alpha_4)\| \le \overline{M}_2 \Big\}\end{aligned}$$

By Nolan and Pollard (1987, Lemma 22) (or Pakes and Pollard (1989, Example 10)), and Pakes and Pollard (1989, Lemma 2.14) and the bounded-variation properties of the weight function $\phi_2(\cdot)$ and the kernel $K(\cdot)$, there exist constants $(A,V)$ such that $\mathscr{H}_{1,F}$ is Euclidean $(A,V)$ for all $F \in \mathcal{F}$, for an envelope of the form $H_1 = C_1 + \overline{C}_2 \cdot |Y_2|$, where $C_1$ and $C_2$ are constant for all $F$. Now define,

$$\mathscr{G}_{1,F} \equiv \left\{ m(y_2, y_1, v) = \zeta_F^{\tau_2}(y_2, y_1, v, u, \widetilde{u}, \beta, h) : u, \widetilde{u} \in \mathbb{R}^{L_V} \times \mathbb{R}^{L_V} \ , \beta \in \Theta \ , h > 0 \right\}.$$

Note that $\mathscr{G}_{1,F} \subseteq \mathscr{H}_{1,F}$. Therefore, there exist constants $(A, V)$ such that $\mathscr{G}_{1,F}$ is Euclidean $(A, V)$ for all $F \in \mathcal{F}$, for an envelope of the form $H_1 = C_1 + C_2 \cdot |Y_2|$, where $C_1$ and $C_2$ are constant for all $F$. Define the empirical process $\nu_n^{\tau_2}(\cdot)$ given by,

$$\left\{ \nu_n^{\tau_2}(u, \widetilde{u}, \beta, h) = \frac{1}{n} \sum_{i=1}^{n} \zeta_F^{\tau_2}(Y_{2i}, Y_{1i}, V_i, u, \widetilde{u}, \beta, h) : u, \widetilde{u} \in \mathbb{R}^{L_V} \times \mathbb{R}^{L_V} \ , \beta \in \Theta \ , h > 0 \right\}.$$

$\nu_n^{\tau_2}(\cdot)$ satisfies the conditions of Result A1. Since there exists a finite constant $\overline{D}_4$ such that $E_F[|Y_2|^4] \leq \overline{D}_4$ for all $F \in \mathcal{F}$ by Assumption 3, Result A1 implies that there exists a constant $\overline{M}$ such that, for each $\varepsilon > 0$,

$$P_F\left( \sup_{\substack{u,\widetilde{u} \in \mathbb{R}^{L_V} \times \mathbb{R}^{L_V} \\ \beta \in \Theta, h > 0}} \left| \nu_n^{\tau_2}(u, \widetilde{u}, \beta, h) \right| > \varepsilon \right) \leq \frac{\overline{M}}{n^{1/2}\varepsilon} \quad \forall \ F \in \mathcal{F}.$$

Therefore,

$$\sup_{\substack{u,\widetilde{u} \in \mathbb{R}^{L_V} \times \mathbb{R}^{L_V} \\ \beta \in \Theta, h > 0}} \left| \nu_n^{\tau_2}(u, \widetilde{u}, \beta, h) \right| = O_p\left( \frac{1}{n^{1/2}} \right) \text{ uniformly over } \mathcal{F}.$$

From here, (A-23) yields,

$$\begin{aligned} \sup_{\substack{v,\widetilde{v} \in \mathbb{R}^{L_V} \times \mathbb{R}^{L_V} \\ \beta \in \Theta}} \left| \widehat{\tau}_2(v, \widetilde{v}, \beta) - \tau_{2F}(v, \widetilde{v}, \beta) \right| &= O_p\left( \frac{1}{h_n^{r_v} \cdot n^{1/2}} \right) + o_p\left( \frac{1}{n^{1/2+\epsilon}} \right) \\ &= o_p\left( \frac{1}{n^{1/4+\epsilon/2}} \right) \quad \text{uniformly over } \mathcal{F}. \end{aligned} \tag{A-24}$$

Where $\epsilon > 0$ is the constant described in Assumption 4. By the conditions in Assumption 2, there exists a finite constant $\overline{\tau}_2$ such that

$$\sup_{\substack{v,\widetilde{v} \in \mathbb{R}^{L_V} \times \mathbb{R}^{L_V} \\ \beta \in \Theta}} \left| \tau_{2F}(v, \widetilde{v}, \beta) \right| \leq \overline{\tau}_2 \quad \forall \ F \in \mathcal{F}.$$

From here and (A-24), we obtain,

$$\sup_{\substack{v,\widetilde{v}\in\mathbb{R}^{L_V}\times\mathbb{R}^{L_V}\\ \beta\in\Theta}} \left|\widehat{\tau}_2(v,\widetilde{v},\beta)\right| = O_p(1) \quad \text{uniformly over } \mathcal{F}. \tag{A-25}$$

The results in (A-12), (A-23), (A-24) and (A-25) summarize the relevant asymptotic properties of $\widehat{\tau}_2(v,\widetilde{v},\beta)$ for our problem.

### A3.4 Asymptotic properties of $\widehat{\mathcal{T}}_2(\beta)$

Recall that,

$$\widehat{\mathcal{T}}_2(\beta) \equiv \frac{1}{n\cdot(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}\widehat{\tau}_2(V_i,V_j,\beta)\cdot\mathbb{1}\left\{\widehat{\tau}_2(V_i,V_j,\beta)\geq -b_n\right\}.$$

Let

$$\widetilde{\mathcal{T}}_2(\beta) \equiv \frac{1}{n\cdot(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}\widehat{\tau}_2(V_i,V_j,\beta)\cdot\mathbb{1}\left\{\tau_{2F}(V_i,V_j,\beta)\geq 0\right\}.$$

Note that $\widetilde{\mathcal{T}}_2(\beta)$ takes $\widehat{\mathcal{T}}_2(\beta)$ and replaces the indicator function $\mathbb{1}\left\{\widehat{\tau}_2(V_i,V_j,\beta)\geq -b_n\right\}$ with the indicator function $\mathbb{1}\left\{\tau_{2F}(V_i,V_j,\beta)\geq 0\right\}$. Our first step is to analyze $\widehat{\mathcal{T}}_2(\beta)-\widetilde{\mathcal{T}}_2(\beta)$. Denote,

$$r_{n,F}^{\mathcal{T}_2}(\beta) \equiv \widehat{\mathcal{T}}_2(\beta)-\widetilde{\mathcal{T}}_2(\beta) = \frac{1}{n\cdot(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}\widehat{\tau}_2(V_i,V_j,\beta)\cdot\left[\mathbb{1}\left\{\widehat{\tau}_2(V_i,V_j,\beta)\geq -b_n\right\}-\mathbb{1}\left\{\tau_{2F}(V_i,V_j,\beta)\geq 0\right\}\right].$$

Thus,

$$\left|r_{n,F}^{\mathcal{T}_2}(\beta)\right| \leq \frac{1}{n\cdot(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}\left|\widehat{\tau}_2(V_i,V_j,\beta)\right|\cdot\left|\mathbb{1}\left\{\widehat{\tau}_2(V_i,V_j,\beta)\geq -b_n\right\}-\mathbb{1}\left\{\tau_{2F}(V_i,V_j,\beta)\geq 0\right\}\right|.$$

As we pointed out in (A-15), we have

$$\begin{aligned}
&\left|\mathbb{1}\left\{\widehat{\tau}_2(V_i,V_j,\beta)\geq -b_n\right\}-\mathbb{1}\left\{\tau_{2F}(V_i,V_j,\beta)\geq 0\right\}\right|\\
=&\mathbb{1}\left\{\widehat{\tau}_2(V_i,V_j,\beta)\geq -b_n\ ,\ -2b_n\leq\tau_{2F}(V_i,V_j,\beta)<0\right\}+\mathbb{1}\left\{\widehat{\tau}_2(V_i,V_j,\beta)\geq -b_n\ ,\ \tau_{2F}(V_i,V_j,\beta)<-2b_n\right\}\\
&+\mathbb{1}\left\{\widehat{\tau}_2(V_i,V_j,\beta)<-b_n\ ,\ \tau_{2F}(V_i,V_j,\beta)\geq 0\right\}\\
\leq&\mathbb{1}\left\{-2b_n\leq\tau_{2F}(V_i,V_j,\beta)<0\right\}+\mathbb{1}\left\{\left|\widehat{\tau}_2(V_i,V_j,\beta)-\tau_{2F}(V_i,V_j,\beta)\right|\geq b_n\right\}
\end{aligned} \tag{A-26}$$

From here, we have,

$$\begin{aligned}
&\left|r^{\mathcal{I}_2}_{n,F}(\beta)\right| \\
\leq& \frac{1}{n\cdot(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}\left|\widehat{\tau}_2(V_i,V_j,\beta)\right|\cdot\mathbb{1}\left\{-2b_n\leq\tau_{2F}(V_i,V_j,\beta)<0\right\} \\
&+\frac{1}{n\cdot(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}\left|\widehat{\tau}_2(V_i,V_j,\beta)\right|\cdot\mathbb{1}\left\{\left|\widehat{\tau}_2(V_i,V_j,\beta)-\tau_{2F}(V_i,V_j,\beta)\right|\geq b_n\right\} \\
\leq& \frac{1}{n\cdot(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}\left(\left|\tau_2(V_i,V_j,\beta)\right|+\left|\widehat{\tau}_2(V_i,V_j,\beta)-\tau_{2F}(V_i,V_j,\beta)\right|\right)\cdot\mathbb{1}\left\{-2b_n\leq\tau_{2F}(V_i,V_j,\beta)<0\right\} \\
&+\frac{1}{n\cdot(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}\left|\widehat{\tau}_2(V_i,V_j,\beta)\right|\cdot\mathbb{1}\left\{\left|\widehat{\tau}_2(V_i,V_j,\beta)-\tau_{2F}(V_i,V_j,\beta)\right|\geq b_n\right\}
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\left|r^{\mathcal{I}_2}_{n,F}(\beta)\right| \\
\leq&\left(2b_n+\sup_{\substack{v,\widetilde{v}\in\mathbb{R}^{L_V}\times\mathbb{R}^{L_V}\\ \beta\in\Theta}}\left|\widehat{\tau}_2(v,\widetilde{v},\beta)-\tau_{2F}(v,\widetilde{v},\beta)\right|\right)\times\frac{1}{n\cdot(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}\mathbb{1}\left\{-2b_n\leq\tau_{2F}(V_i,V_j,\beta)<0\right\} \\
&+\sup_{\substack{v,\widetilde{v}\in\mathbb{R}^{L_V}\times\mathbb{R}^{L_V}\\ \beta\in\Theta}}\left|\widehat{\tau}_2(v,\widetilde{v},\beta)\right|\times\frac{1}{n\cdot(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}\mathbb{1}\left\{\left|\widehat{\tau}_2(V_i,V_j,\beta)-\tau_{2F}(V_i,V_j,\beta)\right|\geq b_n\right\}
\end{aligned}$$

From here and the results in (A-24) and (A-25), uniformly over $\mathcal{F}$, we have

$$\begin{aligned}
\sup_{\beta\in\Theta}\left|r^{\mathcal{I}_2}_{n,F}(\beta)\right|\leq&\left(2b_n+o_p\left(\frac{1}{n^{1/4+\epsilon/2}}\right)\right)\times\sup_{\beta\in\Theta}\left|\frac{1}{n\cdot(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}\mathbb{1}\left\{-2b_n\leq\tau_{2F}(V_i,V_j,\beta)<0\right\}\right| \\
&+O_p(1)\times\sup_{\beta\in\Theta}\left|\frac{1}{n\cdot(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}\mathbb{1}\left\{\left|\widehat{\tau}_2(V_i,V_j,\beta)-\tau_{2F}(V_i,V_j,\beta)\right|\geq b_n\right\}\right|
\end{aligned}\tag{A-27}$$

Where $\epsilon>0$ is the constant described in Assumption 4. Let us analyze each term on the right hand side of (A-27). In what follows, let $V_1,V_2$ be independent draws from the distribution $F$. For a given $\beta\in\Theta$ and $b>0$, let

$$g_{2F}(V_1,V_2,\beta,b)\equiv\frac{1}{2}\Big(\mathbb{1}\left\{-2b\leq\tau_{2F}(V_1,V_2,\beta)<0\right\}+\mathbb{1}\left\{-2b\leq\tau_{2F}(V_2,V_1,\beta)<0\right\}\Big).$$

$g_{2F}(V_1, V_2, \beta, b)$ is symmetric in $V_1, V_2$ by construction. Note that

$$\frac{1}{n\cdot(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}\mathbb{1}\left\{-2b \leq \tau_{2F}(V_i, V_j, \beta) < 0\right\} = \binom{2}{n}^{-1}\sum_{i<j} g_{2F}(V_i, V_j, \beta, b) \equiv S^{g}_{2,n}(\beta, b).$$

We will focus on the properties of the U-process $\left\{S^{g}_{2,n}(\beta, b) : \beta \in \Theta,\ 0 < b \leq \frac{c_0}{2}\right\}$, where $c_0$ is the constant described in Assumption 3. We will proceed by analyzing the Hoeffding decomposition (see Serfling (1980, pages 177-178) or Sherman (1994, equations (6)-(7))) of $S^{g}_{2,n}(\beta, b)$. Let

$$\mu^{g}_{2F}(\beta, b) \equiv E_F\left[\mathbb{1}\{-2b \leq \tau_{2F}(V_1, V_2, \beta) < 0\}\right],$$

Note that $\mu^{g}_{2F}(\beta, b) = E_F\left[g_{2F}(V_1, V_2, \beta, b)\right]$ by symmetry. Let

$$\begin{aligned}
\widetilde{g}_{2F}(V_1, V_2, \beta, b) &\equiv g_{2F}(V_1, V_2, \beta, b) - \mu^{g}_{2F}(\beta, b),\\
\widetilde{m}_{1,F}(V_1, \beta, b) &\equiv E_F\left[\widetilde{g}_{2F}(V_1, V_2, \beta, b)|V_1\right],\\
\widetilde{m}_{2,F}(V_1, V_2, \beta, b) &\equiv \widetilde{g}_{2F}(V_1, V_2, \beta, b) - \widetilde{m}_{1,F}(V_1, \beta, b) - \widetilde{m}_{1,F}(V_2, \beta, b),
\end{aligned}$$

The Hoeffding decomposition of $S^{g}_{2,n}(\beta, b)$ (see Serfling (1980, pages 177-178) or Sherman (1994, equations (6)-(7))) is given by,

$$S^{g}_{2,n}(\beta, b) = \mu^{g}_{2F}(\beta, b) + \frac{2}{n}\sum_{i=1}^{n}\widetilde{m}_{1,F}(V_i, \beta, b) + \binom{n}{2}^{-1}\sum_{i<j}\widetilde{m}_{2,F}(V_i, V_j, \beta, b). \qquad \text{(A-28)}$$

Let us analyze the second and third terms on the right-hand side of (A-28). By the properties of VC classes of sets described, e.g, in Pakes and Pollard (1989, Lemma 2.5), the conditions described in Assumption 3 imply that, for each $F \in \mathcal{F}$, the following class of sets is a VC class, with VC dimension uniformly bounded over $\mathcal{F}$ by a finite constant,

$$\mathscr{D}^{\tau_2}_{2,F} \equiv \left\{(v_1, v_2) \in \mathbb{R}^{L_v} \times \mathbb{R}^{L_v} :\ -c \leq \tau_{2F}(v_1, v_2, \beta) < 0 \ \text{ for some } 0 < c \leq c_0 \text{ and } \beta \in \Theta\right\},$$

where the constant $c_0$ is as described in Assumption 3. From here, the result in Pakes and Pollard (1989, p. 1033) implies that there exist constants $(\overline{A}, \overline{V})$ such that, for each $F \in \mathcal{F}$, the class of indicator functions,

$$\mathscr{M}_F \equiv \left\{m(v_1, v_2) = \mathbb{1}\{-c \leq \tau_{2F}(v_1, v_2, \beta) < 0\} \ \text{ for some } 0 < c \leq c_0 \text{ and } \beta \in \Theta\right\}$$

is Euclidean $(\overline{A}, \overline{V})$ for the constant envelope 1. From here and Sherman (1994, Lemma 5), the conditions for Result A1 are satisfied and, from there, we obtain,

$$\left.\begin{aligned} &\sup_{\substack{\beta\in\Theta\\ 0<b\leq\frac{c_0}{2}}} \left|\tfrac{1}{n}\textstyle\sum_{i=1}^{n}\widetilde{m}_{1,F}(V_i,\beta,b)\right| = O_p\left(\tfrac{1}{n^{1/2}}\right) \\ &\sup_{\substack{\beta\in\Theta\\ 0<b\leq\frac{c_0}{2}}} \left|\tbinom{n}{2}^{-1}\sum_{i<j}\widetilde{m}_{2,F}(V_i,V_j,\beta,b)\right| = O_p\left(\tfrac{1}{n}\right) \end{aligned}\right\} \text{uniformly over } \mathcal{F}. \tag{A-29}$$

Combining (A-29) and (A-28), we have

$$S^g_{2,n}(\beta,b) = \mu^g_{2F}(\beta,b) + \xi^g_n(\beta,b), \quad \text{where} \quad \sup_{\substack{\beta\in\Theta\\ 0<b\leq\frac{c_0}{2}}} \left|\xi^g_n(\beta,b)\right| = O_p\left(\frac{1}{n^{1/2}}\right), \quad \text{uniformly over } \mathcal{F}.$$

Next, recall that, from Assumption 5, there exists $b_0 > 0$ and $\overline{m} < \infty$ such that,

$$\sup_{\beta\in\Theta}\left|\mu^g_{2F}(\beta,b)\right| \leq \overline{m}\cdot b \quad \forall\ 0 < b \leq b_0, \quad \forall\ F \in \mathcal{F}.$$

Next note that there exists $n_0$ such that $b_n < \left(\frac{c_0}{2}\right) \wedge b_0$ for all $n > n_0$. Therefore, for all $n > n_0$,

$$\begin{aligned} \sup_{\beta\in\Theta}\left|\frac{1}{n\cdot(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}\mathbb{1}\left\{-2b_n \leq \tau_{2F}(V_i,V_j,\beta) < 0\right\}\right| &\leq \overline{m}\cdot b_n + \sup_{\substack{\beta\in\Theta\\ 0<b\leq\frac{c_0}{2}}}\left|\xi^g_n(\beta,b)\right| = O(b_n) + O_p\left(\frac{1}{n^{1/2}}\right) \\ &= b_n \times \left(O(1) + o_p\left(\frac{1}{b_n\cdot n^{1/2}}\right)\right) \\ &= b_n \times \left(O(1) + o_p(1)\right) \\ &= O_p(b_n), \quad \text{uniformly over } \mathcal{F}. \end{aligned}$$

Thus, uniformly over $\mathcal{F}$, we have

$$\begin{aligned} \left(2b_n + o_p\left(\frac{1}{n^{1/4+\epsilon/2}}\right)\right)\times\sup_{\beta\in\Theta}\left|\frac{1}{n\cdot(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}\mathbb{1}\left\{-2b_n \leq \tau_{2F}(V_i,V_j,\beta) < 0\right\}\right| &= \left(2b_n + o_p\left(\frac{1}{n^{1/4+\epsilon/2}}\right)\right)\times O_p(b_n) \\ &= O_p\left(b_n^2\right) + o_p\left(\frac{b_n}{n^{1/4+\epsilon/2}}\right) \\ &= o_p\left(\frac{1}{n^{1/2+\epsilon}}\right), \end{aligned}$$

where $\epsilon > 0$ is the constant described in Assumption 4 Going back to (A-27), this result implies that, uniformly over $\mathcal{F}$,

$$\sup_{\beta\in\Theta}\left|r_{n,F}^{\mathcal{T}_2}(\beta)\right| \leq \left(2b_n + o_p\left(\frac{1}{n^{1/4+\epsilon/2}}\right)\right)\times\sup_{\beta\in\Theta}\left|\frac{1}{n\cdot(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}\mathbb{1}\left\{-2b_n \leq \tau_{2F}(V_i,V_j,\beta) < 0\right\}\right|$$
$$+ O_p(1)\times\sup_{\beta\in\Theta}\left|\frac{1}{n\cdot(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}\mathbb{1}\left\{\left|\widehat{\tau}_2(V_i,V_j,\beta)-\tau_{2F}(V_i,V_j,\beta)\right| \geq b_n\right\}\right|$$
$$= O_p(1)\times\sup_{\beta\in\Theta}\left|\frac{1}{n\cdot(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}\mathbb{1}\left\{\left|\widehat{\tau}_2(V_i,V_j,\beta)-\tau_{2F}(V_i,V_j,\beta)\right| \geq b_n\right\}\right| + o_p\left(\frac{1}{n^{1/2+\epsilon}}\right) \tag{A-30}$$

where $\epsilon > 0$ is the constant described in Assumption 4. Take any $C > 0$ and any $\Delta > 0$. We have,

$$P_F\left(\sup_{\beta\in\Theta}\left|\frac{1}{n\cdot(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}\mathbb{1}\left\{\left|\widehat{\tau}_2(V_i,V_j,\beta)-\tau_{2F}(V_i,V_j,\beta)\right| \geq b_n\right\}\right| > \frac{C}{n^{\Delta}}\right)$$
$$\leq P_F\left(\sup_{\substack{v,\widetilde{v}\in\mathbb{R}^{L_V}\times\mathbb{R}^{L_V}\\ \beta\in\Theta}}\left|\widehat{\tau}_2(v,\widetilde{v},\beta)-\tau_{2F}(v,\widetilde{v},\beta)\right| > b_n\right)$$

Since the bandwidth sequence $b_n$ satisfies $n^{1/2}\cdot h_n^r\cdot b_n \longrightarrow \infty$ by Assumption 4, the result we obtained in equation (A-12) yields,

$$\sup_{F\in\mathcal{F}} P_F\left(\sup_{\beta\in\Theta}\left|\frac{1}{n\cdot(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}\mathbb{1}\left\{\left|\widehat{\tau}_2(V_i,V_j,\beta)-\tau_{2F}(V_i,V_j,\beta)\right| \geq b_n\right\}\right| > \frac{C}{n^{\Delta}}\right) \longrightarrow 0,$$

for any $C > 0$ and $\Delta > 0$. In particular, this holds for $\Delta = 1/2+\epsilon$, with $\epsilon > 0$ being the the constant described in Assumption 4. Therefore, under Assumptions 2, 3 and 4,

$$\sup_{\beta\in\Theta}\left|\frac{1}{n\cdot(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}\mathbb{1}\left\{\left|\widehat{\tau}_2(V_i,V_j,\beta)-\tau_{2F}(V_i,V_j,\beta)\right| \geq b_n\right\}\right| = o_p\left(\frac{1}{n^{1/2+\epsilon}}\right), \quad \text{uniformly over } \mathcal{F}. \tag{A-31}$$

Plugging (A-31) into (A-30), we obtain that, under Assumptions 2, 3, 4 and 5,

$$\sup_{\beta\in\Theta}\left|r_{n,F}^{\mathcal{T}_2}(\beta)\right| = o_p\left(\frac{1}{n^{1/2+\epsilon}}\right), \quad \text{uniformly over } \mathcal{F}.$$

Where $\epsilon > 0$ is the constant described in Assumption 4. Since we defined $r_{n,F}^{\mathcal{T}_2}(\beta) \equiv \widehat{\mathcal{T}}_2(\beta) - \widetilde{\mathcal{T}}_2(\beta)$, with $\widetilde{\mathcal{T}}_2(\beta) \equiv \frac{1}{n\cdot(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}\widehat{\tau}_2(V_i,V_j,\beta)\cdot \mathbb{1}\left\{\tau_{2F}(V_i,V_j,\beta)\geq 0\right\}$, we have that, under Assumptions 2, 3, 4 and 5,

$$\widehat{\mathcal{T}}_2(\beta) = \widetilde{\mathcal{T}}_2(\beta) + r_{n,F}^{\mathcal{T}_2}(\beta), \quad \text{where} \quad \sup_{\beta\in\Theta}\left|r_{n,F}^{\mathcal{T}_2}(\beta)\right| = o_p\left(\frac{1}{n^{1/2+\epsilon}}\right), \quad \text{uniformly over } \mathcal{F}.$$

Where $\epsilon > 0$ denotes the constant described in Assumption 4. Our next step is to analyze the asymptotic properties of $\widetilde{\mathcal{T}}_2(\beta)$.

### A3.4.1 Asymptotic properties of $\widetilde{\mathcal{T}}_2(\beta)$

Denote $(A)_+ \equiv \max\{A,0\}$. We have,

$$\begin{aligned}
\widetilde{\mathcal{T}}_2(\beta) &\equiv \frac{1}{n\cdot(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}\widehat{\tau}_2(V_i,V_j,\beta)\cdot\mathbb{1}\left\{\tau_{2F}(V_i,V_j,\beta)\geq 0\right\} \\
&= \frac{1}{n\cdot(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}\Big(\tau_{2F}(V_i,V_j,\beta)\Big)_+ \\
&+ \frac{1}{n\cdot(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}\Big(\widehat{\tau}_2(V_i,V_j,\beta)-\tau_{2F}(V_i,V_j,\beta)\Big)\cdot\mathbb{1}\left\{\tau_{2F}(V_i,V_j,\beta)\geq 0\right\}
\end{aligned}$$

For a pair $v \equiv (x_2,w_1,w_1)$, $\widetilde{v}\equiv(\widetilde{x}_2,\widetilde{w}_1)$, denote,

$$\mathbb{I}_{2F}(v,\widetilde{v},\beta) \equiv \mathbb{1}\{g_{1U}(\widetilde{w}_1,\beta_1)\leq g_{1L}(w_1,\beta_1)\}\cdot\mathbb{1}\{\tau_{2F}(v,\widetilde{v},\beta)\geq 0\}\cdot\phi_2(v)\cdot\phi_2(\widetilde{v}). \tag{A-32}$$

And, for a given $v,\widetilde{v}\in\mathbb{R}^{L_V}\times\mathbb{R}^{L_V}$ and $\beta\in\Theta$, let

$$\begin{aligned}
\delta_{a,F}^{\tau_2}(v,\widetilde{v},\beta) &\equiv (R_{2F}(v)-(x_2'\beta_2-\widetilde{x}_2'\beta_2)\,Q_{2F}(v))\cdot\mathbb{I}_{2F}(v,\widetilde{v},\beta), \\
\delta_{b,F}^{\tau_2}(v,\widetilde{v},\beta) &\equiv (R_{2F}(\widetilde{v})-(\widetilde{x}_2'\beta_2-x_2'\beta_2)\,Q_{2F}(\widetilde{v}))\cdot\mathbb{I}_{2F}(v,\widetilde{v},\beta), \\
\delta_{c,F}^{\tau_2}(v,\widetilde{v},\beta) &\equiv Q_{2F}(\widetilde{v})\cdot\mathbb{I}_{2F}(v,\widetilde{v},\beta), \\
\delta_{d,F}^{\tau_2}(v,\widetilde{v},\beta) &\equiv Q_{2F}(v)\cdot\mathbb{I}_{2F}(v,\widetilde{v},\beta).
\end{aligned}$$

And let $\varphi_F^{Q_2}(u,\widetilde{u},h)$ and $\varphi_F^{R_2}(y_2,u,\widetilde{u},h)$ be as defined in (A-22). As we defined previously, let us group all the observable covariates in the model as $Z\equiv(Y_1,Y_2,V)$. For given $(z,\widetilde{z},\ddot{z})\in\mathbb{R}^{L_v+2}\times\mathbb{R}^{L_v+2}\times\mathbb{R}^{L_v+2}$, $\beta\in\Theta$ and $h>0$, let

$$\begin{aligned}
\varphi_F^{\tau_2}(z,\widetilde{z},\ddot{z},\beta,h) \equiv\; &\delta_{a,F}^{\tau_2}(v,\widetilde{v},\beta)\varphi_F^{Q_2}(\ddot{y}_1,\ddot{v},\widetilde{v},h) - \delta_{b,F}^{\tau_2}(v,\widetilde{v},\beta)\varphi_F^{Q_2}(\ddot{y}_1,\ddot{v},v,h) \\
&+\delta_{c,F}^{\tau_2}(v,\widetilde{v},\beta)\varphi_F^{R_2}(\ddot{y}_2,\ddot{y}_1,\ddot{v},v,h) - \delta_{d,F}^{\tau_2}(v,\widetilde{v},\beta)\varphi_F^{R_2}(\ddot{y}_2,\ddot{y}_1,\ddot{v},\widetilde{v},h)
\end{aligned} \tag{A-33}$$

Note by inspection of the definitions in (A-22) that,

$$E_F\left[\varphi_F^{Q_2}(Y_1,V,v,h)\right]=0, \quad \text{and} \quad E_F\left[\varphi_F^{R_2}(Y_2,Y_1,V,v,h)\right]=0 \quad \forall\, v\in\mathbb{R}^{L_v},\ h>0. \tag{A-34}$$

Recall that we have defined $\mu_{2F}(v)\equiv E_F[Y_2|V=v,Y_1=1]$. By the smoothness conditions in Assumption 2 and the kernel properties in Assumption 4, an $M^{th}$–order approximation implies that there exists a finite $\overline{B}$ such that

$$\begin{aligned}
\varphi_F^{Q_2}(Y_1,V,v,h_n) =& Y_1\phi_2(V)\Gamma(V,v,h_n)-h_n^{r_v}\cdot\phi_2(v)f_{V,1}(v)+B_{n,F}^{Q_2}(v),\\
\varphi_F^{R_2}(Y_2,Y_1,V,v,h_n) =& Y_2Y_1\phi_2(V)\Gamma(V,v,h_n)-h_n^{r_v}\cdot\mu_{2F}(v)\cdot\phi_2(v)f_{V,1}(v)+B_{n,F}^{R_2}(v),\\
&\text{where } \sup_{v\in\mathcal{V}}\left|B_{n,F}^{Q_2}(v)\right|\le\overline{B}\cdot h_n^{r_v+M}, \quad \sup_{v\in\mathcal{V}}\left|B_{n,F}^{R_2}(v)\right|\le\overline{B}\cdot h_n^{r_v+M} \quad \forall\, F\in\mathcal{F}.
\end{aligned} \tag{A-35}$$

From (A-23), we have

$$\begin{aligned}
\widetilde{\mathcal{T}}_2(\beta) =& \frac{1}{n\cdot(n-1)}\sum_{i=1}^n\sum_{j\neq i}\Big(\tau_{2F}(V_i,V_j,\beta)\Big)_+ + \frac{1}{h_n^{r_v}}\cdot\frac{1}{n^2\cdot(n-1)}\sum_{i=1}^n\sum_{j\neq i}\sum_{k=1}^n\varphi_F^{\tau_2}(Z_i,Z_j,Z_k,\beta,h_n)+\xi_{a,n}^{\widetilde{\mathcal{T}}_2}(\beta),\\
&\text{where } \sup_{\beta\in\Theta}\left|\xi_{a,n}^{\widetilde{\mathcal{T}}_2}(\beta)\right|=o_p\left(\frac{1}{n^{1/2+\epsilon}}\right), \text{ uniformly over } \mathcal{F}.
\end{aligned} \tag{A-36}$$

Where $\epsilon>0$ denotes the constant described in Assumption 4. Let

$$\begin{aligned}
U_{a,n}(\beta,h) &\equiv \frac{1}{n\cdot(n-1)\cdot(n-2)}\sum_{i=1}^n\sum_{j\neq i}\sum_{k\neq i,j}\varphi_F^{\tau_2}(Z_i,Z_j,Z_k,\beta,h),\\
U_{b,n}(\beta,h) &\equiv \frac{1}{n\cdot(n-1)}\sum_{i=1}^n\sum_{j\neq i}\Big(\varphi_F^{\tau_2}(Z_i,Z_j,Z_i,\beta,h)+\varphi_F^{\tau_2}(Z_i,Z_j,Z_j,\beta,h)\Big)
\end{aligned}$$

Then, (A-36) can be re-expressed as,

$$\begin{aligned}
\widetilde{\mathcal{T}}_2(\beta) =& \frac{1}{n\cdot(n-1)}\sum_{i=1}^n\sum_{j\neq i}\Big(\tau_{2F}(V_i,V_j,\beta)\Big)_+ + \frac{(n-2)}{n}\cdot\frac{1}{h_n^{r_v}}\cdot U_{a,n}(\beta,h_n)+\frac{1}{n\cdot h_n^{r_v}}\cdot U_{b,n}(\beta,h_n)+\xi_{a,n}^{\widetilde{\mathcal{T}}_2}(\beta),\\
&\text{where } \sup_{\beta\in\Theta}\left|\xi_{a,n}^{\widetilde{\mathcal{T}}_2}(\beta)\right|=o_p\left(\frac{1}{n^{1/2+\epsilon}}\right), \text{ uniformly over } \mathcal{F}.
\end{aligned} \tag{A-37}$$

Where $\epsilon>0$ is the constant described in Assumption 4. Recall from Assumption 3 that the class of sets

$$\mathscr{C}\equiv\Big\{(w_1,w_1)\in\mathbb{R}^{d_U}\times\mathbb{R}^{d_L}: g_{1U}(w_1,\beta_1)\le g_{1L}(w_1,\beta_1)\ \text{ for some } \beta_1\in\Theta\Big\}$$

is a VC class with VC dimension $\overline{V}_C$, and that the following is a VC class of sets for each $F$, with VC dimension uniformly bounded over $\mathcal{F}$ by a finite constant $\overline{V}_D$,

$$\mathscr{D}^{\tau_2}_{1,F} \equiv \Big\{(v_1, v_2) \in \mathbb{R}^{L_v} \times \mathbb{R}^{L_v} : \tau_{2F}(v_1, v_2, \beta) \geq 0 \text{ for some } \beta \in \Theta\Big\}$$

Going back to the definition of $\mathbb{I}_{2F}$ in equation (A-32), these VC properties imply, by the results in Pakes and Pollard (1989, p. 1033) (the result that classes of indicator functions over VC classes of sets are Euclidean $(A, V)$, with $(A, V)$ depending only on the VC-dimension of the underlying class of sets), and Pakes and Pollard (1989, Lemma 2.14) (the product of Euclidean classes of functions is also a Euclidean class) that there exist constants $(\overline{A}, \overline{V})$ such that, for each $F \in \mathcal{F}$, the class of indicator functions

$$\mathscr{I}_{2,F} \equiv \{m(v, \widetilde{v}) = \mathbb{I}_{2F}(v, \widetilde{v}, \beta) : \beta \in \Theta\}, \tag{A-38}$$

is Euclidean $(\overline{A}, \overline{V})$ for the constant envelope 1. From here, let $\varphi^{\tau_2}_F$ be as defined in (A-33) and consider the class of functions,

$$\mathscr{H}_{2,F} \equiv \Big\{m(z_1, z_2, z_3) = \varphi^{\tau_2}_F(z_1, z_2, z_3, \beta, h) : \beta \in \Theta,\, h > 0\Big\}. \tag{A-39}$$

By the conditions in Assumptions 2, 3 and 4 (the bounded properties of the functionals involved, the bounded-variation properties of the weight function $\phi_2(\cdot)$ and the kernel $K(\cdot)$, and the VC property of the classes of sets involved, which led to the Euclidean property of the class of functions described in equation (A-38)), by Nolan and Pollard (1987, Lemma 22) and Pakes and Pollard (1989, Lemma 2.14), there exist constants $(\overline{A}, \overline{V})$ such that $\mathscr{H}_{1,F}$ is Euclidean $(\overline{A}, \overline{V})$ for all $F \in \mathcal{F}$, for an envelope of the form $H_1 = D_1 + D_2 \cdot |Y_2|$, where $D_1$ and $D_2$ are constant for all $F$. Since there exists a finite constant $\overline{D}_4$ such that $E_F[|Y_2|^4] \leq \overline{D}_4$ for all $F \in \mathcal{F}$ by Assumption 3, Result A1 can be used to show that,

$$\sup_{\substack{\beta\in\Theta \\ h>0}} \Big|U_{b,n}(\beta, h)\Big| = O_p(1), \quad \text{uniformly over } \mathcal{F}.$$

Therefore, using the bandwidth convergence conditions described in Assumption 4, equation (A-37) becomes,

$$\begin{gathered}\widetilde{\mathcal{T}}_2(\beta) = \frac{1}{n\cdot(n-1)} \sum_{i=1}^{n} \sum_{j\neq i} \Big(\tau_{2F}(V_i, V_j, \beta)\Big)_+ + \frac{(n-2)}{n} \cdot \frac{1}{h_n^{r_v}} \cdot U_{a,n}(\beta, h_n) + \xi^{\widetilde{\mathcal{T}}_2}_{b,n}(\beta), \\ \text{where } \sup_{\beta\in\Theta}\Big|\xi^{\widetilde{\mathcal{T}}_2}_{b,n}(\beta)\Big| = O_p\left(\frac{1}{n\cdot h_n^{r_v}}\right) + o_p\left(\frac{1}{n^{1/2+\epsilon}}\right) = o_p\left(\frac{1}{n^{1/2+\epsilon}}\right), \text{ uniformly over } \mathcal{F}.\end{gathered} \tag{A-40}$$

Where $\epsilon > 0$ is the constant described in Assumption 4. Next we focus on the Hoeffding decomposition (see Serfling (1980, pages 177-178) or Sherman (1994, equations (6)-(7))) of $U_{a,n}(\beta, h_n)$. In what follows, let $Z_1, Z_2, Z_3$ be iid draws from the distribution $F$. Let

$$\overline{\varphi}_F^{\tau_2}(Z_1, Z_2, Z_3, \beta, h) \equiv \frac{1}{3!} \sum_p \varphi_F^{\tau_2}(Z_{m_1}, Z_{m_2}, Z_{m_3}, \beta, h), \tag{A-41}$$

where $\sum_p$ denotes the sum over the 3! permutations $(m_1, m_2, m_3)$ of $(1,2,3)$. By construction, $\overline{\varphi}_F^{\tau_2}(Z_1, Z_2, Z_3, \beta, h)$ is symmetric in $(Z_1, Z_2, Z_3)$, and $U_{a,n}(\beta, h)$ can be expressed as,

$$U_{a,n}(\beta, h) = \binom{n}{3}^{-1} \sum_{i<j<k} \overline{\varphi}_F^{\tau_2}(Z_1, Z_2, Z_3, \beta, h).$$

Note from (A-34) that $E_F\left[\overline{\varphi}_F^{\tau_2}(Z_1, Z_2, Z_3, \beta, h)\right] = E_F\left[\varphi_F^{\tau_2}(Z_1, Z_2, Z_3, \beta, h)\right] = 0$. For a given $(z, \widetilde{z}, \ddot{z})$, let

$$\begin{aligned}
m_{1F}^{\tau_2}(z, \beta, h) &\equiv E_F\left[\overline{\varphi}_F^{\tau_2}(z, Z_2, Z_3, \beta, h)\right], \\
m_{2F}^{\tau_2}(z, z', \beta, h) &\equiv E_F\left[\overline{\varphi}_F^{\tau_2}(z, z', Z_3, \beta, h)\right] - m_{1F}^{\tau_2}(z, \beta, h) - m_{1F}^{\tau_2}(z', \beta, h), \\
m_{3F}(z, \widetilde{z}, \ddot{z}, \beta, h) &\equiv \overline{\varphi}_F^{\tau_2}(z, \widetilde{z}, \ddot{z}, \beta, h) - m_{2F}^{\tau_2}(z, z', \beta, h) - m_{2F}^{\tau_2}(z, z'', \beta, h) - m_{2F}^{\tau_2}(z', z'', \beta, h) \\
&\quad - m_{1F}^{\tau_2}(z, \beta, h) - m_{1F}^{\tau_2}(z', \beta, h) - m_{1F}^{\tau_2}(z'', \beta, h)
\end{aligned}$$

Let,

$$S_{2,n}^{\tau_2}(\beta, h) \equiv \binom{n}{2}^{-1} \sum_{i<j} m_{2F}^{\tau_2}(Z_i, Z_j, \beta, h), \quad S_{3,n}^{\tau_2}(\beta, h) \equiv \binom{n}{3}^{-1} \sum_{i<j<k} m_{3F}^{\tau_2}(Z_i, Z_j, Z_k, \beta, h)$$

The Hoeffding decomposition of $U_{a,n}(\beta, h_n)$ (see Serfling (1980, pages 177-178) or Sherman (1994, equations (6)-(7))) is given by,

$$U_{a,n}(\beta, h_n) = \frac{3}{n} \sum_{i=1}^{n} m_{1F}^{\tau_2}(Z_i, \beta, h_n) + 3 \cdot S_{2,n}^{\tau_2}(\beta, h_n) + S_{3,n}^{\tau_2}(\beta, h_n) \tag{A-42}$$

$\left\{S_{2,n}^{\tau_2}(\beta, h) : \beta \in \Theta, h > 0\right\}$ is a degenerate U-process of order 2 and $\left\{S_{3,n}^{\tau_2}(\beta, h) : \beta \in \Theta, h > 0\right\}$ is a degenerate U-process of order 3. The Euclidean properties of the class of functions $\mathscr{H}_{2,F}$ defined in (A-39) and described above yield, via Result A1,

$$\sup_{\substack{\beta \in \Theta \\ h > 0}} \left|S_{2,n}^{\tau_2}(\beta, h)\right| = O_p\left(\frac{1}{n}\right), \quad \text{and} \quad \left|S_{3,n}^{\tau_2}(\beta, h)\right| = O_p\left(\frac{1}{n^{3/2}}\right), \quad \text{uniformly over } \mathcal{F}.$$

Therefore, combining (A-42) and (A-40), we have,

$$\widetilde{\mathcal{T}}_2(\beta) = \frac{1}{n\cdot(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}\Big(\tau_{2F}(V_i,V_j,\beta)\Big)_+ + \frac{(n-2)}{n}\cdot\frac{3}{n}\sum_{i=1}^{n}\frac{m_{1F}^{\tau_2}(Z_i,\beta,h_n)}{h_n^{r_v}} + \xi_{c,n}^{\widetilde{\mathcal{T}}_2}(\beta),$$
$$\text{where } \sup_{\beta\in\Theta}\left|\xi_{c,n}^{\widetilde{\mathcal{T}}_2}(\beta)\right| = O_p\left(\frac{1}{n\cdot h_n^{r_v}}\right) + o_p\left(\frac{1}{n^{1/2+\epsilon}}\right) = o_p\left(\frac{1}{n^{1/2+\epsilon}}\right), \text{ uniformly over } \mathcal{F}. \tag{A-43}$$

Where $\epsilon > 0$ denotes the constant described in Assumption 4. Let us turn our attention to $m_{1F}^{\tau_2}(Z_i,\beta,h_n)$. Recall from (A-33) that,

$$\begin{aligned}\varphi_F^{\tau_2}(Z_i,Z_j,Z_k,\beta,h) \equiv\ & \delta_{a,F}^{\tau_2}(V_i,V_j,\beta)\varphi_F^{Q_2}(Y_{1k},V_k,V_j,h) - \delta_{b,F}^{\tau_2}(V_i,V_j,\beta)\varphi_F^{Q_2}(Y_{1k},V_k,V_i,h)\\ &+ \delta_{c,F}^{\tau_2}(V_i,V_j,\beta)\varphi_F^{R_2}(Y_{2k},Y_{1k},V_k,V_i,h) - \delta_{d,F}^{\tau_2}(V_i,V_j,\beta)\varphi_F^{R_2}(Y_{2k},Y_{1k},V_k,V_j,h)\end{aligned}$$

where $\varphi_F^{Q_2}$ and $\varphi_F^{R_2}$ are as described in (A-22) and $\delta_{a,F}^{\tau_2}$, $\delta_{b,F}^{\tau_2}$, $\delta_{c,F}^{\tau_2}$ and $\delta_{d,F}^{\tau_2}$ are as described in (A3.4.1). Note from (A-22) that,

$$\begin{aligned}&E_F\left[\varphi_F^{Q_2}(Y_{1k},V_k,V_j,h)|Z_i,Z_j\right] = E_F\left[\varphi_F^{Q_2}(Y_{1k},V_k,V_i,h)|Z_i,Z_j\right] = 0,\\ &E_F\left[\varphi_F^{R_2}(Y_{2k},Y_{1k},V_k,V_i,h)|Z_i,Z_j\right] = E_F\left[\varphi_F^{R_2}(Y_{2k},Y_{1k},V_k,V_j,h)|Z_i,Z_j\right] = 0.\end{aligned}$$

Thus, from the definition of $\overline{\varphi}_F^{\tau_2}$ in (A-41), we have

$$m_{1F}^{\tau_2}(Z_i,\beta,h) \equiv E_F\left[\overline{\varphi}_F^{\tau_2}(Z_i,Z_j,Z_k,\beta,h)\right] = \frac{1}{3!}\Big(E_F\left[\varphi_F^{\tau_2}(Z_j,Z_k,Z_i,\beta,h)\Big|Z_i\right] + E_F\left[\varphi_F^{\tau_2}(Z_k,Z_j,Z_i,\beta,h)\Big|Z_i\right]\Big) \tag{A-44}$$

As we defined in equation (29) prior to Assumption 2, for a given $v \equiv (x_2, w_1, w_1)$, let

$$\begin{aligned}\eta_{a,F}^{\tau_2}(v,\beta) &\equiv E_F\left[(R_{2F}(V) - (X_2'\beta_2 - x_2'\beta_2)\,Q_{2F}(V))\,\mathbb{1}\{g_{1U}(w_1,\beta_1) \le g_{1L}(W_1,\beta_1)\}\,\mathbb{1}\{\tau_{2F}(V,v,\beta)\ge 0\}\,\phi_2(V)\right],\\ \eta_{b,F}^{\tau_2}(v,\beta) &\equiv E_F\left[(R_{2F}(V) - (X_2'\beta_2 - x_2'\beta_2)\,Q_{2F}(V))\,\mathbb{1}\{g_{1U}(W_1,\beta_1) \le g_{1L}(w_1,\beta_1)\}\,\mathbb{1}\{\tau_{2F}(v,V,\beta)\ge 0\}\,\phi_2(V)\right],\\ \eta_{c,F}^{\tau_2}(v,\beta) &\equiv E_F\left[Q_{2F}(V)\mathbb{1}\{g_{1U}(W_1,\beta_1) \le g_{1L}(w_1,\beta_1)\}\,\mathbb{1}\{\tau_{2F}(v,V,\beta)\ge 0\}\,\phi_2(V)\right],\\ \eta_{d,F}^{\tau_2}(v,\beta) &\equiv E_F\left[Q_{2F}(V)\mathbb{1}\{g_{1U}(w_1,\beta_1) \le g_{1L}(W_1,\beta_1)\}\,\mathbb{1}\{\tau_{2F}(V,v,\beta)\ge 0\}\,\phi_2(V)\right]\end{aligned}$$

Using iterated expectations, we have

$$\begin{aligned}&E_F\left[\varphi_F^{\tau_2}(Z_1,Z_2,Z_3,\beta,h_n)\Big|Z_3\right] =\\ &E_F\left[\eta_{a,F}^{\tau_2}(V_2,\beta)\phi_2(V_2)\varphi_F^{Q_2}(Y_{13},V_3,V_2,h)\Big|Z_3\right] - E_F\left[\eta_{b,F}^{\tau_2}(V_1,\beta)\phi_2(V_1)\varphi_F^{Q_2}(Y_{13},V_3,V_1,h)\Big|Z_3\right]\\ &+E_F\left[\eta_{c,F}^{\tau_2}(V_1,\beta)\phi_2(V_1)\varphi_F^{R_2}(Y_{23},Y_{13},V_3,V_1,h)\Big|Z_3\right] - E_F\left[\eta_{d,F}^{\tau_2}(V_2,\beta)\phi_2(V_2)\varphi_F^{R_2}(Y_{23},Y_{13},V_3,V_2,h)\Big|Z_3\right]\end{aligned} \tag{A-45}$$

We will analyze each of the terms on the right-hand side of (A-45). Using the result in (A-35), we have

$$
\begin{aligned}
&E_F\left[\eta_{a,F}^{\tau_2}(V_2,\beta)\phi_2(V_2)\varphi_F^{Q_2}(Y_{13},V_3,V_2,h_n)\Big|Z_3\right]=\\
&E_F\left[\eta_{a,F}^{\tau_2}(V_2,\beta)\phi_2(V_2)\Gamma(V_3,V_2,h_n)\Big|V_3\right]Y_{13}\phi_2(V_3)-h_n^{r_v}\cdot E_F\left[\eta_{a,F}^{\tau_2}(V,\beta)\phi_2(V)^2 f_{V,1}(V)\right]\\
&+E_F\left[\eta_{a,F}^{\tau_2}(V,\beta)\phi_2(V)B_{n,F}^{Q_2}(V)\right].
\end{aligned}
$$

By the result shown in (A-35) and the boundedness conditions described in Assumption 2, there exists a finite constant $\overline{D}_a$ such that

$$
\sup_{\substack{v\in\mathbb{R}^{L_v}\\ \beta\in\Theta}}\left|\eta_{a,F}^{\tau_2}(v,\beta)\phi_2(v)B_{n,F}^{Q_2}(v)\right|\leq\overline{D}_a\cdot h_n^{r_v+M}\quad\forall\ F\in\mathcal{F}.
$$

Next, by the smoothness conditions in Assumption 2 and the kernel properties in Assumption 4, an $M^{th}$–order approximation implies that there exists a finite $\overline{B}_a$ such that,

$$
E_F\left[\eta_{a,F}^{\tau_2}(V_2,\beta)\phi_2(V_2)\Gamma(V_3,V_2,h_n)\Big|V_3\right]Y_{13}\phi_2(V_3)=h_n^{r_v}\cdot\eta_{a,F}^{\tau_2}(V_3,\beta)\phi_2(V_3)^2Y_{13}f_V(V_3)+B_{n,F}^{a}(V_3,\beta)Y_{13}\phi_2(V_3),
$$

$$
\text{where}\quad\sup_{\substack{v\in\mathbb{R}^{L_v}\\ \beta\in\Theta}}\left|B_{n,F}^{a}(v,\beta)\phi_2(v)\right|\leq\overline{B}_a\cdot h_n^{r_V+M}\quad\forall\ F\in\mathcal{F}.
$$

Combining these results, we obtain that, under Assumptions 2, 3 and 4, there exists a finite constant $\overline{C}$ such that,

$$
\begin{aligned}
&E_F\left[\eta_{a,F}^{\tau_2}(V_2,\beta)\phi_2(V_2)\varphi_F^{Q_2}(Y_{13},V_3,V_2,h_n)\Big|Z_3\right]=\\
&h_n^{r_v}\cdot\left(\eta_{a,F}^{\tau_2}(V_3,\beta)Y_{13}f_V(V_3)\phi_2(V_3)^2-E_F\left[\eta_{a,F}^{\tau_2}(V,\beta)f_{V,1}(V)\phi_2(V)^2\right]\right)+\xi_{a,n}(Y_{13},V_3,\beta),\\
&\text{where}\quad\sup_{\substack{v\in\mathbb{R}^{L_v}\\ \beta\in\Theta}}\left|\xi_{a,n}(Y_{13},v,\beta)\right|\leq\overline{C}\cdot h_n^{r_v+M}\quad\forall\ F\in\mathcal{F}.
\end{aligned}
$$

Note by iterated expectations that $E_F\left[\eta_{a,F}^{\tau_2}(V,\beta)f_{V,1}(V)\phi_2(V)^2\right]=E_F\left[\eta_{a,F}^{\tau_2}(V,\beta)Y_1f_V(V)\phi_2(V)^2\right]$. Therefore, the previous result becomes,

$$
\begin{aligned}
&E_F\left[\eta_{a,F}^{\tau_2}(V_2,\beta)\phi_2(V_2)\varphi_F^{Q_2}(Y_{13},V_3,V_2,h_n)\Big|Z_3\right]=\\
&h_n^{r_v}\cdot\left(\eta_{a,F}^{\tau_2}(V_3,\beta)Y_{13}f_V(V_3)\phi_2(V_3)^2-E_F\left[\eta_{a,F}^{\tau_2}(V,\beta)Y_1f_V(V)\phi_2(V)^2\right]\right)+\xi_{a,n}(Y_{13},V_3,\beta),\\
&\text{where}\quad\sup_{\substack{v\in\mathbb{R}^{L_v}\\ \beta\in\Theta}}\left|\xi_{a,n}(Y_{13},v,\beta)\right|\leq\overline{C}\cdot h_n^{r_v+M}\quad\forall\ F\in\mathcal{F}.
\end{aligned}
\tag{A-46}
$$

Analogous steps can be used to show that, under our assumptions,

$$
\begin{aligned}
&E_F\left[\eta_{b,F}^{\tau_2}(V_1,\beta)\phi_2(V_1)\varphi_F^{Q_2}(Y_{13},V_3,V_1,h_n)\Big|Z_3\right]=\\
&h_n^{r_v}\cdot\left(\eta_{b,F}^{\tau_2}(V_3,\beta)Y_{13}f_V(V_3)\phi_2(V_3)^2-E_F\left[\eta_{b,F}^{\tau_2}(V,\beta)Y_1f_V(V)\phi_2(V)^2\right]\right)+\xi_{b,n}(Y_{13},V_3,\beta),\\
&\text{where}\quad \sup_{\substack{v\in\mathbb{R}^{L_v}\\ \beta\in\Theta}}\left|\xi_{b,n}(Y_{13},v,\beta)\right|\le\overline{C}\cdot h_n^{r_v+M}\quad\forall\ F\in\mathcal{F}.
\end{aligned}
$$

Next, using again the result in (A-35), we have

$$
\begin{aligned}
&E_F\left[\eta_{c,F}^{\tau_2}(V_1,\beta)\phi_2(V_1)\varphi_F^{R_2}(Y_{23},Y_{13},V_3,V_1,h_n)\Big|Z_3\right]=\\
&E_F\left[\eta_{c,F}^{\tau_2}(V_1,\beta)\phi_2(V_1)\Gamma(V_3,V_1,h_n)\Big|V_3\right]Y_{23}Y_{13}\phi_2(V_3)-h_n^{r_v}\cdot E_F\left[\eta_{c,F}^{\tau_2}(V,\beta)\phi_2(V)^2\mu_{2F}(V)f_{V,1}(V)\right]\\
&+E_F\left[\eta_{c,F}^{\tau_2}(V,\beta)\phi_2(V)B_{n,F}^{R_2}(V)\right].
\end{aligned}
$$

By the result shown in (A-35) and the boundedness conditions described in Assumption 2, there exists a finite constant $\overline{D}_c$ such that

$$
\sup_{\substack{v\in\mathbb{R}^{L_v}\\ \beta\in\Theta}}\left|\eta_{c,F}^{\tau_2}(v,\beta)\phi_2(v)B_{n,F}^{R_2}(v)\right|\le\overline{D}_c\cdot h_n^{r_v+M}\quad\forall\ F\in\mathcal{F}.
$$

Next, by the smoothness conditions in Assumption 2 and the kernel properties in Assumption 4, an $M^{th}$–order approximation implies that there exists a finite $\overline{B}_c$ such that,

$$
\begin{aligned}
&E_F\left[\eta_{c,F}^{\tau_2}(V_1,\beta)\phi_2(V_1)\Gamma(V_3,V_1,h_n)\Big|V_3\right]Y_{23}Y_{13}\phi_2(V_3)=h_n^{r_v}\cdot\eta_{c,F}^{\tau_2}(V_3,\beta)Y_{23}Y_{13}f_V(V_3)\phi_2(V_3)^2\\
&+Y_{23}Y_{13}B_{n,F}^{c}(V_3,\beta)\phi_2(V_3),\quad\text{where}\quad\sup_{\substack{v\in\mathbb{R}^{L_v}\\ \beta\in\Theta}}\left|B_{n,F}^{c}(v,\beta)\phi_2(v)\right|\le\overline{B}_c\cdot h_n^{r_V+M}\quad\forall\ F\in\mathcal{F}.
\end{aligned}
$$

By iterated expectations, $E_F\left[\eta_{c,F}^{\tau_2}(V,\beta)\phi_2(V)^2\mu_{2F}(V)f_{V,1}(V)\right]=E_F\left[\eta_{c,F}^{\tau_2}(V,\beta)\phi_2(V)^2Y_2Y_1f_V(V)\right]$. Combining the previous results, we obtain that, under Assumptions 2, 3 and 4, there exists a finite constant $\overline{C}$ such that,

$$
\begin{aligned}
&E_F\left[\eta_{c,F}^{\tau_2}(V_1,\beta)\phi_2(V_1)\varphi_F^{R_2}(Y_{23},Y_{13},V_3,V_1,h_n)\Big|Z_3\right]=\\
&h_n^{r_v}\cdot\left(\eta_{c,F}^{\tau_2}(V_3,\beta)Y_{23}Y_{13}f_V(V_3)\phi_2(V_3)^2-E_F\left[\eta_{c,F}^{\tau_2}(V,\beta)Y_2Y_1f_V(V)\phi_2(V)^2\right]\right)+\xi_{c,n}(Y_{23},Y_{13},V_3,\beta),\\
&\text{where}\quad\sup_{\substack{v\in\mathbb{R}^{L_v}\\ \beta\in\Theta}}\left|\xi_{c,n}(Y_{23},Y_{13},v,\beta)\right|\le\overline{C}\cdot h_n^{r_v+M}\cdot|Y_{23}|\quad\forall\ F\in\mathcal{F}
\end{aligned}
$$

Analogous steps can be used to show that, under our assumptions,

$$\begin{aligned}
&E_F\left[\eta^{\tau_2}_{d,F}(V_2,\beta)\phi_2(V_2)\varphi^{R_2}_F(Y_{23},Y_{13},V_3,V_2,h)\Big|Z_3\right]=\\
&h_n^{r_v}\cdot\Big(\eta^{\tau_2}_{d,F}(V_3,\beta)Y_{23}Y_{13}f_V(V_3)\phi_2(V_3)^2-E_F\left[\eta^{\tau_2}_{d,F}(V,\beta)Y_2Y_1f_V(V)\phi_2(V)^2\right]\Big)+\xi_{d,n}(Y_{23},Y_{13},V_3,\beta),\\
&\text{where}\quad \sup_{\substack{v\in\mathbb{R}^{L_v}\\ \beta\in\Theta}}\left|\xi_{d,n}(Y_{23},Y_{13},v,\beta)\right|\leq\overline{C}\cdot h_n^{r_v+M}\cdot|Y_{23}|\quad\forall\ F\in\mathcal{F}
\end{aligned}\tag{A-47}$$

Let

$$\begin{aligned}
H^{\mathcal{T}_2}_{2F}(Z_i,\beta)\equiv&\Big(\Big(\eta^{\tau_2}_{a,F}(V_i,\beta)-\eta^{\tau_2}_{b,F}(V_i,\beta)\Big)\cdot Y_{1i}+\Big(\eta^{\tau_2}_{c,F}(V_i,\beta)-\eta^{\tau_2}_{d,F}(V_i,\beta)\Big)\cdot Y_{2i}Y_{1i}\Big)\cdot f_V(V_i)\cdot\phi_2(V_i)^2\\
&-E_F\Big[\Big(\Big(\eta^{\tau_2}_{a,F}(V,\beta)-\eta^{\tau_2}_{b,F}(V,\beta)\Big)\cdot Y_1+\Big(\eta^{\tau_2}_{c,F}(V,\beta)-\eta^{\tau_2}_{d,F}(V,\beta)\Big)\cdot Y_2Y_1\Big)\cdot f_V(V)\cdot\phi_2(V)^2\Big].
\end{aligned}\tag{A-48}$$

Note that $E_F\left[H^{\mathcal{T}_2}_{2F}(Z,\beta)\right]=0$. Combining the results in (A-46)-(A-47), we have that, under Assumptions 2, 3 and 4, there exists a finite constant $\overline{C}$ such that,

$$\begin{aligned}
&E_F\left[\varphi^{\tau_2}_F(Z_j,Z_k,Z_i,\beta,h)\Big|Z_i\right]=E_F\left[\varphi^{\tau_2}_F(Z_k,Z_j,Z_i,\beta,h)\Big|Z_i\right]=h_n^{r_v+M}\cdot H^{\mathcal{T}_2}_{2F}(Z_i,\beta)+\xi_{e,n}(Z_i,\beta),\\
&\text{where}\quad\sup_{\beta\in\Theta}\left|\xi_{e,n}(Z_i,\beta)\right|\leq\overline{C}\cdot h_n^{r_v+M}\cdot|Y_{2i}|\quad\forall\ F\in\mathcal{F}
\end{aligned}$$

Plugging this result in to (A-44), we obtain,

$$\begin{aligned}
\frac{1}{h_n^{r_v}}\cdot m^{\tau_2}_{1F}(Z_i,\beta,h)&=\frac{1}{3!}\Big(E_F\left[\varphi^{\tau_2}_F(Z_j,Z_k,Z_i,\beta,h)\Big|Z_i\right]+E_F\left[\varphi^{\tau_2}_F(Z_k,Z_j,Z_i,\beta,h)\Big|Z_i\right]\Big)\\
&=\frac{2}{3!}H^{\mathcal{T}_2}_{2F}(Z_i,\beta)+\xi_{f,n}(Z_i,\beta)\\
&=\frac{1}{3}H^{\mathcal{T}_2}_{2F}(Z_i,\beta)+\xi_{f,n}(Z_i,\beta),\\
&\text{where}\quad\sup_{\beta\in\Theta}\left|\xi_{f,n}(Z_i,\beta)\right|\leq\overline{C}\cdot h_n^{M}\cdot|Y_{2i}|\quad\forall\ F\in\mathcal{F}.
\end{aligned}\tag{A-49}$$

By Assumption 3, there exists a finite constant $\overline{D}_4$ such that $E_F[|Y_2|^4]\leq\overline{D}_4$ for all $F\in\mathcal{F}$. Therefore, using a Chebyshev inequality argument we have $\frac{1}{n}\sum_{i=1}^n|Y_{2i}|=O_p(1)$, uniformly over $\mathcal{F}$, and from the above results, we have

$$\sup_{\beta\in\Theta}\frac{1}{n}\sum_{i=1}^n\left|\xi_{e,n}(Z_i,\beta)\right|=O_p\Big(h_n^{r_v+M}\Big),\quad\text{uniformly over }\mathcal{F}.$$

From here, plugging (A-49) into (A-43), we obtain

$$\widetilde{\mathcal{T}}_2(\beta) = \frac{1}{n\cdot(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}\Big(\tau_{2F}(V_i,V_j,\beta)\Big)_+ + \frac{(n-2)}{n}\cdot\frac{1}{n}\sum_{i=1}^{n}H_{2F}^{\mathcal{T}_2}(Z_i,\beta) + \xi_{d,n}^{\widetilde{\mathcal{T}}_2}(\beta),$$
$$\text{where } \sup_{\beta\in\Theta}\left|\xi_{d,n}^{\widetilde{\mathcal{T}}_2}(\beta)\right| = O_p\left(h_n^{r_v+M}\right) + o_p\left(\frac{1}{n^{1/2+\epsilon}}\right) = o_p\left(\frac{1}{n^{1/2+\epsilon}}\right), \text{ uniformly over } \mathcal{F}. \tag{A-50}$$

Where $\epsilon > 0$ is the constant described in Assumption 4. Consider the class of functions,

$$\mathscr{H}_{3,F} \equiv \left\{m(z) = H_{2F}^{\mathcal{T}_2}(z,\beta) : \beta\in\Theta\right\}.$$

By Assumptions 2 and 3, there exist finite constants $\overline{A}_4$ and $\overline{B}_4$ such that, for all $\beta,\beta'\in\Theta$,

$$\left|H_{2F}^{\mathcal{T}_2}(z,\beta) - H_{2F}^{\mathcal{T}_2}(z,\beta')\right| \leq \left(\overline{A}_4 + \overline{B}_4\cdot|y_2|\right)\cdot\left\|\beta-\beta'\right\| \ \forall\, y_2, v, \quad \forall\, F\in\mathcal{F}.$$

From here, Pakes and Pollard (1989, Lemma 2.13) yields that there exist constants $(\overline{A},\overline{V})$ such that, for each $F\in\mathcal{F}$, the class of functions $\mathscr{H}_{3,F}$ is Euclidean $(\overline{A},\overline{V})$ for the envelope $\overline{H}_3(z) = \left|H_{2F}^{\mathcal{T}_2}(z,\beta_0)\right| + \overline{M}_3\cdot\left(\overline{A}_4 + \overline{B}_4\cdot|y_2|\right)$, where $\beta_0$ is an arbitrary point of $\Theta$ and $\overline{M}_3 \equiv 2\sqrt{k}\sup_\beta\left\|\beta-\beta_0\right\|$ (recall that $k\equiv\dim(\beta)$). By Assumptions 2 and 3, there exists a finite constant $\overline{D}_3$ such that $E_F\left[\overline{H}_3(Z)^4\right] \leq \overline{D}_3$ for all $F\in\mathcal{F}$. Thus, the conditions in Result A1 are satisfied and from there we obtain,

$$\sup_{\beta\in\Theta}\left|\frac{1}{n}\sum_{i=1}^{n}H_{2F}^{\mathcal{T}_2}(Z_i,\beta)\right| = O_p\left(\frac{1}{n^{1/2}}\right), \quad \text{uniformly over } \mathcal{F}.$$

Plugging this result into (A-50), we obtain,

$$\widetilde{\mathcal{T}}_2(\beta) = \frac{1}{n\cdot(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}\Big(\tau_{2F}(V_i,V_j,\beta)\Big)_+ + \frac{1}{n}\sum_{i=1}^{n}H_{2F}^{\mathcal{T}_2}(Z_i,\beta) + \xi_{e,n}^{\widetilde{\mathcal{T}}_2}(\beta),$$
$$\text{where} \quad \xi_{e,n}^{\widetilde{\mathcal{T}}_2}(\beta) \equiv -\left(\frac{2}{n}\right)\cdot\frac{1}{n}\sum_{i=1}^{n}H_{2F}^{\mathcal{T}_2}(Z_i,\beta) + \xi_{d,n}^{\widetilde{\mathcal{T}}_2}(\beta), \quad \text{and} \tag{A-51}$$
$$\sup_{\beta\in\Theta}\left|\xi_{e,n}^{\widetilde{\mathcal{T}}_2}(\beta)\right| = O_p\left(\frac{1}{n^{3/2}}\right) + o_p\left(\frac{1}{n^{1/2+\epsilon}}\right) = o_p\left(\frac{1}{n^{1/2+\epsilon}}\right), \text{ uniformly over } \mathcal{F},$$

where $\epsilon > 0$ is the constant described in Assumption 4. We move on to the last step and focus on $\frac{1}{n\cdot(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}\Big(\tau_{2F}(V_i,V_j,\beta)\Big)_+$ and its Hoeffding decomposition. Let $V_1, V_2$ be iid draws from $F$ and recall that we defined

$$\mathcal{T}_{2F}(\beta) \equiv E_F\left[(\tau_{2F}(V_1,V_2,\beta))_+\right].$$

Let

$$H_{1F}^{\mathcal{T}_2}(V_1,\beta) \equiv \frac{1}{2}\cdot\left(E_F\left[(\tau_{2F}(V_1,V_2,\beta))_+ \middle| V_1\right] + E_F\left[(\tau_{2F}(V_2,V_1,\beta))_+ \middle| V_1\right]\right) - \mathcal{T}_{2F}(\beta) \tag{A-52}$$

and note that $E_F\left[H_{1F}^{\mathcal{T}_2}(V,\beta)\right] = 0$. Let

$$\begin{aligned}\widetilde{g}_F^{\mathcal{T}_2}(V_1,V_2,\beta) &\equiv \left(\frac{1}{2}\cdot\left((\tau_{2F}(V_1,V_2,\beta))_+ + (\tau_{2F}(V_2,V_1,\beta))_+\right) - \mathcal{T}_{2F}(\beta)\right) - H_{1F}^{\mathcal{T}_2}(V_1,\beta) - H_{1F}^{\mathcal{T}_2}(V_2,\beta),\\ S_{2,n}^{\mathcal{T}_2}(\beta) &\equiv \binom{n}{2}^{-1}\sum_{i<j}\widetilde{g}_F^{\mathcal{T}_2}(V_i,V_j,\beta).\end{aligned} \tag{A-53}$$

The Hoeffding decomposition of $\frac{1}{n\cdot(n-1)}\sum_{i=1}^n\sum_{j\neq i}\left(\tau_{2F}(V_i,V_j,\beta)\right)_+$ yields,

$$\frac{1}{n\cdot(n-1)}\sum_{i=1}^n\sum_{j\neq i}\left(\tau_{2F}(V_i,V_j,\beta)\right)_+ = \mathcal{T}_{2F}(\beta) + \frac{2}{n}\sum_{i=1}^n H_{1F}^{\mathcal{T}_2}(V_i,\beta) + S_{2,n}^{\mathcal{T}_2}(\beta) \equiv \binom{n}{2}^{-1}\sum_{i<j}\widetilde{g}_F^{\mathcal{T}_2}(V_i,V_j,\beta). \tag{A-54}$$

We proceed by focusing on the degenerate U-process $\left\{S_{2,n}^{\mathcal{T}_2}(\beta) : \beta\in\Theta\right\}$. Fix any finite $\overline{M}$ and consider the class of functions,

$$\mathscr{H}_4^{\overline{M}} \equiv \left\{m(x_2,\widetilde{x}_2) = \alpha_1 + (x_2-\widetilde{x}_2)'\alpha_2 : \|(\alpha_1,\alpha_2')'\| \leq \overline{M}\right\}.$$

By Pakes and Pollard (1989, Example 2.9), there exist $(\overline{A},\overline{V})$ such that $\mathscr{H}_4$ is a Euclidean $(\overline{A},\overline{V})$ class of functions for envelope $\overline{H}(x_2,\widetilde{x}_2) \equiv \overline{M}\cdot(1\vee\|x_2-\widetilde{x}_2\|)$. Now let

$$\begin{aligned}&\mathscr{H}_{4,F} \equiv\\ &\left\{m(v,\widetilde{v}) = \left(R_{2F}(v)Q_{2F}(\widetilde{v}) - R_{2F}(\widetilde{v})Q_{2F}(v) - (x_2'\beta_2 - \widetilde{x}_2'\beta_2)\,Q_{2F}(v)Q_{2F}(\widetilde{v})\right)\cdot\phi_2(v)\phi_2(\widetilde{v}) : \beta_2\in\Theta\right\}.\end{aligned}$$

Assumptions 2 and 3 imply that there exists $\overline{M}<\infty$ such that $\mathscr{H}_{4,F}\subseteq\mathscr{H}_4^{\overline{M}}$ for all $F\in\mathcal{F}$. Therefore, there exist constants $(\overline{A},\overline{V})$ such that $\mathscr{H}_{4,F}$ is Euclidean $(\overline{A},\overline{V})$ for all $F\in\mathcal{F}$. Next, recall from Assumption 3 that the class of sets

$$\mathscr{C} \equiv \left\{(w_1,w_1)\in\mathbb{R}^{d_U}\times\mathbb{R}^{d_L} : g_{1U}(w_1,\beta_1) \leq g_{1L}(w_1,\beta_1) \text{ for some } \beta_1\in\Theta\right\}$$

is a VC class with VC dimension $\overline{V}_C$, and that the following is a VC class of sets for each $F$, with VC dimension uniformly bounded over $\mathcal{F}$ by a finite constant $\overline{V}_D$,

$$\mathscr{D}_{1,F}^{\tau_2} \equiv \left\{(v_1,v_2)\in\mathbb{R}^{L_v}\times\mathbb{R}^{L_v} : \tau_{2F}(v_1,v_2,\beta) \geq 0 \text{ for some } \beta\in\Theta\right\}$$

These VC properties imply, by the results in Pakes and Pollard (1989, p. 1033) (the result that classes of indicator functions over VC classes of sets are Euclidean $(A, V)$, with $(A, V)$ depending only on the VC-dimension of the underlying class of sets), and Pakes and Pollard (1989, Lemma 2.14) (the product of Euclidean classes of functions is also a Euclidean class) that there exist constants $(\overline{A}', \overline{V}')$ such that, for each $F \in \mathcal{F}$, the class of indicator functions

$$\mathscr{I}_{4,F} \equiv \Big\{ m(v, \widetilde{v}) = \mathbb{1}\{g_{1U}(\widetilde{w}_1, \beta_1) \leq g_{1L}(w_1, \beta_1)\} \cdot \mathbb{1}\{\tau_{2F}(v, \widetilde{v}, \beta) \geq 0\} \Big\},$$

is Euclidean $(\overline{A}', \overline{V}')$ for the constant envelope 1. Recall that

$$\begin{aligned} &\tau_{2F}(v, \widetilde{v}, \beta) = \\ &\Big( (R_{2F}(v)Q_{2F}(\widetilde{v}) - R_{2F}(\widetilde{v})Q_{2F}(v)) - (x_2'\beta_2 - \widetilde{x}_2'\beta_2)\, Q_{2F}(v)Q_{2F}(\widetilde{v}) \Big) \cdot \mathbb{1}\{g_{1U}(\widetilde{w}_1, \beta_1) \leq g_{1L}(w_1, \beta_1)\} \\ &\cdot \phi_2(v)\phi_2(\widetilde{v}). \end{aligned}$$

and $(\tau_{2F}(v, \widetilde{v}, \beta))_+ \equiv \tau_{2F}(v, \widetilde{v}, \beta) \cdot \mathbb{1}\{\tau_{2F}(v, \widetilde{v}, \beta) \geq 0\}$. Using the Euclidean properties of the classes of functions $\mathscr{D}_{1,F}^{\tau_2}$ and $\mathscr{H}_{4,F}$ described above, applying Pakes and Pollard (1989, Lemma 2.14), there exist constants $(\overline{A}_2, \overline{V}_2)$ such that, for each $F \in \mathcal{F}$, the class of functions

$$\mathscr{G}_F^{\tau_2} \equiv \Big\{ m(v, \widetilde{v}) = (\tau_{2F}(v, \widetilde{v}, \beta))_+ \,:\, \beta \in \Theta \Big\}$$

is Euclidean $(\overline{A}_2, \overline{V}_2)$ for an envelope of the form $G(v_1, v_2) = \overline{C}_1 + \overline{C}_2 \cdot \big\|x_2 - x_2'\big\| \cdot \phi(v)\phi(\widetilde{v})$, where $\overline{C}_1$ and $\overline{C}_2$ are finite constants. From the conditions in Assumption 2, there exists a finite constant $\overline{D}$ such that,

$$\sup_{\substack{x_2, \widetilde{x}_2 \in \mathcal{V} \times \mathcal{V} \\ \beta_2 \in \Theta}} \left| x_2'\beta_2 - \widetilde{x}_2'\beta_2 \right| \leq \overline{D}$$

Therefore, trivially there exists a constant $\overline{\mu}_4$ such that $E_F\left[G(V_1, V_2)^4\right] \leq \overline{\mu}_4 \ \forall\, F \in \mathcal{F}$, and the conditions for Result A1 are satisfied, and from there we have that the degenerate U-process $S_{2,n}^{\mathcal{T}_2}(\cdot)$ defined in (A-53) satisfies,

$$\sup_{\beta \in \Theta} \left| S_{2,n}^{\mathcal{T}_2}(\beta) \right| = O_p\left(\frac{1}{n}\right) = o_p\left(\frac{1}{n^{1/2+\epsilon}}\right), \quad \text{uniformly over } \mathcal{F}, \tag{A-55}$$

where $\epsilon > 0$ is the constant described in Assumption 4. Let $H_{2F}^{\mathcal{T}_2}(Z_i, \beta)$ be as defined in (A-48), and denote

$$\psi_F^{\mathcal{T}_2}(Z_i, \beta) \equiv 2 \cdot H_{1F}^{\mathcal{T}_2}(V_i, \beta) + H_{2F}^{\mathcal{T}_2}(Z_i, \beta). \tag{A-56}$$

Note that $E_F\left[\psi_F^{\mathcal{T}_2}(Z,\beta)\right]=0$. Plugging the result in (A-55) into (A-54) and (A-51), we obtain the linear representation result for $\widehat{\mathcal{T}}_2(\beta)$ given in part (A) of Lemma 1,

$$\widehat{\mathcal{T}}_2(\beta)=\mathcal{T}_{2F}(\beta)+\frac{1}{n}\sum_{i=1}^{n}\psi_F^{\mathcal{T}_2}(Z_i,\beta)+\xi_n^{\mathcal{T}_2}(\beta),\quad \text{where}\quad \sup_{\beta\in\Theta}\left|\xi_n^{\mathcal{T}_2}(\beta)\right|=o_p\left(\frac{1}{n^{1/2+\epsilon}}\right),\quad \text{uniformly over } \mathcal{F},$$

where $\epsilon>0$ is the constant described in Assumption 4. **This concludes the proof of part (A) of Lemma 1.** Part (B) is proved following analogous steps. Let

$$\begin{aligned}
\eta_{a,F}^{\tau_1}(w_1,\beta_1)&\equiv E_F\left[R_{1F}(W_1)\mathbb{1}\{g_{1U}(w_1,\beta_1)\leq g_{1L}(W_1,\beta_1)\}\mathbb{1}\{\tau_{1F}(W_1,w_1,\beta)\geq 0\}\phi_1(W_1)\right],\\
\eta_{b,F}^{\tau_1}(w_1,\beta_1)&\equiv E_F\left[R_{1F}(W_1)\mathbb{1}\{g_{1U}(W_1,\beta_1)\leq g_{1L}(w_1,\beta_1)\}\mathbb{1}\{\tau_{1F}(w_1,W_1,\beta)\geq 0\}\phi_1(W_1)\right],\\
\eta_{c,F}^{\tau_1}(w_1,\beta_1)&\equiv E_F\left[Q_{1F}(W_1)\mathbb{1}\{g_{1U}(W_1,\beta_1)\leq g_{1L}(w_1,\beta_1)\}\mathbb{1}\{\tau_{1F}(w_1,W_1,\beta)\geq 0\}\phi_1(W_1)\right],\\
\eta_{d,F}^{\tau_1}(w_1,\beta_1)&\equiv E_F\left[Q_{1F}(W_1)\mathbb{1}\{g_{1U}(w_1,\beta_1)\leq g_{1L}(W_1,\beta_1)\}\mathbb{1}\{\tau_{1F}(W_1,w_1,\beta)\geq 0\}\phi_1(W_1)\right]
\end{aligned}$$

and,

$$\begin{aligned}
H_{1F}^{\mathcal{T}_1}(W_1,\beta_1)&\equiv\frac{1}{2}\cdot\left(E_F\left[(\tau_{1F}(W_1,W_2,\beta_1))_+\middle|W_1\right]+E_F\left[(\tau_{1F}(W_2,W_1,\beta_1))_+\middle|W_1\right]\right)-\mathcal{T}_{1F}(\beta_1),\\
H_{2F}^{\mathcal{T}_1}(Z_i,\beta_1)&\equiv\left(\left(\eta_{a,F}^{\tau_1}(W_{1i},\beta_1)-\eta_{b,F}^{\tau_1}(W_{1i},\beta_1)\right)+\left(\eta_{c,F}^{\tau_1}(W_{1i},\beta_1)-\eta_{d,F}^{\tau_1}(W_{1i},\beta_1)\right)\cdot Y_{1i}\right)\cdot f_{W_1}(W_{1i})\cdot\phi_1(W_{1i})^2\\
&\quad-E_F\left[\left(\left(\eta_{a,F}^{\tau_1}(W_1,\beta_1)-\eta_{b,F}^{\tau_1}(W_1,\beta_1)\right)+\left(\eta_{c,F}^{\tau_1}(W_1,\beta_1)-\eta_{d,F}^{\tau_1}(W_1,\beta_1)\right)\cdot Y_1\right)\cdot f_{W_1}(W_1)\cdot\phi_1(W_1)^2\right],\\
\psi_F^{\mathcal{T}_1}(Z_i,\beta_1)&\equiv 2\cdot H_{1F}^{\mathcal{T}_1}(W_{1i},\beta_1)+H_{2F}^{\mathcal{T}_1}(Z_i,\beta_1).
\end{aligned}\tag{A-57}$$

Using parallel steps to the proof of part (A), we can show that,

$$\begin{aligned}
&\widehat{\mathcal{T}}_1(\beta_1)=\mathcal{T}_{1F}(\beta_1)+\frac{1}{n}\sum_{i=1}^{n}\psi_F^{\mathcal{T}_1}(Z_i,\beta_1)+\xi_n^{\mathcal{T}_1}(\beta_1),\quad\text{where}\\
&\sup_{\beta_1\in\Theta}\left|\xi_n^{\mathcal{T}_1}(\beta_1)\right|=o_p\left(\frac{1}{n^{1/2+\epsilon}}\right),\quad\text{uniformly over } \mathcal{F}.
\end{aligned}$$

where $\epsilon>0$ is the constant described in Assumption 4. **This is the result in part (B) of Lemma 1.** Part (C) follows immediately from (A) and (B). **This completes the proof of Lemma 1. ■**

## A4 Estimation of $\sigma_F^2(\beta)$

In this section we study the asymptotic properties of the estimator for $\sigma_F^2(\beta)\equiv E_F[\psi_F^{\mathcal{T}}(Z,\beta)^2]$ we described in Section 3.9.1 of the paper. Our construction uses the structure of the influence function $\psi_F^{\mathcal{T}}(z,\beta)$ in Lemma 1.

### A4.1 Estimation of the influence function $\psi_F^{\mathcal{T}}(z,\beta)$

We use sample analog estimators of the components described in the structure of the influence function $\psi_F^{\mathcal{T}}(z,\beta)$ in Lemma 1. We will describe separately how we estimated $\psi_F^{\mathcal{T}_2}(z,\beta)$ and $\psi_F^{\mathcal{T}_1}(z,\beta_1)$.

#### A4.1.1 Estimation of $\psi_F^{\mathcal{T}_2}(z,\beta)$

We construct our estimators using sample analogs. Based on the structure described in (A-52), for a given $(v,\beta)$, we estimate $H_{1F}^{\mathcal{T}_2}(v,\beta)$ as,

$$\widehat{H}_1^{\mathcal{T}_2}(v,\beta) \equiv \frac{1}{2}\cdot\frac{1}{n}\sum_{j=1}^{n}\Big[\widehat{\tau}_2(v,V_j,\beta)\mathbb{1}\Big\{\widehat{\tau}_2(v,V_j,\beta)\geq -b_n\Big\} + \widehat{\tau}_2(V_j,v,\beta)\mathbb{1}\Big\{\widehat{\tau}_2(V_j,v,\beta)\geq -b_n\Big\}\Big] - \widehat{\mathcal{T}}_2(\beta).$$

And, based on the structure described in (A-48), for a given $z\equiv(y_1,y_2,v)$, we estimate $H_{2F}^{\mathcal{T}_2}(z,\beta)$ as,

$$\begin{aligned}\widehat{H}_2^{\mathcal{T}_2}(z,\beta) \equiv &\Big(\Big(\widehat{\eta}_a^{\tau_2}(v,\beta)-\widehat{\eta}_b^{\tau_2}(v,\beta)\Big)\cdot y_1 + \Big(\widehat{\eta}_c^{\tau_2}(v,\beta)-\widehat{\eta}_d^{\tau_2}(v,\beta)\Big)\cdot y_2 y_1\Big)\cdot\widehat{f}_V(v)\cdot\phi_2(v)^2 \\ &-\frac{1}{n}\sum_{j=1}^{n}\Big[\Big(\Big(\widehat{\eta}_a^{\tau_2}(V_j,\beta)-\widehat{\eta}_b^{\tau_2}(V_j,\beta)\Big)\cdot Y_{1j} + \Big(\widehat{\eta}_c^{\tau_2}(V_j,\beta)-\widehat{\eta}_d^{\tau_2}(V_j,\beta)\Big)\cdot Y_{2j}Y_{1j}\Big)\cdot\widehat{f}_V(V_j)\cdot\phi_2(V_j)^2\Big].\end{aligned} \tag{A-58}$$

From here, using the definition in (A-56), for a given $z\equiv(y_1,y_2,v)$, we estimate $\psi_F^{\mathcal{T}_2}(z,\beta)$ as

$$\widehat{\psi}^{\mathcal{T}_2}(z,\beta)\equiv 2\cdot\widehat{H}_1^{\mathcal{T}_2}(v,\beta)+\widehat{H}_2^{\mathcal{T}_2}(z,\beta) \tag{A-59}$$

Let us analyze $\widehat{H}_1^{\mathcal{T}_2}(v,\beta)$ first. First, by the results in (A-17) and (A-18), we have

$$\sup_{\substack{v\in\mathbb{R}^{L_V}\\ \beta\in\Theta}}\left|\widehat{H}_1^{\mathcal{T}_2}(v,\beta)-\left(\frac{1}{2}\cdot\frac{1}{n}\sum_{j=1}^{n}\Big[\Big(\tau_{2F}(v,V_j,\beta)\Big)_+ + \Big(\tau_{2F}(V_j,v,\beta)\Big)_+\Big]-\mathcal{T}_{2F}(\beta)\right)\right| = o_p(1),$$

uniformly over $\mathcal{F}$.

As we have pointed out previously (see equation A-16), by the conditions in Assumption 2, there exists a finite constant $\overline{\tau}_2$ such that $\sup_{\substack{v,\widetilde{v}\in\mathbb{R}^{L_V}\times\mathbb{R}^{L_V}\\ \beta\in\Theta}}\Big|\tau_{2F}(v,\widetilde{v},\beta)\Big|\leq\overline{\tau}_2\ \forall\ F\in\mathcal{F}$. By a Chebyshev inequality argument, this implies

$$\sup_{\substack{v\in\mathbb{R}^{L_V}\\ \beta\in\Theta}}\left|\frac{1}{n}\sum_{j=1}^{n}\Big[\Big(\tau_{2F}(v,V_j,\beta)\Big)_+ + \Big(\tau_{2F}(V_j,v,\beta)\Big)_+\Big] - E_F\Big[(\tau_{2F}(v,V,\beta))_+ + (\tau_{2F}(V,v,\beta))_+\Big]\right| = o_p(1),$$

uniformly over $\mathcal{F}$.

Combining both previous results, we obtain

$$\sup_{\substack{v\in\mathbb{R}^{L_V}\\ \beta\in\Theta}} \left|\widehat{H}_1^{\mathcal{T}_2}(v,\beta) - H_{1F}^{\mathcal{T}_2}(v,\beta)\right| = o_p(1), \quad \text{uniformly over } \mathcal{F}. \tag{A-60}$$

Next, we analyze $\widehat{H}_2^{\mathcal{T}_2}(z,\beta)$. We begin by analyzing the estimators used in (A-58). Using the definitions in (29), we construct the estimators in on the right hand side of (A-58) as,

$$\begin{aligned}
\widehat{\eta}_a^{\tau_2}(v,\beta) &\equiv \frac{1}{n}\sum_{j=1}^{n}\Big(\widehat{R}_2(V_j) - \Big(X_{2j}'\beta_2 - x_2'\beta_2\Big)\widehat{Q}_2(V_j)\Big)\mathbb{1}\Big\{g_{1U}(w_1,\beta_1) \le g_{1L}(W_{1j},\beta_1)\Big\}\phi_2(V_j) \\
&\qquad \cdot \mathbb{1}\Big\{\widehat{\tau}_2(V_j,v,\beta) \ge -b_n\Big\}, \\
\widehat{\eta}_b^{\tau_2}(v,\beta) &\equiv \frac{1}{n}\sum_{j=1}^{n}\Big(\widehat{R}_2(V_j) - \Big(X_{2j}'\beta_2 - x_2'\beta_2\Big)\widehat{Q}_2(V_j)\Big)\mathbb{1}\Big\{g_{1U}(W_{1j},\beta_1) \le g_{1L}(w_1,\beta_1)\Big\}\phi_2(V_j) \\
&\qquad \cdot \mathbb{1}\Big\{\widehat{\tau}_2(v,V_j,\beta) \ge -b_n\Big\}, \\
\widehat{\eta}_c^{\tau_2}(v,\beta) &\equiv \frac{1}{n}\sum_{j=1}^{n}\widehat{Q}_2(V_j)\mathbb{1}\Big\{g_{1U}(W_{1j},\beta_1) \le g_{1L}(w_1,\beta_1)\Big\}\phi_2(V_j)\mathbb{1}\Big\{\widehat{\tau}_2(v,V_j,\beta) \ge -b_n\Big\}, \\
\widehat{\eta}_d^{\tau_2}(v,\beta) &\equiv \frac{1}{n}\sum_{j=1}^{n}\widehat{Q}_2(V_j)\mathbb{1}\Big\{g_{1U}(w_1,\beta_1) \le g_{1L}(W_{1j},\beta_1)\Big\}\phi_2(V_j)\mathbb{1}\Big\{\widehat{\tau}_2(V_j,v,\beta) \ge -b_n\Big\}.
\end{aligned} \tag{A-61}$$

Let

$$\begin{aligned}
\varphi^{\eta_a^{\tau_2}}(Z_i,Z_j,v,\beta,h) &\equiv \Big(Y_{2i} - \Big(X_{2j}'\beta_2 - x_2'\beta_2\Big)\Big)Y_{1i}\Gamma(V_i,V_j,h)\phi_2(V_i)\phi_2(V_j) \\
&\qquad \cdot \mathbb{1}\Big\{g_{1U}(w_1,\beta_1) \le g_{1L}(W_{1j},\beta_1)\Big\}, \\
\varphi^{\eta_b^{\tau_2}}(Z_i,Z_j,v,\beta,h) &\equiv \Big(Y_{2i} - \Big(X_{2j}'\beta_2 - x_2'\beta_2\Big)\Big)Y_{1i}\Gamma(V_i,V_j,h)\phi_2(V_i)\phi_2(V_j) \\
&\qquad \cdot \mathbb{1}\Big\{g_{1U}(W_{1j},\beta_1) \le g_{1L}(w_1,\beta_1)\Big\}, \\
\varphi^{\eta_c^{\tau_2}}(Z_i,Z_j,w_1,\beta_1,h) &\equiv \mathbb{1}\Big\{g_{1U}(W_{1j},\beta_1) \le g_{1L}(w_1,\beta_1)\Big\}Y_{1i}\Gamma(V_i,V_j,h)\phi_2(V_i)\phi_2(V_j), \\
\varphi^{\eta_d^{\tau_2}}(Z_i,Z_j,w_1,\beta_1,h) &\equiv \mathbb{1}\Big\{g_{1U}(w_1,\beta_1) \le g_{1L}(W_{1j},\beta_1)\Big\}Y_{1i}\Gamma(V_i,V_j,h)\phi_2(V_i)\phi_2(V_j),
\end{aligned}$$

From the constructions of $\widehat{R}_2$ and $\widehat{Q}_2$ (see (27)), our estimators in (A-61) are,

$$\begin{aligned}
\widehat{\eta_a^{\tau_2}}(v,\beta) &= \frac{1}{h_n^r}\cdot\frac{1}{n(n-1)}\sum_{j=1}^{n}\sum_{i\neq j}\varphi^{\eta_a^{\tau_2}}(Z_i,Z_j,v,\beta,h_n)\mathbb{1}\left\{\widehat{\tau}_2(V_j,v,\beta)\geq -b_n\right\},\\
\widehat{\eta_b^{\tau_2}}(v,\beta) &= \frac{1}{h_n^r}\cdot\frac{1}{n(n-1)}\sum_{j=1}^{n}\sum_{i\neq j}\varphi^{\eta_b^{\tau_2}}(Z_i,Z_j,v,\beta,h_n)\mathbb{1}\left\{\widehat{\tau}_2(v,V_j,\beta)\geq -b_n\right\},\\
\widehat{\eta_c^{\tau_2}}(v,\beta) &= \frac{1}{h_n^r}\cdot\frac{1}{n(n-1)}\sum_{j=1}^{n}\sum_{i\neq j}\varphi^{\eta_c^{\tau_2}}(Z_i,Z_j,w_1,\beta_1,h_n)\mathbb{1}\left\{\widehat{\tau}_2(v,V_j,\beta)\geq -b_n\right\},\\
\widehat{\eta_d^{\tau_2}}(v,\beta) &= \frac{1}{h_n^r}\cdot\frac{1}{n(n-1)}\sum_{j=1}^{n}\sum_{i\neq j}\varphi^{\eta_d^{\tau_2}}(Z_i,Z_j,w_1,\beta_1,h_n)\mathbb{1}\left\{\widehat{\tau}_2(V_j,v,\beta)\geq -b_n\right\}.
\end{aligned}\tag{A-62}$$

If Assumptions 1-5 hold, we have

$$\left.\begin{matrix}
\sup\limits_{\substack{v\in\mathcal{V}\\ \beta\in\Theta}}\left|\widehat{\eta_a^{\tau_2}}(v,\beta)-\eta_{a,F}^{\tau_2}(v,\beta)\right| = o_p(1) & \sup\limits_{\substack{v\in\mathcal{V}\\ \beta\in\Theta}}\left|\widehat{\eta_b^{\tau_2}}(v,\beta)-\eta_{a,F}^{\tau_2}(v,\beta)\right| = o_p(1)\\
\sup\limits_{\substack{v\in\mathcal{V}\\ \beta\in\Theta}}\left|\widehat{\eta_c^{\tau_2}}(v,\beta)-\eta_{a,F}^{\tau_2}(v,\beta)\right| = o_p(1) & \sup\limits_{\substack{v\in\mathcal{V}\\ \beta\in\Theta}}\left|\widehat{\eta_d^{\tau_2}}(v,\beta)-\eta_{a,F}^{\tau_2}(v,\beta)\right| = o_p(1)
\end{matrix}\right\}\text{ uniformly over }\mathcal{F}.\tag{A-63}$$

We will show the above result for $\widehat{\eta_a^{\tau_2}}(v,\beta)$. The proof for the remaining estimators in (A-63) follows analogous steps. Our first step is to express,

$$\begin{aligned}
\widehat{\eta_a^{\tau_2}}(v,\beta) &= \frac{1}{h_n^r}\cdot\frac{1}{n(n-1)}\sum_{j=1}^{n}\sum_{i\neq j}\varphi^{\eta_a^{\tau_2}}(Z_i,Z_j,v,\beta,h_n)\mathbb{1}\left\{\tau_{2F}(V_j,v,\beta)\geq 0\right\}+\xi_n^{\eta_a^{\tau_2}}(v,\beta),\quad\text{where}\\
\xi_n^{\eta_a^{\tau_2}}(v,\beta) &\equiv \frac{1}{h_n^r}\cdot\frac{1}{n(n-1)}\sum_{j=1}^{n}\sum_{i\neq j}\varphi^{\eta_a^{\tau_2}}(Z_i,Z_j,v,\beta,h_n)\Big(\mathbb{1}\left\{\widehat{\tau}_2(V_j,v,\beta)\geq -b_n\right\}-\mathbb{1}\left\{\tau_{2F}(V_j,v,\beta)\geq 0\right\}\Big)
\end{aligned}\tag{A-64}$$

We will first show that $\sup\limits_{\substack{v\in\mathcal{V}\\ \beta\in\Theta}}|\xi_n^{\eta_a^{\tau_2}}(v,\beta)| = o_p(1)$, uniformly over $\mathcal{F}$. Note first that, as we pointed out in equations (A-15) and (A-26), we have

$$\begin{aligned}
&\left|\mathbb{1}\left\{\widehat{\tau}_2(V_j,v,\beta)\geq -b_n\right\}-\mathbb{1}\left\{\tau_{2F}(V_j,v,\beta)\geq 0\right\}\right|\\
&\qquad\leq \mathbb{1}\left\{\left|\widehat{\tau}_2(V_j,v,\beta)-\tau_{2F}(V_j,v,\beta)\right|\geq b_n\right\}+\mathbb{1}\left\{-2b_n\leq\tau_{2F}(V_j,v,\beta)<0\right\}.
\end{aligned}$$

Next, recall from Assumption 2 that, there exists a finite constant $\overline{D}$ such that, $\left|x_2'\beta_2\right|\leq\overline{D}\ \forall\ (x_2,\beta_2)\in\mathcal{V}\times\Theta$. Combined with the bounded properties of the weight function $\phi_2(\cdot)$ and the kernel $K(\cdot)$, Assumption 2 implies,

$$\begin{aligned}\left|\xi_n^{\eta_a^{\tau_2}}(v,\beta)\right| \leq &\left(\frac{1}{h_n^r}\cdot\frac{1}{n}\sum_{j=1}^{n}\left(\mathbb{1}\left\{\left|\widehat{\tau}_2(V_j,v,\beta)-\tau_{2F}(V_j,v,\beta)\right|\geq b_n\right\}+\mathbb{1}\left\{-2b_n\leq\tau_{2F}(V_j,v,\beta)<0\right\}\right)\right)\\ &\times\overline{\phi}^2\overline{K}\left(\frac{1}{n-1}\sum_{i\neq j}|Y_{2i}|+2\overline{D}\right)\end{aligned}\tag{A-65}$$

By Assumption 3, there exists $\overline{D}_4<\infty$ such that $E_F[|Y_2|^4]\leq\overline{D}_4$ for all $F\in\mathcal{F}$. Therefore, a Chebyshev inequality argument yields,

$$\frac{1}{n-1}\sum_{i\neq j}|Y_{2i}|=O_p(1),\quad\text{uniformly over }\mathcal{F}.\tag{A-66}$$

Take any $\delta>0$, note that

$$P_F\left(\sup_{\substack{v\in\mathcal{V}\\ \beta\in\Theta}}\left|\frac{1}{h_n^r}\cdot\frac{1}{n}\sum_{j=1}^{n}\mathbb{1}\left\{\left|\widehat{\tau}_2(V_j,v,\beta)-\tau_{2F}(V_j,v,\beta)\right|\geq b_n\right\}\right|>\delta\right)\leq P_F\left(\sup_{\substack{v,\widetilde{v}\in\mathbb{R}^{L_V}\times\mathbb{R}^{L_V}\\ \beta\in\Theta}}\left|\widehat{\tau}_2(v,\widetilde{v},\beta)-\tau_{2F}(v,\widetilde{v},\beta)\right|>b_n\right)$$

From equation (A-12),

$$\sup_{F\in\mathcal{F}}P_F\left(\sup_{\substack{v,\widetilde{v}\in\mathbb{R}^{L_V}\times\mathbb{R}^{L_V}\\ \beta\in\Theta}}\left|\widehat{\tau}_2(v,\widetilde{v},\beta)-\tau_{2F}(v,\widetilde{v},\beta)\right|>b_n\right)\longrightarrow 0.$$

Therefore,

$$\frac{1}{h_n^r}\cdot\frac{1}{n}\sum_{j=1}^{n}\mathbb{1}\left\{\left|\widehat{\tau}_2(V_j,v,\beta)-\tau_{2F}(V_j,v,\beta)\right|\geq b_n\right\}=o_p(1),\quad\text{uniformly over }\mathcal{F}.\tag{A-67}$$

Next, for a given $(v,\beta)$ and $c>0$, let

$$\overline{m}_F^{\eta_a^{\tau_2}}(v,\beta,c)\equiv\frac{1}{n}\sum_{j=1}^{n}\left(\mathbb{1}\left\{-c\leq\tau_{2F}(V_j,v,\beta)<0\right\}-E_F\left[\mathbb{1}\left\{-c\leq\tau_{2F}(V_j,v,\beta)<0\right\}\right]\right).$$

Note that,

$$\frac{1}{h_n^r}\cdot\frac{1}{n}\sum_{j=1}^{n}\mathbb{1}\left\{-2b_n\leq\tau_{2F}(V_j,v,\beta)<0\right\}=\frac{1}{h_n^r}\cdot\overline{m}_F^{\eta_a^{\tau_2}}(v,\beta,2b_n)+\frac{1}{h_n^r}E_F\left[\mathbb{1}\left\{-2b_n\leq\tau_{2F}(V,v,\beta)<0\right\}\right].\tag{A-68}$$

By the properties of VC classes of sets described, e.g, in Pakes and Pollard (1989, Lemma 2.5), the conditions described in Assumption 3 imply that, for each $F \in \mathcal{F}$, the following class of sets is a VC class, with VC dimension uniformly bounded over $\mathcal{F}$ by a finite constant,

$$\mathscr{C}_{2,F}^{\tau_2} \equiv \left\{ v \in \mathbb{R}^{L_v} : -c \leq \tau_{2F}(v,u,\beta) < 0 \ \text{ for some } 0 < c \leq c_0,\ u \in \mathcal{V}, \text{ and } \beta \in \Theta \right\},$$

where the constant $c_0$ is as described in Assumption 3. From here, the result in Pakes and Pollard (1989, p. 1033) implies that there exist constants $(\overline{A}, \overline{V})$ such that, for each $F \in \mathcal{F}$, the class of indicator functions,

$$\mathscr{H}_F \equiv \left\{ m(u) = \mathbb{1}\{-c \leq \tau_{2F}(v,u,\beta) < 0\} \ \text{ for some } 0 < c \leq c_0,\ u \in \mathcal{V} \text{ and } \beta \in \Theta \right\}$$

is Euclidean $(\overline{A}, \overline{V})$ for the constant envelope 1. From here and Sherman (1994, Lemma 5), the conditions for Result A1 are satisfied and, from there, we obtain,

$$\sup_{\substack{\beta\in\Theta \\ v\in\mathcal{V} \\ 0<c\leq c_0}} \left| \frac{1}{n} \sum_{i=1}^{n} \overline{m}_F^{\eta_a^{\tau_2}}(v,\beta,c) \right| = O_p\left(\frac{1}{n^{1/2}}\right), \quad \text{uniformly over } \mathcal{F}.$$

For $n$ large enough, we have $2b_n \leq c_0$. Therefore, by the above result and the condition in part (ii) of Assumption 5, equation (A-68) yields,

$$\frac{1}{h_n^r} \cdot \sup_{\substack{v\in\mathcal{V} \\ \beta\in\Theta}} \frac{1}{n} \sum_{j=1}^{n} \mathbb{1}\left\{-2b_n \leq \tau_{2F}(V_j,v,\beta) < 0\right\} \leq O_p\left(\frac{1}{h_n^r \cdot n^{1/2}}\right) + 2\overline{m} \cdot \frac{b_n}{h_n^r} = o_p(1), \quad \text{uniformly over } \mathcal{F} \tag{A-69}$$

where $\overline{m}$ is the constant described in Assumption 5. The last line follows from the bandwidth convergence conditions in Assumption 4, which require $h_n^r \cdot n^{1/2} \to \infty$ and $\frac{b_n}{h_n^r} \to 0$. Combining (A-65), (A-66), (A-67), and (A-69), we have

$$\sup_{\substack{v\in\mathcal{V} \\ \beta\in\Theta}} \left| \xi_n^{\eta_a^{\tau_2}}(v,\beta) \right| = o_p(1), \quad \text{uniformly over } \mathcal{F}.$$

Plugging this into (A-64), we obtain,

$$\begin{aligned} &\widehat{\eta_a^{\tau_2}}(v,\beta) = \frac{1}{h_n^r} \cdot \frac{1}{n(n-1)} \sum_{j=1}^{n} \sum_{i\neq j} \varphi^{\eta_a^{\tau_2}}(Z_i,Z_j,v,\beta,h_n) \mathbb{1}\left\{\tau_{2F}(V_j,v,\beta) \geq 0\right\} + \xi_n^{\eta_a^{\tau_2}}(v,\beta), \\ &\text{where} \quad \sup_{\substack{v\in\mathcal{V} \\ \beta\in\Theta}} \left| \xi_n^{\eta_a^{\tau_2}}(v,\beta) \right| = o_p(1), \quad \text{uniformly over } \mathcal{F}. \end{aligned} \tag{A-70}$$

Next, let

$$U_{n,F}^{\eta_a^{\tau_2}}(v,\beta,h) \equiv$$
$$\frac{1}{n(n-1)}\sum_{j=1}^{n}\sum_{i\neq j}\Big(\varphi^{\eta_a^{\tau_2}}(Z_i,Z_j,v,\beta,h)\mathbb{1}\Big\{\tau_{2F}(V_j,v,\beta)\geq 0\Big\}-E_F\Big[\varphi^{\eta_a^{\tau_2}}(Z_i,Z_j,v,\beta,h)\mathbb{1}\Big\{\tau_{2F}(V_j,v,\beta)\geq 0\Big\}\Big]\Big).$$

We can rewrite (A-70) as,

$$\widehat{\eta_a^{\tau_2}}(v,\beta)=\frac{1}{h_n^r}\cdot U_{n,F}^{\eta_a^{\tau_2}}(v,\beta,h_n)+\frac{1}{h_n^r}\cdot E_F\Big[\varphi^{\eta_a^{\tau_2}}(Z_i,Z_j,v,\beta,h_n)\mathbb{1}\Big\{\tau_{2F}(V_j,v,\beta)\geq 0\Big\}\Big]+\xi_n^{\eta_a^{\tau_2}}(v,\beta) \tag{A-71}$$

where, in the above expectation, $Z_i, Z_j$ are two independent draws from $F$. We will analyze $U_{n,F}^{\eta_a^{\tau_2}}(v,\beta,h_n)$ first. Define the class of functions,

$$\mathscr{H}_F^{\eta_a^{\tau_2}}\equiv\Big\{m(z_1,z_2)=\varphi^{\eta_a^{\tau_2}}(z_1,z_2,u,\beta,h)\mathbb{1}\{\tau_{2F}(v_2,u,\beta)\geq 0\}\ \text{ for some } u\in\mathcal{V},\ \beta\in\Theta,\ h>0\Big\}$$

Invoking arguments and results from empirical process theory we have used previously, the smoothness, regularity and manageability conditions in Assumptions 2 and 3, and the bounded-variation properties of the kernel described in Assumption 2 imply, by Pakes and Pollard (1989, Lemma 2.14), that there exist constants $(\overline{A}_2,\overline{V}_2)$ such that, for each $F\in\mathcal{F}$, the class of functions $\mathscr{H}_F^{\eta_a^{\tau_2}}$ is Euclidean $(\overline{A}_2,\overline{V}_2)$ for an envelope $\overline{G}_2(z_1,z_2)$ such that there exists a constant $\overline{C}_2<\infty$ for which $E_F[\overline{G}_2(Z_1,Z_2)^4]\leq\overline{C}_2$ for all $F\in\mathcal{F}$. Thus, the conditions in Result A1 are satisfied and from there we obtain,

$$\sup_{\substack{\beta\in\Theta\\ v\in\mathcal{V}}}\Big|U_{n,F}^{\eta_a^{\tau_2}}(v,\beta,h_n)\Big|=O_p\Big(\frac{1}{n^{1/2}}\Big),\quad\text{uniformly over }\mathcal{F} \tag{A-72}$$

Next, using an $M^{th}$–order approximation, the smoothness conditions in Assumption 2, and the bias-reducing properties of the kernel described in Assumption 4 imply that there exists a constant $\overline{B}^{\eta_a^{\tau_2}}<\infty$ such that,

$$\begin{aligned}&\frac{1}{h_n^r}\cdot E_F\Big[\varphi^{\eta_a^{\tau_2}}(Z_i,Z_j,v,\beta,h_n)\mathbb{1}\Big\{\tau_{2F}(V_j,v,\beta)\geq 0\Big\}\Big]=\eta_{a,F}^{\tau_2}(v,\beta)+\underbrace{B_n^{\eta_a^{\tau_2}}(v,\beta)}_{\text{bias}},\\ &\text{where}\quad\sup_{\substack{\beta\in\Theta\\ v\in\mathcal{V}}}\Big|B_n^{\eta_a^{\tau_2}}(v,\beta)\Big|\leq\overline{B}^{\eta_a^{\tau_2}}\cdot h_n^M\quad\forall\ F\in\mathcal{F}\end{aligned} \tag{A-73}$$

Plugging (A-72) and (A-73) into (A-71), we obtain

$$\begin{aligned}\sup_{\substack{\beta\in\Theta\\ v\in\mathcal{V}}}\left|\widehat{\eta}_a^{\tau_2}(v,\beta)-\eta_{a,F}^{\tau_2}(v,\beta)\right| &\le \frac{1}{h_n^r}\cdot\sup_{\substack{\beta\in\Theta\\ v\in\mathcal{V}}}\left|U_{n,F}^{\eta_a^{\tau_2}}(v,\beta,h_n)\right|+\sup_{\substack{\beta\in\Theta\\ v\in\mathcal{V}}}\left|B_n^{\eta_a^{\tau_2}}(v,\beta)\right|\\ &= O_p\left(\frac{1}{h_n^r\cdot n^{1/2}}\right)+O\left(h_n^M\right)=o_p(1),\quad \text{uniformly over } \mathcal{F}.\end{aligned}$$

which proves the claim in (A-63) for $\widehat{\eta}_a^{\tau_2}(v,\beta)$. Using our assumptions, proving the claim in (A-63) for $\widehat{\eta}_b^{\tau_2}(v,\beta)$, $\widehat{\eta}_c^{\tau_2}(v,\beta)$ and $\widehat{\eta}_d^{\tau_2}(v,\beta)$ follows analogous steps.

Let us continue with $\widehat{f}_V(v)$, which is also used in (A-58). As we have detailed before, for a given $v$, we have

$$\widehat{f}_V(v)\equiv\frac{1}{h_n^r}\cdot\frac{1}{n}\sum_{i=1}^{n}\Gamma(V_i,v,h_n).$$

A result we have used previously is that, by Nolan and Pollard (1987, Lemma 22) (or Pakes and Pollard (1989, Example 10)), the bounded variation nature of our kernel implies that the class of functions $\left\{m(v)=k\left(\frac{v-u}{h}\right)\text{ for some } u\in\mathbb{R},\ h>0\right\}$ is Euclidean $(A_k,V_k)$ for the constant envelope $\overline{k}$ (neither $(A_k,V_k)$, nor $\overline{k}$ depend on $F$). From here and Sherman (1994, Lemma 5), the following empirical process satisfies the conditions of Result A1,

$$\nu_n^{f_V}(v)\equiv\frac{1}{n}\sum_{i=1}^{n}\left(\Gamma(V_i,v,h_n)-E_F\left[\Gamma(V_i,v,h_n)\right]\right),$$

and we have, $\sup_{v\in\mathbb{R}^{L_V}}\left|\nu_n^{f_V}(v)\right|=O_p\left(\frac{1}{n^{1/2}}\right)$, uniformly over $\mathcal{F}$. Next, using an $M^{th}$–order approximation, the smoothness conditions in Assumption 2, and the bias-reducing properties of the kernel described in Assumption 4 imply that there exists a constant $\overline{B}^{f_v}<\infty$ such that,

$$\begin{aligned}&\frac{1}{h_n^r}\cdot E_F\left[\Gamma(V_i,v,h_n)\right]=f_V(v)+\underbrace{B_n^{f_V}(v)}_{\text{bias}},\\ &\text{where}\quad \sup_{v\in\mathcal{V}}\left|B_n^{f_V}(v)\right|\le\overline{B}^{f_v}\cdot h_n^M\quad\forall\ F\in\mathcal{F}\end{aligned}$$

Combining these results, we have

$$\begin{aligned}\sup_{v\in\mathcal{V}}\left|\widehat{f}_V(v)-f_V(v)\right| &\le \frac{1}{h_n^r}\cdot\sup_{v\in\mathcal{V}}\left|\nu_n^{f_V}(v)\right|+\sup_{v\in\mathcal{V}}\left|B_n^{f_V}(v)\right|\\ &= O_p\left(\frac{1}{h_n^r\cdot n^{1/2}}\right)+O\left(h_n^M\right)=o_p(1),\quad \text{uniformly over } \mathcal{F}.\end{aligned}\tag{A-74}$$

Plugging in the results in (A-63) and (A-74) into (A-58), for any $y_1, y_2$, we have[18]

$$\sup_{\substack{\beta\in\Theta\\ v\in\mathbb{R}^{L_V}}} \Bigg|\widehat{H}_2^{\mathcal{T}_2}(z,\beta) - \Bigg\{\bigg(\Big(\eta_{a,F}^{\tau_2}(v,\beta) - \eta_{b,F}^{\tau_2}(v,\beta)\Big)\cdot y_1 + \Big(\eta_{c,F}^{\tau_2}(v,\beta) - \eta_{d,F}^{\tau_2}(v,\beta)\Big)\cdot y_2 y_1\bigg)\cdot f_V(v)\cdot\phi_2(v)^2$$

$$-\frac{1}{n}\sum_{j=1}^{n}\bigg[\bigg(\Big(\eta_{a,F}^{\tau_2}(V_j,\beta) - \eta_{b,F}^{\tau_2}(V_j,\beta)\Big)\cdot Y_{1j} + \Big(\eta_{c,F}^{\tau_2}(V_j,\beta) - \eta_{d,F}^{\tau_2}(V_j,\beta)\Big)\cdot Y_{2j}Y_{1j}\bigg)\cdot f_V(V_j)\cdot\phi_2(V_j)^2\bigg]\Bigg\}\Bigg| = o_p(1),$$

uniformly over $\mathcal{F}$.

(A-75)

By the conditions of Assumption 2, there exists a $\overline{\mu}_4^{\tau_2}$ such that,

$$\sup_{\beta\in\Theta} E_F\left[\left|\bigg(\Big(\eta_{a,F}^{\tau_2}(V,\beta) - \eta_{b,F}^{\tau_2}(V,\beta)\Big)\cdot Y_1 + \Big(\eta_{c,F}^{\tau_2}(V,\beta) - \eta_{d,F}^{\tau_2}(V,\beta)\Big)\cdot Y_2 Y_1\bigg)\cdot f_V(V)\cdot\phi_2(V)^2\right|^4\right] \leq \overline{\mu}_4^{\tau_2}$$

$\forall\ F\in\mathcal{F}$.

From here, a Chebyshev inequality argument yields,

$$\sup_{\beta\in\Theta}\Bigg|\frac{1}{n}\sum_{j=1}^{n}\bigg[\bigg(\Big(\eta_{a,F}^{\tau_2}(V_j,\beta) - \eta_{b,F}^{\tau_2}(V_j,\beta)\Big)\cdot Y_{1j} + \Big(\eta_{c,F}^{\tau_2}(V_j,\beta) - \eta_{d,F}^{\tau_2}(V_j,\beta)\Big)\cdot Y_{2j}Y_{1j}\bigg)\cdot f_V(V_j)\cdot\phi_2(V_j)^2\bigg]$$

$$- E_F\bigg[\bigg(\Big(\eta_{a,F}^{\tau_2}(V,\beta) - \eta_{b,F}^{\tau_2}(V,\beta)\Big)\cdot Y_1 + \Big(\eta_{c,F}^{\tau_2}(V,\beta) - \eta_{d,F}^{\tau_2}(V,\beta)\Big)\cdot Y_2 Y_1\bigg)\cdot f_V(V)\cdot\phi_2(V)^2\bigg] = o_p(1),$$

uniformly over $\mathcal{F}$.

Plugging in this result into (A-75), we have that for any $y_1, y_2$,

$$\sup_{\substack{\beta\in\Theta\\ v\in\mathbb{R}^{L_V}}}\left|\widehat{H}_2^{\mathcal{T}_2}(z,\beta) - H_{2F}^{\mathcal{T}_2}(z,\beta)\right| = o_p(1), \quad \text{uniformly over } \mathcal{F}. \tag{A-76}$$

Combining (A-60) and (A-76) with the definition of $\widehat{\psi}^{\mathcal{T}_2}(z,\beta)$ in (A-59), for any $y_1, y_2$, under Assumptions 1-5, we have

$$\sup_{\substack{\beta\in\Theta\\ v\in\mathbb{R}^{L_V}}}\left|\widehat{\psi}^{\mathcal{T}_2}(z,\beta) - \psi_F^{\mathcal{T}_2}(z,\beta)\right| = o_p(1), \quad \text{uniformly over } \mathcal{F} \tag{A-77}$$

[18]Note that the presence of the weight function $\phi_2(v)$, which is zero for all $v\notin\mathcal{V}$, implies that the results in (A-63) and (A-74), which hold uniformly over $v\in\mathcal{V}$, immediately produce the result in (A-75), which holds uniformly over $v\in\mathbb{R}^{L_V}$ (since any $v\notin\mathcal{V}$ is trimmed away by $\phi_2(\cdot)$).

#### A4.1.2 Estimation of $\psi_F^{\mathcal{T}_1}(z,\beta_1)$

As in our estimation of $\psi_F^{\mathcal{T}_2}(z,\beta)$, we proceed using sample analogs based on the definition of $\psi_F^{\mathcal{T}_1}(z,\beta_1)$. Based on the structure described in (A-57), for a given $(w_1,\beta_1)$, we estimate $H_{1F}^{\mathcal{T}_1}(w_1,\beta_1)$ as,

$$\widehat{H}_1^{\mathcal{T}_1}(w_1,\beta_1) \equiv \frac{1}{2}\cdot\frac{1}{n}\sum_{j=1}^{n}\Big[\widehat{\tau}_1(w_1,W_{1j},\beta_1)\mathbb{1}\Big\{\widehat{\tau}_1(w_1,W_{1j},\beta_1)\geq -b_n\Big\}+\widehat{\tau}_1(W_{1j},w_1,\beta_1)\mathbb{1}\Big\{\widehat{\tau}_1(W_{1j},w_1,\beta_1)\geq -b_n\Big\}\Big] - \widehat{\mathcal{T}}_1(\beta_1).$$

And, for a given $z\equiv(y_1,y_2,w_1)$, we estimate $H_{2F}^{\mathcal{T}_1}(z,\beta_1)$ as,

$$\begin{aligned}\widehat{H}_2^{\mathcal{T}_1}(z,\beta_1) &\equiv \Big(\Big(\widehat{\eta}_a^{\tau_1}(w_1,\beta_1)-\widehat{\eta}_b^{\tau_1}(w_1,\beta_1)\Big)+\Big(\widehat{\eta}_c^{\tau_1}(w_1,\beta_1)-\widehat{\eta}_d^{\tau_1}(w_1,\beta_1)\Big)\cdot y_1\Big)\cdot\widehat{f}_{W_1}(w_1)\cdot\phi_1(w_1)^2\\ &-\frac{1}{n}\sum_{j=1}^{n}\Big[\Big(\Big(\widehat{\eta}_a^{\tau_1}(W_{1j},\beta_1)-\widehat{\eta}_b^{\tau_1}(W_{1j},\beta_1)\Big)+\Big(\widehat{\eta}_c^{\tau_1}(W_{1j},\beta_1)-\widehat{\eta}_d^{\tau_1}(W_{1j},\beta_1)\Big)\cdot Y_{1j}\Big)\cdot\widehat{f}_{W_1}(W_{1j})\cdot\phi_1(W_{1j})^2\Big]\end{aligned} \tag{A-78}$$

From here, using the definition in (A-57), for a given $z$, we estimate $\psi_F^{\mathcal{T}_1}(z,\beta_1)$ as

$$\widehat{\psi}^{\mathcal{T}_1}(z,\beta_1)\equiv 2\cdot\widehat{H}_1^{\mathcal{T}_1}(w_1,\beta_1)+\widehat{H}_2^{\mathcal{T}_1}(z,\beta_1)$$

Using the definitions in (30), we construct the estimators on the right hand side of (A-78) as,

$$\begin{aligned}\widehat{\eta}_a^{\tau_1}(w_1,\beta_1) &\equiv \frac{1}{n}\sum_{j=1}^{n}\widehat{R}_1(W_{1j})\mathbb{1}\Big\{g_{1U}(w_1,\beta_1)\leq g_{1L}(W_{1j},\beta_1)\Big\}\mathbb{1}\Big\{\widehat{\tau}_1(W_{1j},w_1,\beta)\geq -b_n\Big\}\phi_1(W_{1j}),\\ \widehat{\eta}_b^{\tau_1}(w_1,\beta_1) &\equiv \frac{1}{n}\sum_{j=1}^{n}\widehat{R}_1(W_{1j})\mathbb{1}\Big\{g_{1U}(W_{1j},\beta_1)\leq g_{1L}(w_1,\beta_1)\Big\}\mathbb{1}\Big\{\widehat{\tau}_1(w_1,W_{1j},\beta)\geq -b_n\Big\}\phi_1(W_{1j}),\\ \eta_{c,F}^{\tau_1}(w_1,\beta_1) &\equiv \frac{1}{n}\sum_{j=1}^{n}\widehat{Q}_1(W_{1j})\mathbb{1}\Big\{g_{1U}(W_{1j},\beta_1)\leq g_{1L}(w_1,\beta_1)\Big\}\mathbb{1}\Big\{\widehat{\tau}_1(w_1,W_{1j},\beta)\geq -b_n\Big\}\phi_1(W_{1j}),\\ \eta_{d,F}^{\tau_1}(w_1,\beta_1) &\equiv \frac{1}{n}\sum_{j=1}^{n}\widehat{Q}_1(W_{1j})\mathbb{1}\Big\{g_{1U}(w_1,\beta_1)\leq g_{1L}(W_{1j},\beta_1)\Big\}\mathbb{1}\Big\{\widehat{\tau}_1(W_{1j},w_1,\beta)\geq -b_n\Big\}\phi_1(W_{1j})\end{aligned} \tag{A-79}$$

Let

$$\begin{aligned}\varphi^{\eta_a^{\tau_1}}(Z_i,Z_j,w_1,\beta_1,h) &\equiv \mathbb{1}\Big\{g_{1U}(w_1,\beta_1)\leq g_{1L}(W_{1j},\beta_1)\Big\}Y_{1i}\Gamma(W_{1i},W_{1j},h)\phi_1(W_{1i})\phi_1(W_{1j}),\\ \varphi^{\eta_b^{\tau_1}}(Z_i,Z_j,w_1,\beta_1,h) &\equiv \mathbb{1}\Big\{g_{1U}(W_{1j},\beta_1)\leq g_{1L}(w_1,\beta_1)\Big\}Y_{1i}\Gamma(W_{1i},W_{1j},h)\phi_1(W_{1i})\phi_1(W_{1j}),\\ \varphi^{\eta_c^{\tau_1}}(Z_i,Z_j,w_1,\beta_1,h) &\equiv \mathbb{1}\Big\{g_{1U}(W_{1j},\beta_1)\leq g_{1L}(w_1,\beta_1)\Big\}\Gamma(W_{1i},W_{1j},h)\phi_1(W_{1i})\phi_1(W_{1j}),\\ \varphi^{\eta_d^{\tau_1}}(Z_i,Z_j,w_1,\beta_1,h) &\equiv \mathbb{1}\Big\{g_{1U}(w_1,\beta_1)\leq g_{1L}(W_{1j},\beta_1)\Big\}\Gamma(W_{1i},W_{1j},h)\phi_1(W_{1i})\phi_1(W_{1j}).\end{aligned}$$

From the constructions of $\widehat{R}_1$ and $\widehat{Q}_1$ (see (27)), our estimators in (A-79) are,

$$\widehat{\eta}_a^{\tau_1}(w_1,\beta_1) = \frac{1}{h_n^\ell}\cdot\frac{1}{n(n-1)}\sum_{j=1}^n\sum_{i\neq j}\varphi^{\eta_a^{\tau_1}}(Z_i,Z_j,w_1,\beta_1,h_n)\mathbb{1}\left\{\widehat{\tau}_1(W_{1j},w_1,\beta_1)\geq -b_n\right\},$$
$$\widehat{\eta}_b^{\tau_1}(w_1,\beta_1) = \frac{1}{h_n^\ell}\cdot\frac{1}{n(n-1)}\sum_{j=1}^n\sum_{i\neq j}\varphi^{\eta_b^{\tau_1}}(Z_i,Z_j,w_1,\beta_1,h_n)\mathbb{1}\left\{\widehat{\tau}_1(w_1,W_{1j},\beta_1)\geq -b_n\right\},$$
$$\widehat{\eta}_c^{\tau_1}(w_1,\beta_1) = \frac{1}{h_n^\ell}\cdot\frac{1}{n(n-1)}\sum_{j=1}^n\sum_{i\neq j}\varphi^{\eta_c^{\tau_1}}(Z_i,Z_j,w_1,\beta_1,h_n)\mathbb{1}\left\{\widehat{\tau}_1(w_1,W_{1j},\beta_1)\geq -b_n\right\},$$
$$\widehat{\eta}_d^{\tau_1}(w_1,\beta_1) = \frac{1}{h_n^\ell}\cdot\frac{1}{n(n-1)}\sum_{j=1}^n\sum_{i\neq j}\varphi^{\eta_d^{\tau_1}}(Z_i,Z_j,w_1,\beta_1,h_n)\mathbb{1}\left\{\widehat{\tau}_1(W_{1j},w_1,\beta_1)\geq -b_n\right\}.$$

The above expressions are equivalent to those in (A-62). From here, using analogous arguments to those we used in the steps from equation (A-60) to the final result in equation (A-77), we can show that, for any $y_1$, under Assumptions 1-5, we have

$$\sup_{\substack{\beta_1\in\Theta\\ v\in\mathbb{R}^{L_V}}}\left|\widehat{\psi}^{\mathcal{T}_1}(z,\beta_1)-\psi_F^{\mathcal{T}_1}(z,\beta_1)\right| = o_p(1),\quad \text{uniformly over } \mathcal{F} \tag{A-80}$$

### A4.2 Estimation of $\psi_F^{\mathcal{T}}(z,\beta)$

The influence function $\psi_F^{\mathcal{T}}(z,\beta)$ is defined in Lemma 1 as $\psi_F^{\mathcal{T}}(z,\beta)\equiv\psi_F^{\mathcal{T}_2}(z,\beta)+\psi_F^{\mathcal{T}_1}(z,\beta_{1)})$. Accordingly, we estimate it as $\widehat{\psi}^{\mathcal{T}}(z,\beta)\equiv\widehat{\psi}^{\mathcal{T}_2}(z,\beta)+\widehat{\psi}^{\mathcal{T}_1}(z,\beta_{1)})$. From the results in (A-77) and (A-80), for any $y_1, y_2$, we have

$$\sup_{\substack{\beta\in\Theta\\ v\in\mathbb{R}^{L_V}}}\left|\widehat{\psi}^{\mathcal{T}}(z,\beta)-\psi_F^{\mathcal{T}}(z,\beta)\right| = o_p(1),\quad \text{uniformly over } \mathcal{F} \tag{A-81}$$

### A4.3 Our estimator for $\sigma_F^2(\beta)$

We estimate $\sigma_F^2(\beta)\equiv E_F[\psi_F^{\mathcal{T}}(Z,\beta)^2]$ as

$$\widehat{\sigma}^2(\beta)\equiv\frac{1}{n}\sum_{i=1}^n\widehat{\psi}^{\mathcal{T}}(Z_i,\beta)^2$$

Recall that $Y_{1i}\in\{0,1\}$ and also recall that, by Assumption 3, there exists a finite constant $\overline{D}_4$ such that $E_F[|Y_2|^4]\leq\overline{D}_4$ for all $F\in\mathcal{F}$. Combining this with the result in (A-81), we obtain that, under Assumptions 1-5,

$$\sup_{\beta\in\Theta}\left|\widehat{\sigma}^2(\beta)-\sigma_F^2(\beta)\right| = o_p(1),\quad \text{uniformly over } \mathcal{F}.$$

This proves the claim in equation (40) in the paper. ■

## A5 Conditions under which we can let $\kappa_n \to 0$

Assumption 6 allows for $\sigma_F^2(\beta)$ (the relevant measure of the contact sets in our problem) to become arbitrarily close to zero over $(\Theta \times \mathcal{F}) \setminus \overline{\Lambda}_{\Theta,\mathcal{F}}$. If we strengthen Assumption 6 to assume now that $\sigma_F^2(\beta)$ is bounded away from zero uniformly over $(\Theta \times \mathcal{F}) \setminus \overline{\Lambda}_{\Theta,\mathcal{F}}$, we can replace our regularization parameter $\kappa$ with a positive sequence that vanishes asymptotically.

### A5.1 A stronger version of Assumption 6

Suppose we replace Assumption 6 with the following stronger restriction.

**Assumption 6'** ***(A stronger version of Assumption 6)*** *There exist a $B < \infty$ and $\underline{C} > 0$ such that,*

$$E_F[|\psi_F^{\mathcal{T}}(Z_i,\beta)|^3] \leq B, \quad \textit{and} \quad \sigma_F^2(\beta) \geq \underline{C} \quad \forall\, (\beta,F) \in (\Theta \times \mathcal{F}) \setminus \overline{\Lambda}_{\Theta,\mathcal{F}} \quad \blacksquare$$

The Berry-Esseen condition produced by Assumption 6, and the results in Theorem 1 still hold under the stronger restrictions of Assumption 6', but we now also have the following result. Take any positive sequence $\kappa_n \to 0$ such that $\kappa_n \cdot n^{\epsilon} \to \infty$, with $\epsilon > 0$ being the constant described in Assumption 4. Note from (35) that,

$$\sup_{(\beta,F)\in\Theta\times\mathcal{F}} \left| \frac{n^{1/2} \cdot \xi_n^{\mathcal{T}}(\beta)}{(\sigma_F(\beta) \vee \kappa_n)} \right| = o_p\left( \frac{1}{\kappa_n \cdot n^{\epsilon}} \right) = o_p(1). \tag{A-82}$$

If Assumption 6' holds, then for $n$ large enough we have $(\sigma_F(\beta) \vee \kappa_n) = \sigma_F(\beta)\ \forall\, (\beta,F) \in (\Theta \times \mathcal{F}) \setminus \overline{\Lambda}_{\Theta,\mathcal{F}}$. Thus, if we replace the constant regularization parameter $\kappa > 0$ with a sequence $\kappa_n \to 0$ such that $\kappa_n \cdot n^{\epsilon} \to \infty$ and define now,

$$t_n(\beta) \equiv \frac{\sqrt{n} \cdot \widehat{\mathcal{T}}(\beta)}{(\sigma_F(\beta) \vee \kappa_n)}.$$

If we replace Assumption 6 with Assumption 6', the results in equation (36) are strengthened to the following,

$$\begin{aligned} &(i) \quad \lim_{n\to\infty} \sup_{(\beta,F)\in\overline{\Lambda}_{\Theta,\mathcal{F}}} P_F\left(t_n(\beta) > z_{1-\alpha}\right) = 0, \\ &(ii) \quad \lim_{n\to\infty} \sup_{(\beta,F)\in\Lambda_{\Theta,\mathcal{F}}\setminus\overline{\Lambda}_{\Theta,\mathcal{F}}} \left| P_F\left(t_n(\beta) > z_{1-\alpha}\right) - \alpha \right| = 0. \end{aligned} \tag{36'}$$

Thus, the test based on $t_n(\beta)$ would no longer be conservative if $(\beta,F)$ are such that $\sigma_F(\beta) < \kappa$ when $\kappa$ is a constant regularization parameter instead of a sequence vanishing to zero. All the remaining results regarding the construction of our confidence set remain valid.

## References


Andrews, D. (1994). Empirical process methods in econometrics. In R. Engle and D. McFadden (Eds.), *The Handbook of Econometrics*, Volume 4, Chapter 37, pp. 2247–2294. North-Holland.

Dudley, R. (1984). A course on empirical processes. *Ecole d'Été de Probabilitiés de Saint-Flour, XII-1982. Lecture Notes in Math 1097*, 1–142.

Kosorok, M. (2008). *Introduction to Empirical Processes and Semiparametric Inference*. Springer.

Nolan, D. and D. Pollard (1987). U-processes: Rates of convergence. *Annals of Statistics 15*, 780–799.

Pakes, A. and D. Pollard (1989). Simulation and the asymptotics of optimization estimators. *Econometrica 57*(5), 1027–1057.

Pollard, D. (1984). *Convergence of Stochastic Processes*. Springer Verlag.

Pollard, D. (1990). *Empirical Processes: Theory and Applications*. Institute of Mathematical Statistics.

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley. New York, NY.

Sherman, R. (1994). Maximal inequalities for degenerate u-processes with applications to optimization estimators. *Annals of Statistics 22*, 439–459.