Additional Proofs for:

"Large Sample Properties of Matching Estimators for Average Treatment Effects"

by Alberto Abadie & Guido Imbens

Reminder of Proof of Lemma 1: To get the result for $E[U_m U_m']$, notice that

$$\mathbb{E}[U_m U_m'] = N \begin{pmatrix} N-1 \\ m-1 \end{pmatrix} B_m,$$

where

$$B_m = \int_{\mathbb{R}^k} uu'\, f(z+u)\,(1-\Pr(\|X-z\|\le\|u\|))^{N-m}\,(\Pr(\|X-z\|\le\|u\|))^{m-1}\,du.$$

Boundedness of $\mathbb{X}$ implies that B_m converges uniformly. Transforming to polar coordinates again leads to

$$\begin{aligned}
B_m &= \int_0^\infty r^{k-1}\left(\int_{S_k} r^2\omega\omega'\, f(z+r\omega)\lambda_{S_k}(d\omega)\right)\left(1-\int_0^r s^{k-1}\left(\int_{S_k} f(z+s\omega)\lambda_{S_k}(d\omega)\right)ds\right)^{N-m} \\
&\qquad \times \left(\int_0^r s^{k-1}\left(\int_{S_k} f(z+s\omega)\lambda_{S_k}(d\omega)\right)ds\right)^{m-1} dr \\
&= \int_0^\infty e^{-Nb(r)}\,\tilde{a}(r)\,dr,
\end{aligned}$$

where, as before

$$b(r) = -\log\left(1-\int_0^r s^{k-1}\left(\int_{S_k} f(z+s\omega)\lambda_{S_k}(d\omega)\right)ds\right),$$

and

$$\tilde{a}(r) = r^{k+1}\cdot\left(\int_{S_k}\omega\omega'\, f(z+r\omega)\lambda_{S_k}(d\omega)\right)\frac{\left(\int_0^r s^{k-1}\left(\int_{S_k} f(z+s\omega)\lambda_{S_k}(d\omega)\right)ds\right)^{m-1}}{\left(1-\int_0^r s^{k-1}\left(\int_{S_k} f(z+s\omega)\lambda_{S_k}(d\omega)\right)ds\right)^{m}}.$$

That is, $\tilde{a}(r)=\tilde{q}(r)p(r)$, $\tilde{q}(r)=r^{k+1}\tilde{c}(r)$, and, as before, $p(r)=(g(r))^{m-1}$, where

$$\tilde{c}(r) = \frac{\int_{S_k}\omega\omega'\, f(z+r\omega)\lambda_{S_k}(d\omega)}{1-\int_0^r s^{k-1}\left(\int_{S_k} f(z+s\omega)\lambda_{S_k}(d\omega)\right)ds},$$

$$g(r) = \frac{\int_0^r s^{k-1}\left(\int_{S_k} f(z+s\omega)\lambda_{S_k}(d\omega)\right)ds}{1-\int_0^r s^{k-1}\left(\int_{S_k} f(z+s\omega)\lambda_{S_k}(d\omega)\right)ds}.$$

Clearly,

$$\lim_{r\to 0}\tilde{q}(r)r^{-(k+1)} = \lim_{r\to 0}\tilde{c}(r) = \frac{1}{k}f(z)\int_{S_k}\lambda_{S_k}(d\omega)I_k.$$

Hence,

$$\begin{aligned}\lim_{r\to 0}\tilde{a}(r)r^{-(mk+1)} &= \left(\lim_{r\to 0}p(r)r^{-(m-1)k}\right)\left(\lim_{r\to 0}\tilde{q}(r)r^{-(k+1)}\right)\\ &= \left(\frac{1}{k}f(z)\int_{S_k}\lambda_{S_k}(d\omega)\right)^{m-1}\frac{1}{k}f(z)\int_{S_k}\lambda_{S_k}(d\omega)I_k\\ &= \left(\frac{1}{k}f(z)\int_{S_k}\lambda_{S_k}(d\omega)\right)^{m}I_k.\end{aligned}$$

Therefore, the conditions of Lemma A.1 hold for $\alpha = mk+2$, $\beta = k$

$$\begin{aligned}a_0 &= \left(\frac{1}{k}f(z)\int_{S_k}\lambda_{S_k}(d\omega)\right)^{m}I_k\\ b_0 &= \frac{1}{k}f(z)\int_{S_k}\lambda_{S_k}(d\omega).\end{aligned}$$

Applying Lemma A.1, we get

$$\begin{aligned}B_m &= \Gamma\left(\frac{mk+2}{k}\right)\frac{a_0}{kb_0^{(mk+2)/k}}\frac{1}{N^{(mk+2)/k}}+o\left(\frac{1}{N^{(mk+2)/k}}\right)\\ &= \Gamma\left(\frac{mk+2}{k}\right)\frac{1}{k}\left(\frac{f(z)}{k}\int_{S_k}\lambda_{S_k}(d\omega)\right)^{-2/k}\frac{1}{N^{(mk+2)/k}}\cdot I_k+o\left(\frac{1}{N^{(mk+2)/k}}\right).\\ &= \Gamma\left(\frac{mk+2}{k}\right)\frac{1}{k}\left(f(z)\frac{\pi^{k/2}}{\Gamma(1+k/2)}\right)^{-2/k}\frac{1}{N^{(mk+2)/k}}\cdot I_k+o\left(\frac{1}{N^{(mk+2)/k}}\right).\end{aligned}$$

Hence, using the fact that

$$\lim_{N\to\infty}\frac{N^m/(m-1)!}{N\begin{pmatrix}N-1\\ m-1\end{pmatrix}}=1,$$

we have that

$$\begin{aligned}\mathbb{E}[U_mU_m'] &= N\begin{pmatrix}N-1\\ m-1\end{pmatrix}\cdot B_m\\ &= \Gamma\left(\frac{mk+2}{k}\right)\frac{1}{(m-1)!k}\left(f(z)\frac{\pi^{k/2}}{\Gamma\left(1+\frac{k}{2}\right)}\right)^{-2/k}\frac{1}{N^{2/k}}\cdot I+o\left(\frac{1}{N^{2/k}}\right).\end{aligned}$$

Using the same techniques as for the first two moments,

$$E\|U_m\|^3 = N\begin{pmatrix}N-1\\ m-1\end{pmatrix}C_m,$$

where

$$C_m = \int_0^\infty e^{-Nb(r)}\,\bar{a}(r)\,dr,$$

$$b(r) = -\log\left(1-\int_0^r s^{k-1}\left(\int_{S_k}f(z+s\omega)\lambda_{S_k}(d\omega)\right)ds\right),$$

and

$$\bar{a}(r) = r^{k+2} \cdot \left(\int_{S_k} f(z + r\omega)\lambda_{S_k}(d\omega) \right) \frac{\left(\int_0^r s^{k-1} \left(\int_{S_k} f(z + s\omega)\lambda_{S_k}(d\omega) \right) ds \right)^{m-1}}{\left(1 - \int_0^r s^{k-1} \left(\int_{S_k} f(z + s\omega)\lambda_{S_k}(d\omega) \right) ds \right)^m},$$

and C_m converges uniformly. Let $\bar{a}(r) = \bar{q}(r)p(r)$, $\bar{q}(r) = r^{k+2}\bar{c}(r)$, and $p(r) = (g(r))^{m-1}$, where

$$\bar{c}(r) = \frac{\int_{S_k} f(z + r\omega)\lambda_{S_k}(d\omega)}{1 - \int_0^r s^{k-1} \left(\int_{S_k} f(z + s\omega)\lambda_{S_k}(d\omega) \right) ds},$$

$$g(r) = \frac{\int_0^r s^{k-1} \left(\int_{S_k} f(z + s\omega)\lambda_{S_k}(d\omega) \right) ds}{1 - \int_0^r s^{k-1} \left(\int_{S_k} f(z + s\omega)\lambda_{S_k}(d\omega) \right) ds}.$$

Now,

$$\lim_{r\to 0} \tilde{q}(r)r^{-(k+2)} = \lim_{r\to 0} \tilde{c}(r) = f(z) \int_{S_k} \lambda_{S_k}(d\omega).$$

Hence,

$$\begin{aligned}
\lim_{r\to 0} \bar{a}(r)r^{-(mk+2)} &= \left(\lim_{r\to 0} p(r)r^{-(m-1)k} \right) \left(\lim_{r\to 0} \bar{q}(r)r^{-(k+2)} \right) \\
&= \left(\frac{1}{k} f(z) \int_{S_k} \lambda_{S_k}(d\omega) \right)^{m-1} f(z) \int_{S_k} \lambda_{S_k}(d\omega) \\
&= \left(\frac{1}{k} f(z) \int_{S_k} \lambda_{S_k}(d\omega) \right)^{m} k.
\end{aligned}$$

Therefore, the conditions of Lemma A.1 hold for $\alpha = mk + 3$, $\beta = k$

$$\begin{aligned}
a_0 &= \left(\frac{1}{k} f(z) \int_{S_k} \lambda_{S_k}(d\omega) \right)^{m} k \\
b_0 &= \frac{1}{k} f(z) \int_{S_k} \lambda_{S_k}(d\omega).
\end{aligned}$$

Applying Lemma A.1, we get

$$\begin{aligned}
C_m &= \Gamma\left(\frac{mk+3}{k} \right) \frac{a_0}{k b_0^{(mk+3)/k}} \frac{1}{N^{(mk+3)/k}} + o\left(\frac{1}{N^{(mk+3)/k}} \right) \\
&= \Gamma\left(\frac{mk+3}{k} \right) \left(\frac{f(z)}{k} \int_{S_k} \lambda_{S_k}(d\omega) \right)^{-3/k} \frac{1}{N^{(mk+3)/k}} + o\left(\frac{1}{N^{(mk+3)/k}} \right). \\
&= \Gamma\left(\frac{mk+3}{k} \right) \left(f(z) \frac{\pi^{k/2}}{\Gamma(1 + k/2)} \right)^{-3/k} \frac{1}{N^{(mk+3)/k}} + o\left(\frac{1}{N^{(mk+3)/k}} \right).
\end{aligned}$$

Hence, using the fact that

$$\lim_{N\to\infty} \frac{N^m/(m-1)!}{N \binom{N-1}{m-1}} = 1,$$

we have that

$$E[\|U_m\|^3] = N \begin{pmatrix} N-1 \\ m-1 \end{pmatrix} \cdot C_m$$

$$= \Gamma\left(\frac{mk+3}{k}\right) \frac{1}{(m-1)!} \left(f(z)\frac{\pi^{k/2}}{\Gamma(1+k/2)}\right)^{-3/k} \frac{1}{N^{3/k}} + o\left(\frac{1}{N^{3/k}}\right).$$

$$E[\|U_m\|^3] = \Gamma\left(\frac{mk+3}{k}\right) \frac{1}{(m-1)!} \left(f(z)\frac{\pi^{k/2}}{\Gamma(1+k/2)}\right)^{-3/k} \frac{1}{N^{3/k}} + o\left(\frac{1}{N^{3/k}}\right).$$

Therefore

$$E\|U_m\|^3 = O\left(\frac{1}{N^{3/k}}\right).$$

□

Proof of Theorem 5: First, assume without loss of generality that the support of X is $[0,1]$. (If the support of X is $[\underline{x},\overline{x}]$, we can always work with the transform $(X-\underline{x})/(\overline{x}-\underline{x})$ and obtain the same estimator. The results will not change because they are functions of ratios of densities, so the Jacobians of the transformation cancel.)

$$\Pr\big(j \in \mathcal{J}_M(1) | X_j = x_j, W_1 = 1, W_j = 0, \iota_N'\mathbf{W} = N_1\big)$$
$$= \sum_{m=1}^{M} \Pr\big(j = j_m(1) | X_j = x_j, W_1 = 1, W_j = 0, \iota_N'\mathbf{W} = N_1\big).$$

$$\Pr\big(j = j_m(1) | X_j = x_j, W_1 = 1, W_j = 0, \iota_N'\mathbf{W} = N_1\big)$$
$$= \int_0^1 \Pr\big(j = j_m(1) | X_1 = x_1, X_j = x_j, W_1 = 1, W_j = 0, \iota_N'\mathbf{W} = N_1\big) f_1(x_1)\, dx_1$$
$$= \int_0^1 \begin{pmatrix} N_0-1 \\ m-1 \end{pmatrix} \left(\int 1\big\{|x-x_1| \le |x_j - x_1|\big\} f_0(x) dx\right)^{m-1}$$
$$\times \left(1 - \int 1\big\{|x-x_1| \le |x_j - x_1|\big\} f_0(x) dx\right)^{N_0-m} f_1(x_1)\, dx_1$$
$$= \begin{pmatrix} N_0-1 \\ m-1 \end{pmatrix} \int_0^{x_j} \left(\int_{2x_1-x_j}^{x_j} f_0(x)\,dx\right)^{m-1} \left(1 - \int_{2x_1-x_j}^{x_j} f_0(x)\,dx\right)^{N_0-m} f_1(x)\,dx$$
$$+ \begin{pmatrix} N_0-1 \\ m-1 \end{pmatrix} \int_{x_j}^{1} \left(\int_{x_j}^{2x_1-x_j} f_0(x)\,dx\right)^{m-1} \left(1 - \int_{x_j}^{2x_1-x_j} f_0(x)\,dx\right)^{N_0-m} f_1(x)\,dx$$
$$= \begin{pmatrix} N_0-1 \\ m-1 \end{pmatrix} \int_0^{x_j} \left(\int_{x_j-2u}^{x_j} f_0(x)\,dx\right)^{m-1} \left(1 - \int_{x_j-2u}^{x_j} f_0(x)\,dx\right)^{N_0-m} f_1(x_j-u)\,du$$
$$+ \begin{pmatrix} N_0-1 \\ m-1 \end{pmatrix} \int_0^{1-x_j} \left(\int_{x_j}^{x_j+2v} f_0(x)\,dx\right)^{m-1} \left(1 - \int_{x_j}^{x_j+2v} f_0(x)\,dx\right)^{N_0-m} f_1(x_j+v)\,dv,$$

where $u = x_j - x_1$, $v = x_1 - x_j$. Notice that

$$\int_0^{x_j} \left(\int_{x_j-2u}^{x_j} f_0(x)\,dx\right)^{m-1} \left(1 - \int_{x_j-2u}^{x_j} f_0(x)\,dx\right)^{N_0-m} f_1(x_j-u)\,du$$
$$= \int a(u) \exp\big(-N_0 b(u)\big)\, du,$$

where

$$a(u) = f_1(x_j - u) \frac{\left(\int_{x_j - 2u}^{x_j} f_0(x)\, dx \right)^{m-1}}{\left(1 - \int_{x_j - 2u}^{x_j} f_0(x)\, dx \right)^{m}},$$

and

$$b(u) = -\log \left(1 - \int_{x_j - 2u}^{x_j} f_0(x)\, dx \right).$$

It is easy to see that $\lim_{u\to 0} a(u)/u^{m-1} = f_1(x_j)(2f_0(x_j))^{m-1}$, and $\lim_{u \to 0} b(u)/u = 2f_0(x_j)$. Applying Lemma A.1 we obtain:

$$\int_0^{x_j} \left(\int_{x_j - 2u}^{x_j} f_0(x)\, dx \right)^{m-1} \left(1 - \int_{x_j - 2u}^{x_j} f_0(x)\, dx \right)^{N_0 - m} f_1(x_j - u)\, du$$
$$= (m-1)! \frac{f_1(x_j)}{2f_0(x_j)} \frac{1}{N_0^m} + o\left(\frac{1}{N_0^m} \right).$$

Now, because,

$$\lim_{N_0 \to \infty} \frac{N_0^m/(m-1)!}{N_0 \begin{pmatrix} N_0 - 1 \\ m-1 \end{pmatrix}} = 1,$$

we obtain:

$$\begin{pmatrix} N_0 - 1 \\ m-1 \end{pmatrix} \int_0^{x_j} \left(\int_{x_j - 2u}^{x_j} f_0(x)\, dx \right)^{m-1} \left(1 - \int_{x_j - 2u}^{x_j} f_0(x)\, dx \right)^{N_0 - m} f_1(x_j - u)\, du$$
$$= \frac{f_1(x_j)}{2f_0(x_j)} \frac{1}{N_0} + o\left(\frac{1}{N_0} \right).$$

Similarly,

$$\begin{pmatrix} N_0 - 1 \\ m-1 \end{pmatrix} \int_0^{1-x_j} \left(\int_{x_j}^{x_j + 2v} f_0(x)\, dx \right)^{m-1} \left(1 - \int_{x_j}^{x_j + 2v} f_0(x)\, dx \right)^{N_0 - m} f_1(x_j + v)\, dv$$
$$= \frac{f_1(x_j)}{2f_0(x_j)} \frac{1}{N_0} + o\left(\frac{1}{N_0} \right).$$

Therefore:

$$\Pr\big(j = j_m(1) | X_j = x_j, W_1 = 1, W_j = 0, \iota_N' \mathbf{W} = N_1\big) = \frac{f_1(x_j)}{f_0(x_j)} \frac{1}{N_0} + o\left(\frac{1}{N_0} \right),$$

and

$$\Pr\big(j \in \mathcal{J}_M(1) | X_j = x_j, W_1 = 1, W_j = 0, \iota_N' \mathbf{W} = N_1\big) = M \frac{f_1(x_j)}{f_0(x_j)} \frac{1}{N_0} + o\left(\frac{1}{N_0} \right).$$

Now, let us calculate the joint probability of two matches:

$$
\begin{aligned}
&\Pr\big(j \in \mathcal{J}_M(1), j \in \mathcal{J}_M(2) | X_j = x_j, W_1 = W_2 = 1, W_j = 0, \iota_N' \mathbf{W} = N_1\big) \\
&= \Pr\big(j \in \mathcal{J}_M(1), j \in \mathcal{J}_M(2) | X_j = x_j, W_1 = W_2 = 1, W_j = 0, \iota_N' \mathbf{W} = N_1, X_2 \leq X_1\big) \\
&= \int_0^1 \int_0^{x_1} \Pr\big(j \in \mathcal{J}_M(1), j \in \mathcal{J}_M(2) | X_j = x_j, W_1 = W_2 = 1, W_j = 0, \iota_N' \mathbf{W} = N_1, \\
&\qquad\qquad X_1 = x_1, X_2 = x_2\big) \times \frac{f_1(x_1) f_1(x_2)}{1/2}\, dx_2\, dx_1.
\end{aligned}
$$

Let $P_j^{1,2} = \Pr\big(j \in \mathcal{J}_M(1), j \in \mathcal{J}_M(2) | X_j = x_j, W_1 = W_2 = 1, W_j = 0, \iota_N' \mathbf{W} = N_1, X_1 = x_1, X_2 = x_2\big)$. Now, let us calculate $P_j^{1,2}$ for $x_2 \leq x_1$. Three cases:

1. $[x_2 \leq x_1 \leq x_j]$: Notice that

$$
P_j^{1,2} = \sum_{m=1}^{M} \binom{N_0 - 1}{m - 1} \left(\int_{2x_2 - x_j}^{x_j} f_0(x)\, dx \right)^{m-1} \left(1 - \int_{2x_2 - x_j}^{x_j} f_0(x)\, dx \right)^{N_0 - m}.
$$

Let

$$
\begin{aligned}
P_1 &= \int_0^{x_j} \left[\int_{x_2}^{x_j} \sum_{m=1}^{M} \binom{N_0 - 1}{m - 1} \left(\int_{2x_2 - x_j}^{x_j} f_0(x)\, dx \right)^{m-1} \right. \\
&\quad \times \left. \left(1 - \int_{2x_2 - x_j}^{x_j} f_0(x)\, dx \right)^{N_0 - m} f_1(x_1)\, dx_1 \right] f_1(x_2)\, dx_2 \\
&= \sum_{m=1}^{M} \binom{N_0 - 1}{m - 1} \int_0^{x_j} \left(\int_{2x_2 - x_j}^{x_j} f_0(x)\, dx \right)^{m-1} \\
&\qquad \times \left(1 - \int_{2x_2 - x_j}^{x_j} f_0(x)\, dx \right)^{N_0 - m} \left(\int_{x_2}^{x_j} f_1(x_1)\, dx_1 \right) f_1(x_2)\, dx_2.
\end{aligned}
$$

Now let $u_2 = x_j - x_2$. Then,

$$
\begin{aligned}
&\int_0^{x_j} \left(\int_{2x_2 - x_j}^{x_j} f_0(x)\, dx \right)^{m-1} \left(1 - \int_{2x_2 - x_j}^{x_j} f_0(x)\, dx \right)^{N_0 - m} \left(\int_{x_2}^{x_j} f_1(x_1)\, dx_1 \right) f_1(x_2)\, dx_2 \\
&= \int_0^{x_j} \left(\int_{x_j - 2u_2}^{x_j} f_0(x)\, dx \right)^{m-1} \left(1 - \int_{x_j - 2u_2}^{x_j} f_0(x)\, dx \right)^{N_0 - m} \\
&\qquad \times \left(\int_{x_j - u_2}^{x_j} f_1(x_1)\, dx_1 \right) f_1(x_j - u_2)\, du_2 = \int_0^{x_j} a_1(u_2) \exp(-N_0\, b_1(u_2))\, du_2,
\end{aligned}
$$

where

$$
\begin{aligned}
a_1(u_2) &= f_1(x_j - u_2) \frac{\displaystyle\int_{x_j - u_2}^{x_j} f_1(x_1)\, dx_1 \left(\int_{x_j - 2u_2}^{x_j} f_0(x)\, dx \right)^{m-1}}{\displaystyle\left(1 - \int_{x_j - 2u_2}^{x_j} f_0(x)\, dx \right)^{m}}, \\
b_1(u_2) &= -\log\left(1 - \int_{x_j - 2u_2}^{x_j} f_0(x)\, dx \right).
\end{aligned}
$$

It is easy to see that

$$\lim_{u_2\to 0}\frac{a_1(u_2)}{u_2^m}=\big(f_1(x_j)\big)^2\big(2f_0(x_j)\big)^{m-1},$$

and

$$\lim_{u_2\to 0}\frac{b_1(u_2)}{u_2}=2f_0(x_j).$$

As a result, Lemma A.1 holds with $\alpha=m+1$, $\beta=1$, $a_0=(f_1(x_j))^2(2f_0(x_j))^{m-1}$, and $b_0=2f_0(x_j)$:

$$\begin{aligned}
&\int_0^{x_j}\left(\int_{2x_2-x_j}^{x_j}f_0(x)\,dx\right)^{m-1}\left(1-\int_{2x_2-x_j}^{x_j}f_0(x)\,dx\right)^{N_0-m}\left(\int_{x_2}^{x_j}f_1(x_1)\,dx_1\right)f_1(x_2)\,dx_2\\
&=m!\frac{\big(f_1(x_j)\big)^2\big(2f_0(x_j)\big)^{m-1}}{\big(2f_0(x_j)\big)^{m+1}}\frac{1}{N_0^{m+1}}+o\left(\frac{1}{N_0^{m+1}}\right)\\
&=m!\left(\frac{f_1(x_j)}{2f_0(x_j)}\right)^2\frac{1}{N_0^{m+1}}+o\left(\frac{1}{N_0^{m+1}}\right).
\end{aligned}$$

Now, because,

$$\lim_{N_0\to\infty}\frac{N_0^{m+1}/m!}{\frac{N_0^2}{m}\begin{pmatrix}N_0-1\\ m-1\end{pmatrix}}=1,$$

we obtain:

$$\begin{aligned}
P_1 &= \left(\sum_{m=1}^{M}m\right)\left(\frac{f_1(x_j)}{2f_0(x_j)}\right)^2\frac{1}{N_0^2}+o\left(\frac{1}{N_0^2}\right)\\
&= \frac{M(M+1)}{2}\left(\frac{f_1(x_j)}{2f_0(x_j)}\right)^2\frac{1}{N_0^2}+o\left(\frac{1}{N_0^2}\right).
\end{aligned}$$

2. $[x_2\le x_j\le x_1]$: Notice that

$$\begin{aligned}
P_j^{1,2}=\sum_{m_2=1}^{M}\sum_{m_1=1}^{M}&\frac{(N_0-1)!}{(m_2-1)!(m_1-1)!(N_0-m_1-m_2+1)!}\left(\int_{2x_2-x_j}^{x_j}f_0(x)\,dx\right)^{m_2-1}\\
&\times\left(\int_{x_j}^{2x_1-x_j}f_0(x)\,dx\right)^{m_1-1}\left(1-\int_{2x_2-x_j}^{2x_1-x_j}f_0(x)\,dx\right)^{N_0-m_1-m_2+1}.
\end{aligned}$$

Let

$$\begin{aligned}
P_2=&\int_0^{x_j}\int_{x_j}^{1}\sum_{m_2=1}^{M}\sum_{m_1=1}^{M}\frac{(N_0-1)!}{(m_2-1)!(m_1-1)!(N_0-m_1-m_2+1)!}\left(\int_{2x_2-x_j}^{x_j}f_0(x)\,dx\right)^{m_2-1}\\
&\times\left(\int_{x_j}^{2x_1-x_j}f_0(x)\,dx\right)^{m_1-1}\left(1-\int_{2x_2-x_j}^{2x_1-x_j}f_0(x)\,dx\right)^{N_0-m_1-m_2+1}f_1(x_1)f_1(x_2)\,dx_1dx_2.
\end{aligned}$$

Let $u_1 = x_1 - x_j$, $u_2 = x_j - x_2$. Notice that for $0 \le x_2 \le x_j \le x_1 \le 1$, $\sup\|(u_1, u_2)\| = 1$. Also, let $S_2^{++} = \{\omega = (\omega_1, \omega_2) \in \mathbb{R}^2 : \omega_1 \ge 0, \omega_2 \ge 0, \|\omega\| = 1\}$.

$$
\begin{aligned}
&\int_0^{x_j}\int_{x_j}^1 \left(\int_{2x_2-x_j}^{x_j} f_0(x)\,dx\right)^{m_2-1} \left(\int_{x_j}^{2x_1-x_j} f_0(x)\,dx\right)^{m_1-1} \\
&\times \left(1 - \int_{2x_2-x_j}^{2x_1-x_j} f_0(x)\,dx\right)^{N_0-m_1-m_2+1} f_1(x_1) f_1(x_2)\,dx_1 dx_2 \\
&= \int_0^{x_j}\int_0^{1-x_j} \left(\int_{x_j-2u_2}^{x_j} f_0(x)\,dx\right)^{m_2-1} \left(\int_{x_j}^{x_j+2u_1} f_0(x)\,dx\right)^{m_1-1} \\
&\times \left(1 - \int_{x_j-2u_2}^{x_j+2u_1} f_0(x)\,dx\right)^{N_0-m_1-m_2+1} f_1(x_j+u_1) f_1(x_j-u_2)\,du_1 du_2 \\
&= \int_{S_2^{++}} \left(\int_0^{\min\{(1-x_j)/\omega_1, x_j/\omega_2\}} r \left(\int_{x_j-2r\omega_2}^{x_j} f_0(x)\,dx\right)^{m_2-1} \left(\int_{x_j}^{x_j+2r\omega_1} f_0(x)\,dx\right)^{m_1-1}\right. \\
&\times \left.\left(1 - \int_{x_j-2r\omega_2}^{x_j+2r\omega_1} f_0(x)\,dx\right)^{N_0-m_1-m_2+1} f_1(x_j+r\omega_1) f_1(x_j-r\omega_2)\,dr\right) \lambda_{S_2}(d\omega) \\
&\qquad\qquad = \int_{S_2^{++}} \left(\int_0^{\min\{(1-x_j)/\omega_1, x_j/\omega_2\}} a_2(r)\exp(-N_0\, b_2(r))\,dr\right) \lambda_{S_2}(d\omega),
\end{aligned}
$$

where

$$
\begin{aligned}
a_2(r) &= \frac{r f_1(x_j+r\omega_1) f_1(x_j-r\omega_2) \left(\int_{x_j-2r\omega_2}^{x_j} f_0(x)\,dx\right)^{m_2-1} \left(\int_{x_j}^{x_j+2r\omega_1} f_0(x)\,dx\right)^{m_1-1}}{\left(1 - \int_{x_j-2r\omega_2}^{x_j+2r\omega_1} f_0(x)\,dx\right)^{m_1+m_2-1}}, \\
b_2(r) &= -\log\left(1 - \int_{x_j-2r\omega_2}^{x_j+2r\omega_1} f_0(x)\,dx\right).
\end{aligned}
$$

It can be easily seen that for $\omega_1 > 0$ and $\omega_2 > 0$

$$
\begin{aligned}
&\lim_{r\to 0} \frac{a_2(r)}{r^{m_1+m_2-1}} = \Big(f_1(x_j)\Big)^2 \Big(2\omega_2 f_0(x_j)\Big)^{m_2-1} \Big(2\omega_1 f_0(x_j)\Big)^{m_1-1}, \\
&\lim_{r\to 0} \frac{b_2(r)}{r} = 2 f_0(x_j)(\omega_1 + \omega_2).
\end{aligned}
$$

Therefore, Lemma A.1 holds with $\alpha = m_1 + m_2$, $\beta = 1$,

$$
a_0 = (f_1(x_j))^2 (2 f_0(x_j))^{m_1+m_2-2} \omega_1^{m_1-1} \omega_2^{m_2-1}, \quad \text{and} \quad b_0 = 2 f_0(x_j)(\omega_1 + \omega_2).
$$

$$
\begin{aligned}
&\int_0^{\min\{(1-x_j)/\omega_1, x_j/\omega_2\}} r \left(\int_{x_j-2r\omega_2}^{x_j} f_0(x)\,dx\right)^{m_2-1} \left(\int_{x_j}^{x_j+2r\omega_1} f_0(x)\,dx\right)^{m_1-1} \\
&\times \left(1 - \int_{x_j-2r\omega_2}^{x_j+2r\omega_1} f_0(x)\,dx\right)^{N_0-m_1-m_2+1} f_1(x_j+r\omega_1) f_1(x_j-r\omega_2)\,dr \\
&\qquad\qquad = (m_1+m_2-1)! \left(\frac{f_1(x_j)}{2 f_0(x_j)}\right)^2 \frac{\omega_1^{m_1-1}\omega_2^{m_2-1}}{(\omega_1+\omega_2)^{m_1+m_2}} \frac{1}{N_0^{m_1+m_2}} + o\left(\frac{1}{N_0^{m_1+m_2}}\right).
\end{aligned}
$$

Therefore,

$$\int_0^{\min\{(1-x_j)/\omega_1, x_j/\omega_2\}} \frac{(N_0-1)!\ N_0^2}{(N_0-m_1-m_2+1)!}\, r \left(\int_{x_j-2r\omega_2}^{x_j} f_0(x)\,dx\right)^{m_2-1} \left(\int_{x_j}^{x_j+2r\omega_1} f_0(x)\,dx\right)^{m_1-1}$$
$$\times \left(1-\int_{x_j-2r\omega_2}^{x_j+2r\omega_1} f_0(x)\,dx\right)^{N_0-m_1-m_2+1} f_1(x_j+r\omega_1)f_1(x_j-r\omega_2)\,dr$$
$$= (m_1+m_2-1)!\left(\frac{f_1(x_j)}{2f_0(x_j)}\right)^2 \frac{\omega_1^{m_1-1}\omega_2^{m_2-1}}{(\omega_1+\omega_2)^{m_1+m_2}} + o(1).$$

In addition, it can be shown that last integral is bounded by a constant. To see that, notice that

$$\left(\int_{x_j-2r\omega_2}^{x_j} f_0(x)\,dx\right)^{m_2-1} \le \left(2\bar{f}r\omega_2\right)^{m_2-1},$$

In addition, for $0 \le r \le (1-x_j)/2\omega_1$ we have

$$\left(\int_{x_j}^{x_j+2r\omega_1} f_0(x)\,dx\right)^{m_1-1} \le \left(2\bar{f}r\omega_1\right)^{m_1-1},$$

and

$$\left(1-\int_{x_j-2r\omega_2}^{x_j+2r\omega_1} f_0(x)\,dx\right)^{N_0-m_1-m_2+1} \le \left(1-\int_{x_j-r\omega_2}^{x_j+r\omega_1} f_0(x)\,dx\right)^{N_0-m_1-m_2+1}$$
$$\le \left(1-\underline{f}r(\omega_1+\omega_2)\right)^{N_0-m_1-m_2+1}.$$

Let $s = \underline{f}(\omega_1+\omega_2)r$.

$$\int_0^{\min\{(1-x_j)/\omega_1, x_j/\omega_2\}} \frac{(N_0-1)!\ N_0^2}{(N_0-m_1-m_2+1)!}\, r \left(\int_{x_j-2r\omega_2}^{x_j} f_0(x)\,dx\right)^{m_2-1} \left(\int_{x_j}^{x_j+2r\omega_1} f_0(x)\,dx\right)^{m_1-1}$$
$$\times \left(1-\int_{x_j-2r\omega_2}^{x_j+2r\omega_1} f_0(x)\,dx\right)^{N_0-m_1-m_2+1} f_1(x_j+r\omega_1)f_1(x_j-r\omega_2)\,dr$$
$$\le \bar{f}^2(2\bar{f})^{m_1+m_2-2}\omega_1^{m_1-1}\omega_2^{m_2-1}$$
$$\times \int_0^{\min\{(1-x_j)/\omega_1, x_j/\omega_2\}} \frac{(N_0-1)!\ N_0^2}{(N_0-m_1-m_2+1)!} r^{m_1+m_2-1}\left(1-\underline{f}(\omega_1+\omega_2)r\right)^{N_0-m_1-m_2+1} dr$$
$$= \frac{\bar{f}^2(2\bar{f})^{m_1+m_2-2}}{\underline{f}^{m_1+m_2}} \frac{\omega_1^{m_1-1}\omega_2^{m_2-1}}{(\omega_1+\omega_2)^{m_1+m_2}}$$
$$\times \frac{(N_0-1)!\ N_0^2}{(N_0-m_1-m_2+1)!}\int_0^{\underline{f}(\omega_1+\omega_2)\min\{(1-x_j)/\omega_1, x_j/\omega_2\}} s^{m_1+m_2-1}\left(1-s\right)^{N_0-m_1-m_2+1} ds.$$

It is easy to see that $\min\{(1-x_j)/\omega_1, x_j/\omega_2\} \le (\omega_1+\omega_2)^{-1}$. (If $x_j/\omega_2 \le (1-x_j)/\omega_1$, then $\omega_2 - \omega_2 x_j \ge \omega_1 x_j$ so $\omega_2 \ge (\omega_1+\omega_2)x_j$. An analogous argument applies for the case $x_j/\omega_2 \ge (1-x_j)/\omega_1$.) In addition, because the support of X has length one, it has to be the case that $\underline{f} \le 1$.

Therefore, the upper limit of the integral is smaller than one. We obtain

$$\int_0^{\min\{(1-x_j)/\omega_1, x_j/\omega_2\}} \frac{(N_0-1)!\, N_0^2}{(N_0-m_1-m_2+1)!}\, r \left(\int_{x_j-2r\omega_2}^{x_j} f_0(x)\,dx\right)^{m_2-1} \left(\int_{x_j}^{x_j+2r\omega_1} f_0(x)\,dx\right)^{m_1-1}$$
$$\times \left(1-\int_{x_j-2r\omega_2}^{x_j+2r\omega_1} f_0(x)\,dx\right)^{N_0-m_1-m_2+1} f_1(x_j+r\omega_1) f_1(x_j-r\omega_2)\,dr$$
$$\leq \frac{\bar{f}^2(2\bar{f})^{m_1+m_2-2}}{\underline{f}^{m_1+m_2}} \frac{\omega_1^{m_1-1}\omega_2^{m_2-1}}{(\omega_1+\omega_2)^{m_1+m_2}}$$
$$\times \frac{(N_0-1)!\, N_0^2}{(N_0-m_1-m_2+1)!} \int_0^1 s^{m_1+m_2-1}\,(1-s)^{N_0-m_1-m_2+1}\,ds.$$

The integral in the right hand side is a Beta function with parameters m_1+m_2 and $N_0-m_1-m_2+2$, which is equal to $[(m_1+m_2-1)!(N_0-m_1-m_2+1)!]/[(N_0+1)!]$. Therefore,

$$\int_0^{\min\{(1-x_j)/\omega_1, x_j/\omega_2\}} \frac{(N_0-1)!\, N_0^2}{(N_0-m_1-m_2+1)!}\, r \left(\int_{x_j-2r\omega_2}^{x_j} f_0(x)\,dx\right)^{m_2-1} \left(\int_{x_j}^{x_j+2r\omega_1} f_0(x)\,dx\right)^{m_1-1}$$
$$\times \left(1-\int_{x_j-2r\omega_2}^{x_j+2r\omega_1} f_0(x)\,dx\right)^{N_0-m_1-m_2+1} f_1(x_j+r\omega_1) f_1(x_j-r\omega_2)\,dr$$
$$\leq \frac{\bar{f}^2(2\bar{f})^{m_1+m_2-2}}{\underline{f}^{m_1+m_2}} \frac{\omega_1^{m_1-1}\omega_2^{m_2-1}}{(\omega_1+\omega_2)^{m_1+m_2}} (m_1+m_2-1)! \frac{N_0^2}{(N_0+1)N_0}$$
$$\leq \frac{\bar{f}^2(2\bar{f})^{m_1+m_2-2}}{\underline{f}^{m_1+m_2}} \frac{\omega_1^{m_1-1}\omega_2^{m_2-1}}{(\omega_1+\omega_2)^{m_1+m_2}} (m_1+m_2-1)!,$$

which is integrable over $(\omega_1,\omega_2)\in S_2^{++}$, because

$$\int_{S_2^{++}} \frac{\omega_1^{m_1-1}\omega_2^{m_2-1}}{(\omega_1+\omega_2)^{m_1+m_2}} \lambda_{S_2}(d\omega) = \frac{(m_1-1)!\,(m_2-1)!}{(m_1+m_2-1)},$$

(see, e.g., Gradshteyn and Ryzhik, 2000, eq. 3.667 7).

Then, applying Lebesgue's Dominated Convergence Theorem, and using again the result in last equation we obtain:

$$P_2 = M^2\left(\frac{f_1(x_j)}{2f_0(x_j)}\right)^2 \frac{1}{N_0^2} + o\left(\frac{1}{N_0^2}\right).$$

3. $[x_j \leq x_2 \leq x_1]$: This case is analogous to the case with $[x_2 \leq x_1 \leq x_j]$ and contributes the same amount to the integral:

$$P_3 = \frac{M(M+1)}{2}\left(\frac{f_1(x_j)}{2f_0(x_j)}\right)^2 \frac{1}{N_0^2} + o\left(\frac{1}{N_0^2}\right).$$

Hence,

$$\Pr\big(j\in\mathcal{J}_M(1), j\in\mathcal{J}_M(2) | X_j=x_j, W_1=W_2=1, W_j=0, \iota_N'\mathbf{W}=N_1\big)$$
$$= 2\big(P_1+P_2+P_3\big) = \frac{M(2M+1)}{2}\left(\frac{f_1(x_j)}{f_0(x_j)}\right)^2 \frac{1}{N_0^2} + o\left(\frac{1}{N_0^2}\right).$$

As a result the first two moments of $K(i)$ are given by:

$$E[K(i)|W_i=0,X_i=x_i,\iota_N'\mathbf{W}=N_1]=M\frac{f_1(x_i)}{f_0(x_i)}\frac{N_1}{N_0}+\frac{N_1}{N_0}\,o(1).$$

$$\begin{aligned}E[K(i)^2|W_i=0,X_i=x_i,\iota_N'\mathbf{W}=N_1] &= M\frac{f_1(x_i)}{f_0(x_i)}\frac{N_1}{N_0}\\ &+ \frac{M(2M+1)}{2}\left(\frac{f_1(x_i)}{f_0(x_i)}\right)^2\frac{N_1(N_1-1)}{N_0^2}\\ &+ \frac{N_1}{N_0}\,o(1).\end{aligned}$$

Let $p=\Pr(W_i=1)$.

$$\begin{aligned}\mathbb{E}\left[\left(1+\frac{K_M(i)}{M}\right)^2\sigma^2_{W_i}(X_i)\right] &= \mathbb{E}\left[\left(1+\frac{K_M(i)}{M}\right)^2\sigma^2_{W_i}(X_i)\middle| W_i=0\right]\Pr(W_i=0)\\ &+ \mathbb{E}\left[\left(1+\frac{K_M(i)}{M}\right)^2\sigma^2_{W_i}(X_i)\middle| W_i=1\right]\Pr(W_i=1).\end{aligned}$$

We focus on the first term. The second term can be calculated analogously. First, notice that $N_1/N_0\to p/(1-p)$ almost surely. From the proof of Lemma 3, we obtain:

$$\mathbb{E}[K_M^q(i)|W_i=0,X_i=x,\iota_N'\mathbf{W}]\le C\sum_{n=1}^{q}S(q,n)\cdot\left(\frac{N_1}{N_0}\right)^n,$$

for some $C>0$ and all $q\ge 1$. Using Chernoff's Inequality, it is easy to show that all positive moments of (N_1/N_0) exist, conditional of $X_i=x$ and $W_i=0$ and are bounded uniformly in N. Markov's Inequality implies that $(N_1/N_0)^n$ is asymptotically uniformly integrable, which, in turn, implies convergence of moments:

$$E[K_M(i)|W_i=0,X_i=x]=M\frac{f_1(x)}{f_0(x)}\frac{p}{1-p}+o(1).$$

$$\begin{aligned}E[K_M(i)^2|W_i=0,X_i=x] &= M\frac{f_1(x)}{f_0(x)}\frac{p}{1-p}\\ &+ \frac{M(2M+1)}{2}\left(\frac{f_1(x)}{f_0(x)}\right)^2\frac{p^2}{(1-p)^2}+o(1).\end{aligned}$$

Therefore,

$$\begin{aligned}&\mathbb{E}\left[\left(1+\frac{K_M(i)}{M}\right)^2\middle| W_i=0,X_i=x\right]=\mathbb{E}\left[\left(1+2\frac{K_M(i)}{M}+\frac{K_M^2(i)}{M^2}\right)\middle| W_i=0,X_i=x\right]\\ &=1+2\frac{f_1(x)}{f_0(x)}\frac{p}{1-p}+\frac{1}{M}\frac{f_1(x)}{f_0(x)}\frac{p}{1-p}+\frac{2M+1}{2M}\left(\frac{f_1(x)}{f_0(x)}\right)^2\frac{p^2}{(1-p)^2}+o(1)\\ &\qquad=\left(1+\frac{p}{1-p}\frac{f_1(x)}{f_0(x)}\right)^2+\frac{1}{M}\left(\frac{p}{1-p}\frac{f_1(x)}{f_0(x)}+\frac{1}{2}\frac{p^2}{(1-p)^2}\frac{f_1(x)^2}{f_0(x)^2}\right)+o(1).\end{aligned}$$

Because $e(x)=pf_1(x)/(pf_1(x)+(1-p)f_0(x))$, and thus $(1-e(x))/e(x)=(1-p)f_0(x)/(pf_1(x))$, this can be written as

$$\mathbb{E}\left[\left(1+\frac{K_M(i)}{M}\right)^2\middle| W_i=0,X_i=x\right]=\frac{1}{(1-e(x))^2}+\frac{1}{M}\left(\frac{e(x)}{1-e(x)}+\frac{1}{2}\frac{e(x)^2}{(1-e(x))^2}\right)+o(1).$$

Because the bound on the conditional moments of $K_M(i)$, and because the moments of $(N_1/N_0)^n$ conditional on $X_i = x$ and $W_i = 0$ do not depend on x, and are bounded uniformly in N, we obtain:

$$\mathbb{E}\left[\left(1+\frac{K_M(i)}{M}\right)^2 \sigma^2_{W_i}(X_i)\middle| W_i=0\right]$$
$$= \mathbb{E}\left[\left(\frac{1}{(1-e(X_i))^2}+\frac{1}{M}\left(\frac{e(X_i)}{1-e(X_i)}+\frac{1}{2}\frac{e(X_i)^2}{(1-e(X_i))^2}\right)\right)\sigma^2_0(X_i)\middle| W_i=0\right]+o(1).$$

After some algebra it can be shown that:

$$\mathbb{E}\left[\left(\frac{1}{(1-e(X_i))^2}+\frac{1}{M}\left(\frac{e(X_i)}{1-e(X_i)}+\frac{1}{2}\frac{e(X_i)^2}{(1-e(X_i))^2}\right)\right)\sigma^2_0(X_i)\middle| W_i=0\right](1-p)$$
$$= \mathbb{E}\left[\frac{\sigma^2_0(X_i)}{1-e(X_i)}\right]+\frac{1}{2M}\mathbb{E}\left[\left(\frac{1}{1-e(X_i)}-(1-e(X_i))\right)\sigma^2_0(X_i)\right].$$

The analogous result holds conditioning on $W_i = 1$, so the result follows. □

Although it is not necessary for the proof of Theorem 5, it can be seen that V^E converges to its expectation. Before proving that result we give a couple of preliminary lemmas. As before, without loss of generality assume that the support of X is $[0, 1]$. Define

$$P_i = \begin{cases} \int_{A_M(i)} f_0(x)dx & \text{if } W_i = 1, \\ \int_{A_M(i)} f_1(x)dx & \text{if } W_i = 0, \end{cases}$$

Let $X_{0,i}$, for $i = 1, \ldots, N_0$ be the covariates for the N_0 control observations, with $P_{0,i}$ and $K_{0,M}(i)$ the corresponding versions of P_i and $K_M(i)$. Also define $X_{0,(i)}$ to be the order statistics so that $X_{0,(i)} < X_{0,(i+1)}$, and let $A_M(i)$ be the corresponding catchment areas:

$$A_M((i)) = \big((X_{0,(i)}+X_{0,(i-M)})/2, (X_{0,(i)}+X_{0,(i+M)})/2\big) \qquad \text{for} \qquad M \le i \le N_0 - M,$$

and $P_{0,(i)}$ the corresponding probabilities (that is, based on the ordering of the covariates, not based on the ordering of the probabilities themselves):

$$P_{0,(i)} = \int_{A_M((i))} f_1(x)dx = \int_{(X_{0,(i)}+X_{0,(i-M)})/2}^{(X_{0,(i)}+X_{0,(i+M)})/2} f_1(x)dx.$$

Define

$$\tilde{P}_{0,(i)} = \frac{f_1(X_{0,(i)})}{f_0(X_{0,(i)})}\cdot (F_0(X_{0,(i+M)} - F_0(X_{0,(i-M)}))/2.$$

Lemma A.6: *For all $\delta > 0$, and all fixed K*
(i)

$$\max_i (X_{0,(i+K)} - X_{0,(i)}) = o_p(N_0^{-1+\delta}),$$

(ii)

$$\max_i |\tilde{P}_{0,i} - P_{0,i}| = o_p(N_0^{-2+\delta}).$$

and (iii)

$$\max_i P_{0,i} = o_p(N_0^{-1+\delta}),$$

Proof of Lemma A.6:
First, consider part (i). Because $F_0(X_{0,(i+M)}) - F_0(X_{0,(i)}) = f_0(x)\cdot(X_{0,(i+K)} - X_{0,(i)})$ for some $x \in [X_{0,(i)}, X_{0,(i+K)}]$, it follows that

$$\max_i |X_{0,(i+K)} - X_{0,(i)}| \le \sup_x \frac{1}{f_0(x)} \cdot \max_i \left|F_0(X_{0,(i+K)}) - F_0(X_{0,(i)})\right|$$

With the support of X compact and the density bounded away from zero the first factor is bounded. By the fact that the distribution function of a continuous random variable has a uniform distribution, and by the fact that the order statistics of uniform random variables can be written as ratios of sums of iid unit exponentials the second factor has the same distribution as

$$\max_{i=1,\ldots,N_0-K} \sum_{m=i}^{i+K} \epsilon_m \Big/ \sum_{m=0}^{N_0} \epsilon_m,$$

for i.i.d. unit exponential ϵ_m. Hence it will be sufficient to show that $\max_{i=1,\ldots,N_0-K} \sum_{m=1}^{K} \epsilon_{i+m} / \sum_{j=0}^{N_0} \epsilon_j = o_p(N_0^{-1+\delta})$. To show this we first show that

$$\max_i \left| \sum_{m=1}^{K} \epsilon_{i+m} \Big/ \sum_{j=0}^{N_0} \epsilon_j - \sum_{m=1}^{K} \epsilon_{i+m} \Big/ (N_0+1) \right| = o_p(N_0^{-1+\delta}), \tag{A.16}$$

and second that

$$\max_i \sum_{m=1}^{K} \epsilon_{i+m}/(N_0+1) = o_p(N_0^{-1+\delta}). \tag{A.17}$$

To show that (A.16) holds, consider

$$N_0^{1-\delta} \cdot \max_i \left| \sum_{m=1}^{K} \epsilon_{i+m} \Big/ \sum_{j=0}^{N_0} \epsilon_j - \sum_{m=1}^{K} \epsilon_{i+m} \Big/ (N_0+1) \right| = N_0^{1-\delta} \cdot \max_i \sum_{m=1}^{K} \epsilon_{i+m} \cdot \left| \frac{1}{N_0+1} - \frac{1}{\sum_j \epsilon_j} \right|$$

$$\le \left| \frac{1}{1+1/N_0} - \frac{1}{\sum_j \epsilon_j / N_0} \right| \cdot N_0^{-\delta} \cdot M \cdot \max_i \epsilon_i. \tag{A.18}$$

The second factor in (A.18), $N_0^{-\delta} \cdot M \cdot \max_i \epsilon_i$ is $o_p(1)$ because all moments of ϵ_m exists because it is unit exponential. The first factor in (A.18) converges to zero, so (A.18) is $o_p(1)$ and (A.16) holds.
To show that (A.17) holds, consider

$$\Pr\left(N_0^{1-\delta} \cdot \max_i \sum_{m=1}^{K} \epsilon_{i+m} \Big/ (N_0+1) > C \right)$$

$$\le \Pr\left(N_0^{-\delta} \cdot \max_i \sum_{m=1}^{K} \epsilon_{i+m} > C \right)$$

$$\le \sum_{i=1}^{N_0} \Pr\left(N_0^{-\delta} \cdot \sum_{m=1}^{K} \epsilon_{i+m} > C \right)$$

$$\le N_0 \cdot \Pr\left(\sum_{m=1}^{M} \epsilon_m \ge C N^{\delta} \right).$$

Pick $k > 1/\delta$. Then the right hand side is equal to

$$N_0 \cdot \Pr\left(\left|\sum_{m=1}^{M} \epsilon_m\right|^k \geq C^k N^{k\cdot\delta}\right).$$

By Chebyshev's inequality this can be bounded from above by

$$N_0 \cdot \frac{\mathbb{E}\left[\left|\sum_{m=1}^{M} \epsilon_m\right|^k\right]}{C^k N_0^{k\cdot\delta}} = \mathbb{E}\left[\left|\sum_{m=1}^{M} \epsilon_m\right|^k\right] C^{-k} N_0^{1-k\cdot\delta},$$

which converges to zero because $\mathbb{E}\left[\left|\sum_{m=1}^{M} \epsilon_m\right|^k\right]$ is finite given that the ϵ_i are iid unit exponential. This finishes part (i) of the Lemma.

Next, consider part (ii). Because

$$\max_i |\tilde{P}_{0,i} - P_{0,i}| = \max_i |\tilde{P}_{0,(i)} - P_{0,(i)}|,$$

it is sufficient to show that

$$\max_i |\tilde{P}_{0,(i)} - P_{0,(i)}| = o_p(N_0^{-2+\delta}).$$

By the triangle inequality

$$N^{2-\delta} \cdot \max_i |\tilde{P}_{0,(i)} - P_{0,(i)}|$$

$$\leq N^{2-\delta} \cdot \max_i \left| \int_{(X_{0,(i)}+X_{0,(i-M)})/2}^{(X_{0,(i)}+X_{0,(i+M)})/2} f_1(x)dx - f_1(X_{0,(i)}) \cdot (X_{0,(i+M)} - X_{0,(i-M)})/2 \right| \tag{A.19}$$

$$+N^{2-\delta} \cdot \max_i \left| f_1(X_{0,(i)}) \cdot (X_{0,(i+M)} - X_{0,(i-M)})/2 - \frac{f_1(X_{0,(i)})}{f_0(X_{0,(i)})} \cdot (F_0(X_{0,(i+M)} - F_0(X_{0,(i-M)}))/2 \right|. \tag{A.20}$$

Consider the first term, (A.19):

$$N^{2-\delta} \cdot \max_i \left| \int_{(X_{0,(i)}+X_{0,(i-M)})/2}^{(X_{0,(i)}+X_{0,(i+M)})/2} f_1(x)dx - f_1(X_{0,(i)}) \cdot (X_{0,(i+M)} - X_{0,(i-M)})/2 \right|$$

$$\leq N^{2-\delta} \cdot \max_i \left|(X_{0,(i+M)} - X_{0,(i-M)})/2\right| \cdot \max_i \sup_{x\in[X_{0,(i-M)},X_{0,(i+M)}]} \left|f_1(x) - f_1(X_{0,(i)})\right|$$

$$\leq \frac{1}{2} \cdot N^{2-\delta} \cdot \max_i \left|X_{0,(i+M)} - X_{0,(i-M)}\right| \cdot \max_i \left|X_{0,(i+M)} - X_{0,(i-M)}\right| \cdot \sup_x |f_1'(x)|$$

$$= \frac{1}{2} \cdot N^{2-\delta} \cdot \max_i \left|X_{0,(i+M)} - X_{0,(i-M)}\right|^2 \cdot \sup_x |f_1'(x)|.$$

The last factor is bounded. Because for all $\delta' > 0$ we have $N^{1-\delta'} \cdot \max_i \left|(X_{0,(i+M)} - X_{0,(i-M)})\right| = o_p(1)$, it follows that for $\delta > \delta'/2$

$$N^{2-\delta} \cdot \max_i \left|(X_{0,(i+M)} - X_{0,(i-M)})\right|^2 = N^{2\delta'-\delta} \cdot \left(N^{1-\delta'} \cdot \max_i \left|(X_{0,(i+M)} - X_{0,(i-M)})\right|\right)^2 = o_p(1).$$

Hence (A.19) is $o_p(1)$.
To show that the second term, (A.20), is $o_p(1)$ it is sufficient to show that

$$N^{2-\delta}\cdot\max_i\left|(X_{0,(i+M)}-X_{0,(i-M)})-\frac{1}{f_0(X_{0,(i)})}\cdot(F_0(X_{0,(i+M)}-F_0(X_{0,(i-M)}))\right|=o_p(1),$$

because $f_1(x)$ is bounded. By a mean value theorem it follows that $(X_{0,(i+M)}-X_{0,(i-M)})\cdot f_0(x)=(F_0(X_{0,(i+M)}-F_0(X_{0,(i-M)}))$ for some $x\in[X_{0,(i-M)},X_{0,(i+M)}]$. So, for this value of x,

$$\begin{aligned}
&N^{2-\delta}\cdot\max_i\left|(X_{0,(i+M)}-X_{0,(i-M)})-\frac{1}{f_0(X_{0,(i)})}\cdot(F_0(X_{0,(i+M)}-F_0(X_{0,(i-M)}))\right|\\
&\qquad=N^{2-\delta}\cdot\max_i\left|(X_{0,(i+M)}-X_{0,(i-M)})\cdot\left(1-\frac{f_0(x)}{f_0(X_{0,(i)}}\right)\right|\\
&\qquad=N^{2-\delta}\cdot\max_i\left|(X_{0,(i+M)}-X_{0,(i-M)})\cdot\left(1-\frac{f_0(X_{0,(i)})+f_0'(\tilde{x})\cdot(x-X_{0,(i)})}{f_0(X_{0,(i)})}\right)\right|\\
&\qquad\le N^{1-\delta/2}\cdot\max_i\left|(X_{0,(i+M)}-X_{0,(i-M)})\right|\cdot N_0^{1-\delta/2}\cdot\left|(x-X_{0,(i)})\cdot\frac{f_0'(\tilde{x})}{f_0(X_{0,(i)})}\right|\\
&\qquad\le N^{1-\delta/2}\cdot\max_i\left|(X_{0,(i+M)}-X_{0,(i-M)})\right|\cdot N_0^{1-\delta/2}\cdot\left|(x-X_{0,(i)})\right|\cdot\sup_{x',x''}\frac{f_0'(x')}{f_0(x'')}.
\end{aligned}$$

Because $x\in[X_{0,(i-M)},X_{0,(i+M)}]$ and $(X_{0,(i+M)}-X_{0,(i-M)})=o_p(N_0^{1-\delta/2})$, it follows that $N^{1-\delta/2}\cdot\max_i\left|(X_{0,(i+M)}-X_{0,(i-M)})\right|$ and $N_0^{1-\delta/2}\cdot\left|(x-X_{0,(i)})\right|$ are $o_p(1)$. Because the last factor is bounded the product is $o_p(1)$ and thus (A.20) is $o_p(1)$. This finished part *(ii)* of the Lemma.
Finally, consider part *(iii)* of the Lemma. By part *(ii)* of the Lemma it is sufficient to show that $\tilde{P}_{0,(i)}=o_p(N_0^{1-\delta})$. Because $f_0(x)$ and $f_1(x)$ are bounded and bounded away from zero it is sufficient to show that $\max_i(F_0(X_{0,(i+M)})-F_0(X_{0,(i-M)})=o_p(N_0^{1-\delta})$. This follows by the same argument as in part *(i)* of the Lemma. □

LEMMA A.7: *Suppose that $g(\cdot)$ is continuously differentiable on the interval $[0,1]$. If $\epsilon_0,\ldots,\epsilon_N$ are iid unit-exponential, then for fixed M, as $N\to\infty$, (i)*

$$\frac{1}{2}\cdot\sum_{i=M}^{N-M}g\left(\sum_{j=0}^{i}\epsilon_j\Big/\sum_{l=0}^{N}\epsilon_l\right)\cdot\left(\sum_{j=-M+1}^{M}\epsilon_{i+j}\Big/\sum_{l=0}^{N}\epsilon_l\right)\xrightarrow{p}M\cdot\int_0^1g(x)dx.\tag{A.21}$$

and (ii)

$$\frac{1}{4}\cdot\sum_{i=M}^{N-M}g\left(\sum_{j=0}^{i}\epsilon_j\Big/\sum_{l=0}^{N}\epsilon_l\right)\cdot\left(\sum_{j=-M+1}^{M}\epsilon_{i+j}\Big/\sum_{l=0}^{N}\epsilon_l\right)^2\xrightarrow{p}\frac{M(2M+1)}{2}\cdot\int_0^1g(x)dx.\tag{A.22}$$

PROOF: We show that for arbitrary fixed integers K_1, K_2, and $K_3\le\min(K_1,K_2)$ we have

$$\sum_{i=K_1}^{N-K_2}g\left(\sum_{j=0}^{i}\epsilon_j\Big/\sum_{l=0}^{N}\epsilon_l\right)\cdot\left(\epsilon_{i+K_3}\Big/\sum_{l=0}^{N}\epsilon_l\right)\xrightarrow{p}\int_0^1g(x)dx.\tag{A.23}$$

Applying to each of the $2\cdot M$ terms in (A.21) then gives the desired result. We prove that (A.23) holds for $K_1=K_2=K_3=0$. The difference with terms with other values of K_1, K_2, and K_3 is of order $o_p(1)$. By the triangle inequality we have

$$\left|\sum_{i=1}^{N}g\left(\sum_{j=0}^{i}\epsilon_j\Big/\sum_{l=0}^{N}\epsilon_l\right)\cdot\left(\epsilon_i\Big/\sum_{l=0}^{N}\epsilon_l\right)-\int_0^1g(x)dx\right|$$

$$\leq \left|\sum_{i=1}^{N} g\left(\sum_{j=0}^{i}\epsilon_j \Big/ \sum_{l=0}^{N}\epsilon_l\right)\cdot\left(\epsilon_i \Big/ \sum_{l=0}^{N}\epsilon_l\right) - \sum_{i=1}^{N} g\left(i/(N+1)\right)\cdot\left(\epsilon_i \Big/ \sum_{l=0}^{N}\epsilon_l\right)\right| \tag{A.24}$$

$$+\left|\sum_{i=1}^{N} g\left(i/(N+1)\right)\cdot\left(\epsilon_i \Big/ \sum_{l=0}^{N}\epsilon_l\right) - \sum_{i=1}^{N} g\left(i/(N+1)\right)\cdot\frac{\epsilon_i}{N+1}\right| \tag{A.25}$$

$$+\left|\frac{1}{N+1}\sum_{i=1}^{N} g(i/(N+1))\cdot\epsilon_i - \frac{1}{N+1}\sum_{i=1}^{N} g(i/(N+1))\right| \tag{A.26}$$

$$+\left|\frac{1}{N+1}\sum_{i=1}^{N} g(i/(N+1)) - \int_0^1 g(x)dx\right|. \tag{A.27}$$

We will show for each of the terms (A.24)-(A.27) that they are of order $o_p(1)$.
First note that

$$\max_{i=1,\ldots,N}\left|\frac{1}{N+1}\sum_{j=0}^{i}\epsilon_i - \frac{i}{N+1}\right| = o_p(1). \tag{A.28}$$

This follows because

$$\max_{i=1,\ldots,N}\left|\frac{1}{N+1}\sum_{j=0}^{i}\epsilon_i - \frac{i}{N+1}\right| = \max_{i=1,\ldots,N}\left|\frac{1}{N+1}\sum_{j=0}^{i}(\epsilon_i - 1)\right|.$$

The expectation of $\epsilon_i - 1$ is zero, with second moment equal to μ_2 and fourth moment finite. Then define $Z_i = \frac{1}{N+1}\sum_{j=0}^{i}(\epsilon_i - 1)$. The fourth moment of Z_i is bounded by $C_1 \cdot i^2/N^4$ for some finite C_1. Then, for $0 < \alpha < 1/4$,

$$\Pr(\max_i Z_i \geq C_2 \cdot N^{-\alpha}) \leq \sum_{i=1}^{N}\Pr(Z_i \geq C_2 \cdot N^{-\alpha})$$

$$= \sum_{i=1}^{N}\Pr(Z_i^4 \geq C_2^2 \cdot N^{-4\cdot\alpha}) \leq \sum_{i=1}^{N}\mathbb{E}[Z_i^4]/(C_2^4 \cdot N^{-4\alpha})$$

$$= \sum_{i=1}^{N} C_1 \cdot (i^2/N^4)\cdot N^{4\alpha}/C_2^4 \leq C_1 \cdot N^{3+4\alpha-4}/C_2^4 \to 0.$$

Next, note that

$$\max_{i=1,\ldots,N}\left|\sum_{j=0}^{i}\epsilon_i \Big/ \sum_{l=0}^{N}\epsilon_l - \sum_{j=0}^{i}\epsilon_j/(N+1)\right| = o_p(1). \tag{A.29}$$

This follows from

$$\max_{i=1,\ldots,N}\left|\sum_{j=0}^{i}\epsilon_i \Big/ \sum_{l=0}^{N}\epsilon_l - \sum_{j=0}^{i}\epsilon_j/(N+1)\right| \leq \max_i \frac{1}{N+1}\sum_{j=0}^{i}\epsilon_i \cdot \left|\frac{1}{\sum_{l=0}^{N}\epsilon_l/(N+1)} - 1\right|$$

$$= \frac{1}{N+1}\sum_{j=0}^{N}\epsilon_i \cdot \left|\frac{1}{\sum_{l=0}^{N}\epsilon_l/(N+1)} - 1\right| = o_p(1).$$

Combined these two results imply that

$$\max_{i=1,\ldots,N}\left|\sum_{j=0}^{i}\epsilon_i\Big/\sum_{l=0}^{N}\epsilon_l - i/(N+1)\right| = o_p(1). \tag{A.30}$$

Now consider (A.24).

$$\left|\sum_{i=1}^{N} g\left(\sum_{j=0}^{i}\epsilon_j\Big/\sum_{l=0}^{N}\epsilon_l\right)\cdot\left(\epsilon_i\Big/\sum_{l=0}^{N}\epsilon_l\right) - \sum_{i=1}^{N} g\left(i/(N+1)\right)\cdot\left(\epsilon_i\Big/\sum_{l=0}^{N}\epsilon_l\right)\right|$$

$$\leq \max_x g'(x)\cdot\max_i\left|\sum_{j=0}^{i}\epsilon_j\Big/\sum_{l=0}^{N}\epsilon_l - i/(N+1)\right|\cdot\left|\sum_{i=1}^{N}\left(\epsilon_i\Big/\sum_{l=0}^{N}\epsilon_l\right)\right|$$

$$= \max_x g'(x)\cdot\max_i\left|\sum_{j=0}^{i}\epsilon_j\Big/\sum_{l=0}^{N}\epsilon_l - i/(N+1)\right| = o_p(1).$$

Next, consider (A.25).

$$\left|\sum_{i=1}^{N} g\left(i/(N+1)\right)\cdot\left(\epsilon_i\Big/\sum_{l=0}^{N}\epsilon_l\right) - \sum_{i=1}^{N} g\left(i/(N+1)\right)\cdot\frac{\epsilon_i}{N+1}\right|$$

$$\leq \frac{1}{N+1}\sum_{i=1}^{N}\left|g\left(i/(N+1)\right)\right|\cdot\epsilon_i\cdot\left|\frac{1}{\sum_{l=0}^{N}\epsilon_l/(N+1)} - 1\right| = o_p(1).$$

Next, consider (A.26). Because $\mathbb{E}[\epsilon_i]=1$, it follows that the expectation of this term is zero. Its variance is $\sum_{i=1}^{N} g(i/N)^2/(N+1)^2 \leq \max_{x\in[0,1]} g(x)^2/(N+1)$ which converges to zero, implying that this term is $o_p(1)$.
Finally, consider (A.27). By continuity of $g(\cdot)$, this non-stochastic term converges to zero. This finishes the proof for part (i)
The same argument shows that

$$\frac{1}{4}\cdot\sum_{i=M}^{N-M} g\left(\sum_{j=0}^{i}\epsilon_j\Big/\sum_{l=0}^{N}\epsilon_l\right)\cdot\left(\sum_{j=-M+1}^{M}\epsilon_{i+j}\Big/\sum_{l=0}^{N}\epsilon_l\right)^2 - \frac{1}{4}\cdot\sum_{i=M}^{N-M} g\left(i/N\right)\cdot\left(\sum_{j=-M+1}^{M}\epsilon_{i+j}\Big/N\right)^2 \xrightarrow{p} 0$$

The expected value for $(\sum_{j=-M+1}^{M}\epsilon_{i+j})$ is $4M^2+2M$. Using the existence of fourth moments one can then show that

$$\frac{1}{4}\cdot\sum_{i=M}^{N-M} g\left(i/N\right)\cdot\left(\sum_{j=-M+1}^{M}\epsilon_{i+j}\Big/N\right)^2 - \frac{1}{4}\cdot\sum_{i=M}^{N-M} g\left(i/N\right)\cdot\left(4M^2+2M\right)^2\Big/N^2 \xrightarrow{p} 0,$$

which in turn implies the second part of the Lemma. □.

Lemma A.8:

$$\frac{1}{N_1}\sum_{i=1}^{N}(1-W_i)\cdot\sigma_0^2(X_i)\cdot K_M(i)^2 - \frac{1}{N_1}\sum_{i=1}^{N}(1-W_i)\cdot\sigma_0^2(X_i)\cdot\left(N_1^2\cdot P_i^2 + N_1\cdot P_i\cdot(1-P_i)\right) \xrightarrow{p} 0.$$

Proof of Lemma A.8: We will show that

$$\left|\frac{1}{N_1}\sum_{i=1}^{N}(1-W_i)\cdot\sigma_0^2(X_i)\cdot\big(K_M(i)^2-N_1^2\cdot P_i^2-N_1\cdot P_i\cdot(1-P_i)\big)\right|$$
$$=\left|\frac{1}{N_1}\sum_{i=1}^{N_0}\sigma_0^2(X_{0,i})\cdot\big(K_{0,M}(i)^2-N_1^2\cdot P_{0,i}^2-N_1\cdot P_{0,i}\cdot(1-P_{0,i})\big)\right|\xrightarrow{p}0.$$

We first prove this for the single match case with $M=1$. With $M=1$, conditional on N_0, N_1, and $\mathbf{X}_1=(W_1X_1,\ldots,W_NX_N)'$ the N_0-vector $K_{0,1}$ with ith element $K_{0,1}(i)$ has a multinomial distribution with parameters $P_{0,1},\ldots,P_{0,N_0-1}$ and N_1. The moment generating function of the first N_0-1 components of the vector K_1 is

$$M(t_1,\ldots,t_{N_0-1})=\mathbb{E}\left[\exp\left(\sum_{i=1}^{N_0-1}K_1(i)\cdot t_i\right)\right]=\left(\sum_{i=1}^{N_0-1}P_{0,i}\cdot\exp(t_i)+P_{0,N_0}\right)^{N_1}.$$

We need the following moments:

$$\mu_{1,i}=\mathbb{E}\left[K_{0,1}(i)\Big|\mathbf{X}_1\right]=N_1\cdot P_{0,i},$$

$$\mu_{2,i}=\mathbb{E}\left[K_{0,1}(i)^2\Big|\mathbf{X}_1\right]=N_1^2\cdot P_{0,i}^2+N_1\cdot P_{0,i}\cdot(1-P_{0,i}),$$

$$\mu_{4,i}=\mathbb{E}\left[K_{0,1}(i)^4\Big|\mathbf{X}_1\right]=N_1\cdot P_{0,i}+7\cdot N_1\cdot(N_1-1)\cdot P_{0,i}^2$$
$$+4\cdot N_1\cdot(N_1-1)\cdot(N_1-2)\cdot P_{0,i}^3+N_1\cdot(N_1-1)\cdot(N_1-2)\cdot(N_1-3)\cdot P_{0,i}^4,$$

$$\mu_{2,2,i,j}=\mathbb{E}\left[K_{0,1}(i)^2\cdot K_{0,1}(j)^2\Big|\mathbf{X}_1\right]=N_1\cdot(N_1-1)\cdot P_{0,i}\cdot P_{0,j}$$
$$+N_1\cdot(N_1-1)\cdot(N_1-2)\cdot P_{0,i}\cdot P_{0,j}^2+N_1\cdot(N_1-1)\cdot(N_1-2)\cdot P_{0,i}^2\cdot P_{0,j}$$
$$+N_1\cdot(N_1-1)\cdot(N_1-2)\cdot(N_1-3)\cdot P_{N_0,i}^2\cdot P_{0,j}^2,$$

where the last expectation is for $i\neq j$.
Then

$$\mathbb{E}\left[\left(\frac{1}{N_1}\sum_{i=1}^{N_0}\sigma_0^2(X_{0,i})\cdot\big(K_{0,1}(i)^2-N_1^2\cdot P_{0,i}^2+N_1\cdot P_{0,i}\cdot(1-P_{0,i})\big)\right)^2\right]$$

$$=\frac{1}{N_1^2}\sum_{i=1}^{N_0}\sum_{j=1}^{N_0}\mathbb{E}\left[\sigma_0^2(X_{0,i})\cdot\sigma_0^2(X_{0,j})\cdot\big(K_{0,1}(i)^2-\mu_2\big)\cdot\big(K_{0,1}(j)^2-\mu_2\big)\right]$$

$$=\frac{1}{N_1^2}\left(\sum_{i=1}^{N_0}\sigma_0^4(X_{0,i})\cdot\mu_{4,i}+\sum_{i=1}^{N_0}\sum_{j\neq i}\sigma_0^2(X_{0,i})\cdot\sigma_0^2(X_{0,j})\cdot(\mu_{2,2,i,j}-\mu_{2,i}\cdot\mu_{2,j})\right)$$

$$=\frac{1}{N_1^2}\sum_{i=1}^{N_0}\sigma_0^4(X_{0,i})\cdot\big(P_{0,i}+7\cdot(N_1-1)\cdot P_{0,i}^2$$
$$+4\cdot(N_1-1)\cdot(N_1-2)\cdot P_{0,i}^3+(N_1-1)\cdot(N_1-2)\cdot(N_1-3)\cdot P_{0,i}^4\big)$$
$$+\frac{1}{N_1}\sum_{i=1}^{N_0}\sum_{j\neq i}\sigma_0^2(X_{0,i})\cdot\sigma_0^2(X_{0,j})\cdot\big(-P_{0,i}\cdot P_{0,j}+(-4N_1+2)\cdot P_{0,i}\cdot P_{0,j}^2$$

$$+(-4N_1+2)\cdot P_{0,i}^2\cdot P_{0,j}$$
$$+((N_1-1)\cdot(N_1-2)\cdot(N_1-3)-N_1^3-2N_1^2+N_1)\cdot P_{0,i}^2\cdot P_{0,j}^2\big).$$

Using the fact that for all $\delta > 0$, $\max_i P_{0,i} = o_p(N^{-1+\delta})$, all sums in this expression can be shown to be converging to zero in probability. For example, choosing $\delta < 1/2$

$$\left|\frac{1}{N_1^2}\sum_{i=1}^{N_0}\sigma_0^4(X_{0,i})\cdot(N_1-1)\cdot(N_1-2)\cdot(N_1-3)\cdot P_{0,i}^4\right|$$
$$\leq \sup_x \sigma_0^4(x)\cdot\left|\frac{1}{N_1^2}N_0\cdot N_1^3\cdot\max_i P_{0,i}\right| = o_p(N^{2-4+4\delta}) = o_p(1).$$

Now consider the case with general M. The marginal distribution of $K_{0,M}(i)$ remains binomial with parameters N_1 and $P_{0,i}$, and so the moments $\mu_{1,i}$, $\mu_{2,i}$, and $\mu_{4,i}$ are as before. The difference is in the moment $\mu_{2,2,i,j}$. There are two possibilities. First, $A_M(i)\cap A_M(j)=\emptyset$. In that case $\mu_{2,2,i,j}$ is as before. Second, $A_M(i)\cap A_M(j)\neq\emptyset$. In that case $|\mu_{2,2,i,j}|\leq\max(\mu_{4,i},\mu_{4,j})$. First we shall show that that out of the $N_0\cdot(N_0-1)$ pairs (i,j) the number of pairs with $A_M(i)\cap A_M(j)\neq\emptyset$ is less than $2\cdot N_0\cdot(M-1)$. Consider the catchment area for unit (i), $A_M((i)) = \big((X_{0,(i)}+X_{0,(i-M)})/2, (X_{0,(i)}+X_{0,(i+M)})/2\big)$. This overlaps only with the catchment areas for units $(i-M+1)$ to $(i+M-1)$, a total of $2(M-1)$ units. Hence the total number of pairs (i,j) with overlap in the catchment areas is less than or equal to $N_0 2(M-1)$. For the units with overlapping catchment areas the correlation between $K_{0,M}(i)$ and $K_{0,M}(j)$ is higher than for units with non-overlapping catchment areas. Hence we can bound the second cross moment for such units from above by $\sqrt{\mu_{4,i}\mu_{4,j}}$. Hence the difference with the expression for $M=1$ is bounded by

$$\frac{1}{N_1^2}N_0\cdot 2(M-1)\cdot\max_i\mu_{4,i}.$$

This converges to zero using the expression for $\mu_{4,i}$ given above. □

Lemma A.9:

$$\frac{1}{N_1}\sum_{i=1}^{N}(1-W_i)\cdot\sigma_0^2(X_i)\cdot\big(N_1^2\cdot P_i^2+N_1\cdot P_i\cdot(1-P_i)\big)\xrightarrow{p} M\cdot\mathbb{E}\left[\sigma_0^2(X)\frac{f_1(X)}{f_0(X)}\middle| W=0\right]$$
$$+\frac{M(2M+1)}{2}\cdot\mathbb{E}\left[\sigma_0^2(X)\cdot\left(\frac{f_1(X)}{f_0(X)}\right)^2\middle| W=0\right].$$

Proof of Lemma A.9: To prove this result we show first that

$$\frac{1}{N_1}\sum_{i=1}^{N}(1-W_i)\cdot N_1\cdot P_i\cdot(1-P_i)\xrightarrow{p} M\cdot\mathbb{E}\left[\sigma_0^2(X)\cdot\frac{f_1(X)}{f_0(X)}\middle| W=0\right], \quad (A.31)$$

and second that

$$\frac{1}{N_1}\sum_{i=1}^{N}(1-W_i)\cdot N_1^2\cdot P_i^2\xrightarrow{p}\frac{p}{2}\cdot M\cdot(2M+1)\cdot\mathbb{E}\left[\sigma_0^2(X)\cdot\left(\frac{f_1(X)}{f_0(X)}\right)^2\middle| W=0\right]. \quad (A.32)$$

the combination of which gives the desired result.
For (A.31) consider:

$$-\frac{1}{N_1}\sum_{i=1}^{N}\left[(1-W_i)\cdot\sigma_0^2(X_i)\cdot N_1\cdot P_i^2\right] = -\sum_{i=1}^{N_0}\sigma_0^2(X_{0,i})\cdot P_{0,i}^2.$$

This can be bounded in absolute value by $N_0 \cdot \sup_x \sigma_0^2(x) \cdot \max_i P_{0,i}^2$. By Lemma A.6 $\max_i P_{0,i} = o_p(N^{-1+\delta})$ for any $\delta > 0$, so $\max_i P_{0,i}^2 = o_p(N_0^{-2+2\delta})$, and $N_0 \cdot \sup_x(\sigma_0^2(x) \cdot \max_i P_{0,i}^2 = o_p(N_0^{-1+2\delta})$ for all $\delta > 0$. Choose $\delta < 1/2$, so that $N_0 \cdot \sup_x \sigma_0^2(x) \cdot \max_i P_{0,i}^2 = o_p(1)$ and thus $\sum_i \sigma_0^2(X_{0,i}) \cdot P_{0,i}^2 = o_p(1)$. Hence in order to show that (A.31) holds it remains to show that

$$\frac{1}{N_1} \sum_{i=1}^{N} (1 - W_i) \cdot \sigma_0^2(X_i) \cdot N_1 \cdot P_i \xrightarrow{p} M \cdot \mathbb{E}\left[\sigma_0^2(X) \cdot \frac{f_1(X)}{f_0(X)}\middle| W = 0\right]. \tag{A.33}$$

We can write

$$\frac{1}{N_1} \sum_{i=1}^{N} (1 - W_i) \cdot \sigma_0^2(X_i) \cdot N_1 \cdot P_i = \sum_{i=1}^{N_0} \sigma_0^2(X_{0,i}) \cdot P_{0,i}$$

$$= \sum_{i=1}^{N_0} \sigma_0^2(X_{0,(i)}) \cdot P_{0,(i)} = \sum_{i=1}^{N_0} \sigma_0^2(X_{0,(i)}) \cdot \tilde{P}_{0,(i)} + o_p(1),$$

$$= \sum_{i=1}^{N_0} \sigma_0^2(X_{0,(i)}) \cdot \frac{f_1(X_{0,(i)})}{f_0(X_{0,(i)})} \cdot \big(F_0(X_{0,(i+M)}) - F_0(X_{0,(i-M)})\big) + o_p(1).$$

where the second to last equality follows from Lemma A.6(i). With X_i iid with cdf $F_0(x)$, the vector $(F_0(X_{(1)}), \ldots, F_0(X_{(N_0)})))'$ has the same distribution as the vector $(\epsilon_1 / \sum_{j=1}^{N_0} \epsilon_j, \ldots, \epsilon_{N_0} / \sum_{j=1}^{N_0} \epsilon_j)'$, with the ϵ_j iid unit exponential. Hence to show (A.33) it suffices to show, for iid unit exponential ϵ_j, that

$$\sum_{i=1}^{N_0} \sigma_0^2(F_0^{-1}(\sum_{j=0}^{i} \epsilon_j / \sum_{l=0}^{N_0} \epsilon_l)) \cdot \frac{f_1(F_0^{-1}(\sum_{j=0}^{i} \epsilon_j / \sum_{l=0}^{N_0} \epsilon_l))}{f_0(F_0^{-1}(\sum_{j=0}^{i} \epsilon_j / \sum_{l=0}^{N_0} \epsilon_l))} \cdot \left(\sum_{j=-M}^{M} \epsilon_{i+j} / \sum_{l=0}^{N_0} \epsilon_l\right)$$

$$\xrightarrow{p} M \cdot \mathbb{E}\left[\sigma_0^2(X) \cdot \frac{f_1(X)}{f_0(X)}\middle| W = 0\right].$$

By Lemma A.7, for $g(\cdot)$ continuously differentiable on $[0, 1]$ and iid unit-exponential ϵ_i we have

$$\frac{1}{2} \cdot \sum_{i=1}^{N_0} g\left(\sum_{j=0}^{i} \epsilon_j \Big/ \sum_{l=0}^{N_0} \epsilon_l\right) \cdot \left(\sum_{j=-M}^{M} \epsilon_{i+j} \Big/ \sum_{l=0}^{N_0} \epsilon_l\right) \xrightarrow{p} M \cdot \int_0^1 g(z) dz.$$

With $g(z) = \sigma_0^2(F_0^{-1}(z)) \cdot f_1(F_0^{-1}(z)) / f_0(F_0^{-1}(z))$, this equals

$$M \cdot \int_0^1 g(z) dz = M \cdot \int_0^1 \sigma_0^2(F_0^{-1}(z)) \cdot \frac{f_1(F_0^{-1}(z))}{f_0(F_0^{-1}(z))} dz$$

$$= M \cdot \int_0^1 \sigma_0^2(z) \frac{f_1(z)}{f_0(z)} dF_0(z) = M \cdot \mathbb{E}\left[\sigma_0^2(X) \cdot \frac{f_1(X)}{f_0(X))}\middle| W = 0\right].$$

Thus (A.33) holds, and therefore (A.31).
For (A.32) we have:

$$\frac{1}{N_1} \sum_{i=1}^{N} (1 - W_i) \cdot \sigma_0^2(X_i) \cdot N_1^2 \cdot P_i^2 = N_1 \cdot \sum_{i=1}^{N_0} \sigma_0^2(X_{0,i}) \cdot P_{0,i}^2 = N_1 \cdot \sum_{i=1}^{N_0} \sigma_0^2(X_{0,i}) \cdot \tilde{P}_{0,i}^2 + o_p(1)$$

$$= N_1 \cdot \sum_{i=1}^{N_0} \sigma_0^2(F_0^{-1}(\sum_{j=0}^{i} \epsilon_j / \sum_{l=0}^{N_0} \epsilon_l)) \cdot \left(\frac{f_1(F_0^{-1}(\sum_{j=0}^{i} \epsilon_j / \sum_{l=0}^{N_0} \epsilon_l))}{f_0(F_0^{-1}(\sum_{j=0}^{i} \epsilon_j / \sum_{l=0}^{N_0} \epsilon_l))}\right)^2 \cdot \left(\frac{\sum_{m=-M+1}^{M} \epsilon_{i+m}}{2 \cdot \sum_{j=0}^{N_0} \epsilon_j}\right)^2 + o_p(1)$$

$$= N_1 \cdot \sum_{i=1}^{N_0} g\left(F_0^{-1}(\sum_{j=0}^{i}\epsilon_j / \sum_{l=0}^{N_0}\epsilon_l)\right) \cdot \left(\frac{\sum_{m=-M+1}^{M}\epsilon_{i+m}}{2 \cdot \sum_{j=0}^{N_0}\epsilon_j}\right)^2 + o_p(1),$$

with $g(z) = \sigma_0^2(z) f_1(z)^2 / f_0(z)^2$. Then using part (ii) of Lemma A.7 shows that (A.32) holds. □
Proof of Theorem 5:
Conditional on $\mathbf{X}$ and $\mathbf{W}$ the variance of the simple matching estimator with M matches is

$$V^E = \frac{1}{N}\sum_{i=1}^{N}\left(1 + \frac{K_M(i)}{M}\right)^2 \sigma_{W_i}^2(X_i)$$

$$= \frac{1}{N}\sum_{i=1}^{N}(1-W_i)\cdot\sigma_0^2(X_i) + 2\cdot\frac{1}{N}\sum_{i=1}^{N}(1-W_i)\cdot\frac{K_M(i)}{M}\cdot\sigma_0^2(X_i) + \frac{1}{N}\sum_{i=1}^{N}(1-W_i)\cdot\left(\frac{K_M(i)}{M}\right)^2\cdot\sigma_0^2(X_i)$$

$$+\frac{1}{N}\sum_{i=1}^{N}W_i\cdot\sigma_1^2(X_i) + 2\cdot\frac{1}{N}\sum_{i=1}^{N}W_i\cdot\frac{K_M(i)}{M}\cdot\sigma_1^2(X_i) + \frac{1}{N}\sum_{i=1}^{N}W_i\cdot\left(\frac{K_M(i)}{M}\right)^2\cdot\sigma_1^2(X_i).$$

First we shall show that this converges to

$$V^{E*} = (1-p)\cdot\mathbb{E}[\sigma_0^2(X)|W=0] + 2\cdot(1-p)\cdot\mathbb{E}\left[\sigma_0^2(X)\cdot\frac{p\cdot f_1(X)}{(1-p)\cdot f_0(X)}\middle| W=0\right]$$

$$+p\cdot\mathbb{E}\left[\frac{\sigma_0^2(X)}{M}\cdot\frac{f_1(X)}{f_0(X)} + \sigma_0^2(X)\cdot\frac{2M+1}{2M}\cdot\left(\frac{f_1(X)}{f_0(X)}\right)^2\cdot\frac{p}{1-p}\middle| W=0\right]$$

$$+p\cdot\mathbb{E}[\sigma_1^2(X)|W=1] + 2\cdot p\cdot\mathbb{E}\left[\sigma_1^2(X)\cdot\frac{(1-p)\cdot f_0(X)}{p\cdot f_1(X)}\middle| W=1\right]$$

$$+(1-p)\cdot\mathbb{E}\left[\frac{\sigma_1^2(X)}{M}\cdot\frac{f_0(X)}{f_1(X)} + \sigma_1^2(X)\cdot\frac{2M+1}{2M}\cdot\left(\frac{f_0(X)}{f_1(X)}\right)^2\cdot\frac{1-p}{p}\middle| W=1\right].$$

We will look at convergence of one of the terms in V^E to the corresponding term in V^{E*}. Specifically, we will show that

$$\frac{1}{N}\sum_{i=1}^{N}(1-W_i)\cdot\left(\frac{K_M(i)}{M}\right)^2\sigma_0^2(X_i)$$

$$\xrightarrow{p} p\cdot\mathbb{E}\left[\sigma_0^2(X)\cdot\frac{1}{M}\cdot\frac{f_1(X)}{f_0(X)} + \sigma_0^2(X)\cdot\frac{2M+1}{2M}\cdot\left(\frac{f_1(X)}{f_0(X)}\right)^2\cdot\frac{p}{1-p}\middle| W=0\right]. \quad \text{(A.34)}$$

The others follow by the same argument.
By Lemma A.8

$$\frac{1}{N_1}\sum_{i=1}^{N}(1-W_i)\cdot\sigma_0^2(X_i)\cdot\left(K_M(i)^2 - N_1^2\cdot P_i^2 - N_1\cdot P_i\cdot(1-P_i)\right) \xrightarrow{p} 0.$$

By Lemma A.9

$$\frac{1}{N_1}\sum_{i=1}^{N}(1-W_i)\cdot\sigma_0^2(X_i)\cdot\left(N_1^2\cdot P_i^2 + N_1\cdot P_i\cdot(1-P_i)\right) \xrightarrow{p} M\cdot\mathbb{E}\left[\sigma_0^2(X)\frac{p\cdot f_1(X)}{(1-p)\cdot f_0(X)}\middle| W=0\right]$$

$$+\frac{M(2M+1)}{2}\cdot\frac{p}{1-p},$$

so that

$$\frac{1}{N\cdot M^2}\sum_{i=1}^{N}(1-W_i)\cdot\sigma_0^2(X_i)\cdot\left(N_1^2\cdot P_i^2+N_1\cdot P_i\cdot(1-P_i)\right)\xrightarrow{p}\frac{p}{M}\cdot\mathbb{E}\left[\sigma_0^2(X)\frac{p\cdot f_1(X)}{(1-p)\cdot f_0(X)}\middle|W=0\right]$$
$$+\frac{2M+1}{2M}\cdot\frac{p}{1-p}.$$

The final step consists of showing that V^{E*} is equal to

$$\mathbb{E}\left[\frac{\sigma_0^2(X)}{1-e(X)}\right]+\frac{1}{2M}\cdot\mathbb{E}\left[\left(\frac{1}{1-e(X)}-(1-e(X))\right)\sigma_0^2(X)\right]$$
$$+\,\mathbb{E}\left[\frac{\sigma_1^2(X)}{e(X)}\right]+\frac{1}{2M}\cdot\mathbb{E}\left[\left(\frac{1}{e(X)}-e(X)\right)\sigma_1^2(X)\right].$$

This follows from simple algebra. For example, collecting the terms with $\sigma_0^2(X)$ and no factor $(1/M)$ in V^{E*}:

$$(1-p)\cdot\mathbb{E}[\sigma_0^2(X)|W=0]+2\cdot(1-p)\cdot\mathbb{E}\left[\sigma_0^2(X)\cdot\frac{p\cdot f_1(X)}{(1-p)\cdot f_0(X)}\middle|W=0\right]$$
$$+p\cdot\mathbb{E}\left[\sigma_0^2(X)\cdot\left(\frac{f_1(X)}{f_0(X)}\right)^2\cdot\frac{p}{1-p}\middle|W=0\right]$$
$$=(1-p)\cdot\mathbb{E}\left[\sigma_0^2(X)\cdot\left(1+2\frac{p\cdot f_1(X)}{(1-p)\cdot f_0(X)}+\frac{p^2\cdot f_1(X)^2}{(1-p)^2\cdot f_0(X)^2}\right)\middle|W=0\right]$$
$$=(1-p)\cdot\mathbb{E}\left[\sigma_0^2(X)\cdot\left(1+2\frac{e(X)}{1-e(X)}+(1-p)\frac{e(X)^2}{(1-e(X))^2}\right)\middle|W=0\right]$$
$$=(1-p)\cdot\mathbb{E}\left[\sigma_0^2(X)\cdot\left(1+2\frac{e(X)}{1-e(X)}+\frac{e(X)^2}{(1-e(X))^2}\right)\middle|W=0\right]$$
$$=(1-p)\cdot\mathbb{E}\left[\frac{\sigma_0^2(X)}{(1-e(X))^2}\middle|W=0\right]$$
$$=\mathbb{E}\left[\frac{\sigma_0^2(X)}{1-e(X)}\right].$$

□

Remainder of proof of Lemma A.5:
Let $\bar{\mu}_4$ be a bound to the fourth centered conditional moments of ε_i. Because $\bar{\mu}_4<\infty$, it is easy to show that the variance of the first term is $o(1)$. The variance of the second term multiplied times J^4 is:

$$E\left[\left(\frac{1}{N}\sum_{i=1}^{N}K_M(i)^q\sum_{j=1}^{J}\left(\varepsilon_{l_j(i)}^2-\sigma_{W_i}^2(X_{l_j(i)})\right)\right)^2\right]$$
$$=\frac{1}{N}E\left[\frac{1}{N}\sum_{i=1}^{N}K_M(i)^{2q}\left(\sum_{j=1}^{J}\left(\varepsilon_{l_j(i)}^2-\sigma_{W_i}^2(X_{l_j(i)})\right)\right)^2\right]$$
$$+\frac{2}{N}E\left[\frac{1}{N}\sum_{i=1}^{N}\sum_{t>i}K_M(i)^qK_M(t)^q\sum_{j=1}^{J}\left(\varepsilon_{l_j(i)}^2-\sigma_{W_i}^2(X_{l_j(i)})\right)\sum_{j=1}^{J}\left(\varepsilon_{l_j(t)}^2-\sigma_{W_i}^2(X_{l_j(t)})\right)\right]$$
$$\leq\frac{1}{N}E\left[\frac{1}{N}\sum_{i=1}^{N}K_M(i)^{2q}J\bar{\mu}_4\right]+\frac{2}{N}E\left[\frac{1}{N}\sum_{i=1}^{N}\left(\max_{i=1\ldots N}K_M(i)^{2q}\right)J\bar{L}(k)\bar{\mu}_4\right]$$
$$=\frac{J\bar{\mu}_4}{N}E\left[K_M(i)^{2q}\right]+\frac{2J\bar{L}(k)\bar{\mu}_4}{\sqrt{N}}E\left[\frac{1}{\sqrt{N}}\max_{i=1\ldots N}K_M(i)^{2q}\right].$$

Using Bonferroni's Inequality:

$$
\begin{aligned}
E\left[\left(\frac{1}{\sqrt{N}}\max_{i=1\dots N} K_M(i)^{2q}\right)^2\right] &= \frac{1}{N}E\left[\max_{i=1\dots N} K_M(i)^{4q}\right] = \frac{1}{N}\sum_{n=0}^{\infty}\Pr\left(\max_{i=1\dots N} K_M(i)^{4q} > n\right) \\
&\leq \frac{1}{N}\sum_{n=0}^{\infty} N\Pr\left(K_M(i)^{4q} > n\right) = E\Big[K_M(i)^{4q}\Big],
\end{aligned}
$$

which is uniformly bounded for all N. Because the first moment of a random variable is bounded by the square-root of the second moment we obtain that $E[N^{-1/2}\max_{i=1\dots N} K_M(i)^{2q}]$ is uniformly bounded for all N, and the second term on the right hand side of equation (A.10) is $o_p(1)$.
Let $\bar{\sigma}^2$ be a uniform bound for the conditional variance of ε_i, and $\bar{\sigma}^4 = (\bar{\sigma}^2)^2$. The variance of the third term on the right hand side of equation (A.10) divided by $4/J^4$ is:

$$
\begin{aligned}
&E\left[\left(\frac{1}{N}\sum_{i=1}^{N} K_M(i)^q \sum_{j=1}^{J}\sum_{h>j}\varepsilon_{l_j(i)}\varepsilon_{l_h(i)}\right)^2\right] = \frac{1}{N}E\left[\frac{1}{N}\sum_{i=1}^{N} K_M(i)^{2q}\left(\sum_{j=1}^{J}\sum_{h>j}\varepsilon_{l_j(i)}\varepsilon_{l_h(i)}\right)^2\right] \\
&+ \frac{2}{N}E\left[\frac{1}{N}\sum_{i=1}^{N}\sum_{t>i} K_M(i)^q K_M(t)^q \left(\sum_{j=1}^{J}\sum_{h>j}\varepsilon_{l_j(i)}\varepsilon_{l_h(i)}\right)\left(\sum_{j=1}^{J}\sum_{h>j}\varepsilon_{l_j(t)}\varepsilon_{l_h(t)}\right)\right] \\
&\qquad \leq \frac{J(J-1)\bar{\sigma}^4}{2N}E\Big[K_M(i)^{2q}\Big] + \frac{2J\bar{L}(k)(J-1)\bar{\sigma}^4}{\sqrt{N}}E\left[\frac{1}{\sqrt{N}}\max_{i=1\dots N} K_M(i)^{2q}\right] = o(1).
\end{aligned}
$$

The last inequality holds because within the same treatment group, each observation is used as a match at most $J\cdot\bar{L}(k)$ times; and because there are $J-1$ possible combinations of that observation and the other $J-1$ matches.
The variance of the fourth term on the right hand side of equation (A.10) divided by $4/J^2$ is:

$$
\begin{aligned}
&E\left[\left(\frac{1}{N}\sum_{i=1}^{N} K_M(i)^q \varepsilon_i \sum_{j=1}^{J}\varepsilon_{l_j(i)}\right)^2\right] = \frac{1}{N}E\left[\frac{1}{N}\sum_{i=1}^{N} K_M(i)^{2q}\left(\varepsilon_i\sum_{j=1}^{J}\varepsilon_{l_j(i)}\right)^2\right] \\
&+ \frac{2}{N}E\left[\frac{1}{N}\sum_{i=1}^{N}\sum_{t>i} K_M(i)^q K_M(t)^q\left(\varepsilon_i\sum_{j=1}^{J}\varepsilon_{l_j(i)}\right)\left(\varepsilon_t\sum_{j=1}^{J}\varepsilon_{l_j(t)}\right)\right] \\
&\qquad \leq \frac{J\bar{\sigma}^4}{N}E\Big[K_M(i)^{2q}\Big] + \frac{2J\bar{L}(k)\bar{\sigma}^4}{\sqrt{N}}E\left[\frac{1}{\sqrt{N}}\max_{i=1\dots N} K_M(i)^{2q}\right] = o(1).
\end{aligned}
$$

The variance of the fifth term on the right hand side of equation (A.10) divided by $4/J^2$ is:

$$E\left[\left(\frac{1}{N}\sum_{i=1}^{N}K_M(i)^q\varepsilon_i\sum_{j=1}^{J}\mu_{W_i}(X_i)-\mu_{W_i}(X_{l_j(i)})\right)^2\right]$$

$$=\frac{1}{N}E\left[\frac{1}{N}\sum_{i=1}^{N}K_M(i)^{2q}\left(\varepsilon_i\sum_{j=1}^{J}\mu_{W_i}(X_i)-\mu_{W_i}(X_{l_j(i)})\right)^2\right]$$

$$+\frac{2}{N}E\left[\frac{1}{N}\sum_{i=1}^{N}\sum_{t>i}K_M(i)^qK_M(t)^q\left(\varepsilon_i\sum_{j=1}^{J}\mu_{W_i}(X_i)-\mu_{W_i}(X_{l_j(i)})\right)\left(\varepsilon_t\sum_{j=1}^{J}\mu_{W_t}(X_t)-\mu_{W_t}(X_{l_j(t)})\right)\right]$$

$$\leq\frac{\bar{\sigma}^2}{N}E\left[K_M(i)^{2q}\left(\sum_{j=1}^{J}\mu_{W_i}(X_i)-\mu_{W_i}(X_{l_j(i)})\right)^2\right]=o(1).$$

The variance of the sixth term on the right hand side of equation (A.10) divided by $4/J^4$ is:

$$E\left[\left(\frac{1}{N}\sum_{i=1}^{N}K_M(i)^q\sum_{j=1}^{J}\varepsilon_{l_j(i)}\sum_{j=1}^{J}\mu_{W_i}(X_i)-\mu_{W_i}(X_{l_j(i)})\right)^2\right]$$

$$=\frac{1}{N}E\left[\frac{1}{N}\sum_{i=1}^{N}K_M(i)^{2q}\left(\sum_{j=1}^{J}\varepsilon_{l_j(i)}\right)^2\left(\sum_{j=1}^{J}\mu_{W_i}(X_i)-\mu_{W_i}(X_{l_j(i)})\right)^2\right]$$

$$+\frac{2}{N}E\Bigg[\frac{1}{N}\sum_{i=1}^{N}\sum_{t>i}K_M(i)^qK_M(t)^q\left(\sum_{j=1}^{J}\varepsilon_{l_j(i)}\sum_{j=1}^{J}\mu_{W_i}(X_i)-\mu_{W_i}(X_{l_j(i)})\right)$$

$$\times\left(\sum_{j=1}^{J}\varepsilon_{l_j(t)}\sum_{j=1}^{J}\mu_{W_t}(X_t)-\mu_{W_t}(X_{l_j(t)})\right)\Bigg]$$

$$\leq\frac{J\bar{\sigma}^2}{N}E\left[\frac{1}{N}\sum_{i=1}^{N}K_M(i)^{2q}\max_{i=1\ldots N}\left(\sum_{j=1}^{J}\mu_{W_i}(X_i)-\mu_{W_i}(X_{l_j(i)})\right)^2\right]$$

$$+\frac{2J\bar{L}(k)\bar{\sigma}^2}{\sqrt{N}}E\left[\frac{1}{\sqrt{N}}\max_{i=1\ldots N}K_M(i)^{2q}\max_{i=1\ldots N}\left(\sum_{j=1}^{J}\mu_{W_i}(X_i)-\mu_{W_i}(X_{l_j(i)})\right)^2\right]=o(1).$$

As a result, the variance of each term on the right hand side of equation (A.10) is $o(1)$, so equation (A.9) holds. This result, along with the result in equation (A.8) guarantees that the result of the lemma holds. Here we also prove the second part of the lemma. The analysis from the beginning of this proof to equation (A.7) applies here without change. Recall also from the proof of Lemma 3 that for any $q\geq 1$, $(N_0/N_1)\mathbb{E}[K_M(i)^q|W_i=0]$ is uniformly bounded. As a result, we obtain:

$$\left|\frac{1}{N_1}\sum_{i=1}^{N}(1-W_i)K_M(i)^q\Big(\mathbb{E}\big[\widehat{\sigma}^2_{W_i}(X_i)|\mathbf{X},\mathbf{W}\big]-\sigma^2_{W_i}(X_i)\Big)\right|$$

$$\leq\max_{W_i=0}\left|\mathbb{E}\big[\widehat{\sigma}^2_{W_i}(X_i)|\mathbf{X},\mathbf{W}\big]-\sigma^2_{W_i}(X_i)\right|\left(\frac{N_0}{N_1}\right)E\Big[K_M(i)^q\Big|W_i=0\Big]=o_p(1).\quad\text{(A.35)}$$

To obtain the second result of the lemma, it is left to be proven that for $q \geq 1$:

$$\frac{1}{N_1}\sum_{i=1}^{N}(1-W_i)K_M(i)^q\Big(E\big[\widehat{\sigma}^2_{W_i}(X_i)|\mathbf{X},\mathbf{W}\big]-\widehat{\sigma}^2_{W_i}(X_i)\Big)=o_p(1). \tag{A.36}$$

Notice that:

$$\left(\frac{J+1}{J}\right)\frac{1}{N_1}\sum_{W_i=0}K_M(i)^q\Big(E\big[\widehat{\sigma}^2_{W_i}(X_i)|\mathbf{X},\mathbf{W}\big]-\widehat{\sigma}^2_{W_i}(X_i)\Big)=\frac{1}{N_1}\sum_{W_i=0}K_M(i)^q\Big(\varepsilon_i^2-\sigma^2_{W_i}(X_i)\Big)$$
$$+\frac{1}{N_1}\sum_{W_i=0}K_M(i)^q\frac{1}{J^2}\sum_{j=1}^{J}\Big(\varepsilon^2_{l_j(i)}-\sigma^2_{W_i}(X_{l_j(i)})\Big)+\frac{2}{J^2N_1}\sum_{W_i=0}K_M(i)^q\sum_{j=1}^{J}\sum_{h>j}\varepsilon_{l_j(i)}\varepsilon_{l_h(i)}$$
$$-\frac{2}{JN_1}\sum_{W_i=0}K_M(i)^q\varepsilon_i\sum_{j=1}^{J}\varepsilon_{l_j(i)}+\frac{2}{JN_1}\sum_{W_i=0}K_M(i)^q\varepsilon_i\sum_{j=1}^{J}\Big(\mu_{W_i}(X_i)-\mu_{W_i}(X_{l_j(i)})\Big)$$
$$-\frac{2}{J^2N_1}\sum_{W_i=0}K_M(i)^q\sum_{j=1}^{J}\varepsilon_{l_j(i)}\sum_{j=1}^{J}\Big(\mu_{W_i}(X_i)-\mu_{W_i}(X_{l_j(i)})\Big). \tag{A.37}$$

The means of the terms on the right-hand side of equation (A.37) are zero. It is left to be shown that the variances are $o(1)$. The variance of the first term on the right hand side of equation (A.37) is

$$\begin{aligned}E\left[\left(\frac{1}{N_1}\sum_{W_i=0}K_M(i)^q\Big(\varepsilon_i^2-\sigma^2_{W_i}(X_i)\Big)\right)^2\right] &= \frac{1}{N_1}E\left[\frac{1}{N_1}\sum_{W_i=0}K_M(i)^{2q}\Big(\varepsilon_i^2-\sigma^2_{W_i}(X_i)\Big)^2\right]\\ &\leq \frac{\bar{\mu}_4}{N_1}\left(\frac{N_0}{N_1}\right)E\Big[K_M(i)^{2q}\Big|W_i=0\Big]=o_p(1).\end{aligned}$$

The variance of the second term multiplied times J^4 is:

$$\begin{aligned}&E\left[\left(\frac{1}{N_1}\sum_{W_i=0}K_M(i)^q\sum_{j=1}^{J}\Big(\varepsilon^2_{l_j(i)}-\sigma^2(X_{l_j(i)})\Big)\right)^2\right]\\
&=\frac{1}{N_1}E\left[\frac{1}{N_1}\sum_{W_i=0}K_M(i)^{2q}\left(\sum_{j=1}^{J}\Big(\varepsilon^2_{l_j(i)}-\sigma^2(X_{l_j(i)})\Big)\right)^2\right]\\
&+\frac{2}{N_1}E\left[\frac{1}{N_1}\sum_{W_i=0}K_M(i)^q\sum_{W_t=0,t>i}K_M(t)^q\sum_{j=1}^{J}\Big(\varepsilon^2_{l_j(i)}-\sigma^2(X_{l_j(i)})\Big)\sum_{j=1}^{J}\Big(\varepsilon^2_{l_j(t)}-\sigma^2(X_{l_j(t)})\Big)\right]\\
&\leq\frac{J\bar{\mu}_4}{N_1}\left(\frac{N_0}{N_1}\right)E\Big[K_M(i)^{2q}\Big|W_i=0\Big]+\frac{2J(J\bar{L}(k)-1)\bar{\mu}_4}{N_1}E\left[\frac{1}{N_1}\sum_{W_i=0}K_M(i)^q\max_{W_t=0}K_M(t)^q\right]\\
&=\frac{2J(J\bar{L}(k)-1)\bar{\mu}_4}{N_1}E\left[\frac{1}{N_1}\sum_{W_i=0}1_{[1,\infty)}(K_M(i))K_M(i)^q\max_{W_t=0}K_M(t)^q\right]+o(1)\\
&\leq\frac{2J(J\bar{L}(k)-1)\bar{\mu}_4}{N_1}E\left[\frac{1}{N_1}\left(\max_{W_i=0}K_M(t)^{2q}\right)\sum_{W_i=0}1_{[1,\infty)}(K_M(i))\right]+o(1)\\
&\leq\frac{2J(J\bar{L}(k)-1)M\bar{\mu}_4}{\sqrt{N_1}}E\left[\frac{1}{\sqrt{N_1}}\max_{W_i=0}K_M(t)^{2q}\right]+o(1),\end{aligned}$$

where 1_A is the indicator function of the set A (that is $1_A(x) = 1$ if $x \in A$, zero otherwise). Using Bonferroni's Inequality:

$$\begin{aligned} E\left[\left(\frac{1}{\sqrt{N_1}}\max_{W_i=0} K_M(i)^{2q}\right)^2\right] &= \frac{1}{N_1}E\left[\max_{W_i=0} K_M(i)^{4q}\right] \\ &= \frac{1}{N_1}\sum_{n=0}^{\infty}\Pr\left(\max_{W_i=0} K_M(i)^{4q} > n\right) \\ &\leq \frac{1}{N_1}\sum_{n=0}^{\infty} N_0 \Pr\left(K_M(i)^{4q} > n \middle| W_i = 0\right) \\ &= \left(\frac{N_0}{N_1}\right) E\Big[K_M(i)^{4q} \Big| W_i = 0\Big], \end{aligned}$$

which is uniformly bounded. Because the first moment of a random variable is bounded by the square-root of the second moment we obtain that $E[N_1^{-1/2}\max_{W_i=0} K_M(i)^{2q}]$ is uniformly bounded, and the second term on the right hand side of equation (A.37) is $o_p(1)$. The variance of the third term on the right hand side of equation (A.37) divided by $4/J^4$ is:

$$\begin{aligned} & E\left[\left(\frac{1}{N_1}\sum_{W_i=0} K_M(i)^q \sum_{j=1}^{J}\sum_{h>j}\varepsilon_{l_j(i)}\varepsilon_{l_h(i)}\right)^2\right] = \frac{1}{N_1}E\left[\frac{1}{N_1}\sum_{W_i=0} K_M(i)^{2q}\left(\sum_{j=1}^{J}\sum_{h>j}\varepsilon_{l_j(i)}\varepsilon_{l_h(i)}\right)^2\right] \\ & + \frac{2}{N_1}E\left[\frac{1}{N}\sum_{W_i=0} K_M(i)^q \sum_{W_t=0,t>i} K_M(t)^q \left(\sum_{j=1}^{J}\sum_{h>j}\varepsilon_{l_j(i)}\varepsilon_{l_h(i)}\right)\left(\sum_{j=1}^{J}\sum_{h>j}\varepsilon_{l_j(t)}\varepsilon_{l_h(t)}\right)\right] \\ & \leq \frac{J(J-1)\bar{\sigma}^4}{2N_1}\left(\frac{N_0}{N_1}\right)E\Big[K_M(i)^{2q}\Big| W_i = 0\Big] + \frac{J(J-1)(J\bar{L}(k)-1)\bar{\sigma}^4}{N_1}E\left[\frac{1}{N_1}\sum_{W_i=0} K_M(i)^q \max_{W_t=0} K_M(t)^q\right] \\ & = \frac{J(J-1)(J\bar{L}(k)-1)\bar{\sigma}^4}{N_1}E\left[\frac{1}{N_1}\sum_{W_i=0} 1_{[1,\infty)}(K_M(i))K_M(i)^q \max_{W_t=0} K_M(t)^q\right] + o(1) \\ & \leq \frac{J(J-1)(J\bar{L}(k)-1)\bar{\sigma}^4}{N_1}E\left[\max_{W_i=0} K_M(i)^{2q}\frac{1}{N_1}\sum_{W_i=0} 1_{[1,\infty)}(K_M(i))\right] + o(1) \\ & \qquad \leq \frac{J(J-1)(J\bar{L}(k)-1)M\bar{\sigma}^4}{\sqrt{N_1}}E\left[\frac{1}{\sqrt{N_1}}\max_{W_i=0} K_M(i)^{2q}\right] + o(1) = o(1). \end{aligned}$$

The variance of the fourth term on the right hand side of equation (A.37) divided by $4/J^2$ is:

$$\begin{aligned} & E\left[\left(\frac{1}{N_1}\sum_{W_i=0} K_M(i)^q \varepsilon_i \sum_{j=1}^{J}\varepsilon_{l_j(i)}\right)^2\right] = \frac{1}{N_1}E\left[\frac{1}{N_1}\sum_{W_i=0} K_M(i)^{2q}\left(\varepsilon_i\sum_{j=1}^{J}\varepsilon_{l_j(i)}\right)^2\right] \\ & + \frac{2}{N_1}E\left[\frac{1}{N_1}\sum_{W_i=0} K_M(i)^q \sum_{W_t=0,t>i} K_M(t)^q \left(\varepsilon_i\sum_{j=1}^{J}\varepsilon_{l_j(i)}\right)\left(\varepsilon_t\sum_{j=1}^{J}\varepsilon_{l_j(t)}\right)\right] \\ & \quad \leq \frac{J\bar{\sigma}^4}{N_1}\left(\frac{N_0}{N_1}\right)E\Big[K_M(i)^{2q}\Big| W_i = 0\Big] + \frac{2J^2\bar{L}(k)\bar{\sigma}^4}{N_1}E\left[\frac{1}{N_1}\sum_{W_i=0} K_M(i)^q \max_{W_t=0} K_M(i)^q\right] = o(1). \end{aligned}$$

The variance of the fifth term on the right hand side of equation (A.37) divided by $4/J^2$ is:

$$E\left[\left(\frac{1}{N_1}\sum_{W_i=0} K_M(i)^q \varepsilon_i \sum_{j=1}^{J} \mu_{W_i}(X_i) - \mu_{W_i}(X_{l_j(i)})\right)^2\right]$$

$$= \frac{1}{N_1} E\left[\frac{1}{N_1}\sum_{W_i=0} K_M(i)^{2q} \left(\varepsilon_i \sum_{j=1}^{J} \mu_{W_i}(X_i) - \mu_{W_i}(X_{l_j(i)})\right)^2\right]$$

$$+\frac{2}{N_1} E\left[\frac{1}{N_1}\sum_{W_i=0} K_M(i)^q \sum_{t>i} K_M(t)^q \left(\varepsilon_i \sum_{j=1}^{J} \mu_{W_i}(X_i) - \mu_{W_i}(X_{l_j(i)})\right)\left(\varepsilon_t \sum_{j=1}^{J} \mu_{W_t}(X_t) - \mu_{W_t}(X_{l_j(t)})\right)\right]$$

$$\leq \frac{\bar{\sigma}^2}{N_1} E\left[\frac{1}{N_1}\sum_{W_i=0} K_M(i)^{2q} \left(\sum_{j=1}^{J} \mu_{W_i}(X_i) - \mu_{W_i}(X_{l_j(i)})\right)^2\right] = o(1).$$

The variance of the sixth term on the right hand side of equation (A.37) divided by $4/J^4$ is:

$$E\left[\left(\frac{1}{N_1}\sum_{W_i=0} K_M(i)^q \sum_{j=1}^{J} \varepsilon_{l_j(i)} \sum_{j=1}^{J} \mu_{W_i}(X_i) - \mu_{W_i}(X_{l_j(i)})\right)^2\right]$$

$$= \frac{1}{N_1} E\left[\frac{1}{N_1}\sum_{W_i=0} K_M(i)^{2q} \left(\sum_{j=1}^{J} \varepsilon_{l_j(i)}\right)^2 \left(\sum_{j=1}^{J} \mu_{W_i}(X_i) - \mu_{W_i}(X_{l_j(i)})\right)^2\right]$$

$$+ \frac{2}{N_1} E\Bigg[\frac{1}{N_1}\sum_{W_i=0} K_M(i)^q \sum_{W_t=0,t>i} K_M(t)^q \left(\sum_{j=1}^{J} \varepsilon_{l_j(i)} \sum_{j=1}^{J} \mu_{W_i}(X_i) - \mu_{W_i}(X_{l_j(i)})\right)$$

$$\times \left(\sum_{j=1}^{J} \varepsilon_{l_j(t)} \sum_{j=1}^{J} \mu_{W_t}(X_t) - \mu_{W_t}(X_{l_j(t)})\right)\Bigg]$$

$$\leq \frac{J\bar{\sigma}^2}{N_1} E\left[\frac{1}{N_1}\sum_{W_i=0} K_M(i)^{2q} \max_{W_t=0} \left(\sum_{j=1}^{J} \mu_{W_t}(X_t) - \mu_{W_t}(X_{l_j(t)})\right)^2\right]$$

$$+\frac{2J(J\bar{L}(k)-1)\bar{\sigma}^2}{N_1} E\left[\frac{1}{N_1}\sum_{W_i=0} K_M(i)^q \max_{W_t=0} K_M(t)^q \max_{W_t=0} \left(\sum_{j=1}^{J} \mu_{W_t}(X_t) - \mu_{W_t}(X_{l_j(t)})\right)^2\right] = o(1).$$

As a result, the variance of each term on the right hand side of equation (A.37) is $o(1)$, so equation (A.36) holds. This result, along with the result in equation (A.35) guarantees that the result of the lemma holds.□

Reminder of proof of Theorem 8: For the variance of $\widehat{\tau}_M^{sm,t}$ we obtain:

$$\begin{aligned}\frac{1}{N_1}\sum_{W_i=1}\left(Y_i - \widehat{Y}_i(0) - \widehat{\tau}_M^{sm,t}\right)^2 &= \frac{1}{N_1}\sum_{W_i=1}\left(Y_i - \widehat{Y}_i(0) - \tau^t\right)^2 - (\widehat{\tau}_M^{sm,t} - \tau^t)^2 \\ &= \frac{1}{N_1}\sum_{W_i=1}\left(Y_i - \widehat{Y}_i(0) - \tau^t\right)^2 + o_p(1). \qquad \text{(A.38)}\end{aligned}$$

In addition,

$$\frac{1}{N_1}\sum_{W_i=1}\left(Y_i-\widehat{Y}_i(0)-\tau^t\right)^2$$
$$=\frac{1}{N_1}\sum_{W_i=1}\left(\frac{1}{M}\sum_{m=1}^{M}\mu_1(X_i)-\mu_0(X_{j_m(i)})-\tau^t\right)^2+\frac{1}{N_1}\sum_{W_i=1}\left(\varepsilon_i-\frac{1}{M}\sum_{m=1}^{M}\varepsilon_{j_m(i)}\right)^2$$
$$+\frac{2}{N_1}\sum_{W_i=1}\left(\frac{1}{M}\sum_{m=1}^{M}\mu_1(X_i)-\mu_0(X_{j_m(i)})-\tau^t\right)\left(\varepsilon_i-\frac{1}{M}\sum_{m=1}^{M}\varepsilon_{j_m(i)}\right). \quad \text{(A.39)}$$

Because the sample maximum of the norms of the matching discrepancies, $\|X_i-X_{j_m(i)}\|$, is $o_p(1)$, and the regression function μ_0, is Lipschitz, we obtain

$$\frac{1}{N_1}\sum_{W_i=1}\left(\frac{1}{M}\sum_{m=1}^{M}\mu_0(X_i)-\mu_0(X_{j_m(i)})\right)^2=o_p(1). \quad \text{(A.40)}$$

Consider the first term on the right hand side of equation (A.39):

$$\frac{1}{N_1}\sum_{W_i=1}\left(\frac{1}{M}\sum_{m=1}^{M}\mu_1(X_i)-\mu_0(X_{j_m(i)})-\tau^t\right)^2$$
$$=\frac{1}{N_1}\sum_{W_i=1}\left((\mu_1(X_i)-\mu_0(X_i)-\tau^t)+\left(\frac{1}{M}\sum_{m=1}^{M}\mu_0(X_i)-\mu_0(X_{j_m(i)})\right)\right)^2$$
$$=\frac{1}{N_1}\sum_{W_i=1}(\mu_1(X_i)-\mu_0(X_i)-\tau^t)^2+\frac{1}{N_1}\sum_{W_i=1}\left(\frac{1}{M}\sum_{m=1}^{M}\mu_0(X_i)-\mu_0(X_{j_m(i)})\right)^2$$
$$+\frac{1}{N_1}\sum_{W_i=1}(\mu_1(X_i)-\mu_0(X_i)-\tau^t)\left(\frac{1}{M}\sum_{m=1}^{M}\mu_0(X_i)-\mu_0(X_{j_m(i)})\right)$$
$$=\frac{1}{N_1}\sum_{W_i=1}(\mu_1(X_i)-\mu_0(X_i)-\tau^t)^2+o_p(1), \quad \text{(A.41)}$$

by Hölder's Inequality and equation (A.40). Next, consider the second term on the right hand side of equation (A.39):

$$\frac{1}{N_1}\sum_{W_i=1}\left(\varepsilon_i-\frac{1}{M}\sum_{m=1}^{M}\varepsilon_{j_m(i)}\right)^2=\frac{1}{N_1}\sum_{W_i=1}\left(\varepsilon_i^2+\frac{1}{M^2}\left(\sum_{m=1}^{M}\varepsilon_{j_m(i)}^2+2\sum_{m=1}^{M}\sum_{n>m}\varepsilon_{j_m(i)}\varepsilon_{j_n(i)}\right)-\frac{2}{M}\sum_{m=1}^{M}\varepsilon_i\varepsilon_{j_m(i)}\right)$$

Therefore,

$$\frac{1}{N_1}\sum_{W_i=1}\left(\varepsilon_i-\frac{1}{M}\sum_{m=1}^{M}\varepsilon_{j_m(i)}\right)^2-\frac{1}{N_1}\sum_{i=1}^{N}\left(W_i+(1-W_i)\frac{K_M(i)}{M^2}\right)\sigma^2(X_i,W_i)$$
$$=\frac{1}{N_1}\sum_{i=1}^{N}\left(W_i+(1-W_i)\frac{K_M(i)}{M^2}\right)\left(\varepsilon_i^2-\sigma^2(X_i,W_i)\right)$$
$$+\frac{1}{N_1}\sum_{W_i=1}\frac{2}{M^2}\left(\sum_{m=1}^{M}\sum_{n>m}\varepsilon_{j_m(i)}\varepsilon_{j_n(i)}\right)-\frac{1}{N_1}\sum_{W_i=1}\frac{2}{M}\sum_{m=1}^{M}\varepsilon_i\varepsilon_{j_m(i)}. \quad \text{(A.42)}$$

The expectations conditional on $\mathbf{X}$ and $\mathbf{W}$ of each of the three terms on the right hand side of last expression are zero, so the unconditional expectations are also zero. Because the fourth conditional moments of ε_i are uniformly bounded, and because $(N_0/N_1)E[K_M(i)^2|W_i = 0]$ is uniformly bounded (see Proof of Lemma 3), we obtain:

$$E\left[\left(\frac{1}{N_1}\sum_{i=1}^{N}\left(W_i + (1-W_i)\frac{K_M(i)}{M^2}\right)\left(\varepsilon_i^2 - \sigma^2(X_i, W_i)\right)\right)^2\right]$$

$$= \frac{1}{N_1}E\left[\frac{1}{N_1}\sum_{i=1}^{N}\left(W_i + (1-W_i)\frac{K_M(i)^2}{M^4}\right)\left(\varepsilon_i^2 - \sigma^2(X_i, W_i)\right)^2\right] = o(1).$$

The variance of the second term divided by $4/M^4$ is

$$\begin{aligned} E\left[\left(\frac{1}{N_1}\sum_{W_i=1}\sum_{m=1}^{M}\sum_{n>m}\varepsilon_{j_m(i)}\varepsilon_{j_n(i)}\right)^2\right] &= \frac{1}{N_1}E\left[\frac{1}{N_1}\sum_{W_i=1}\left(\sum_{m=1}^{M}\sum_{n>m}\varepsilon_{j_m(i)}\varepsilon_{j_n(i)}\right)^2\right] \\ &+ \frac{2}{N_1}E\left[\frac{1}{N_1}\sum_{W_i=1}\sum_{W_j=1,j>i}\left(\sum_{m=1}^{M}\sum_{n>m}\varepsilon_{j_m(i)}\varepsilon_{j_n(i)}\right)\left(\sum_{m=1}^{M}\sum_{n>m}\varepsilon_{j_m(j)}\varepsilon_{j_n(j)}\right)\right] \\ &\le \frac{1}{N_1}E\left[\frac{1}{N_1}\sum_{W_i=1}\frac{(M-1)M}{2}\bar{\sigma}^4\right] \\ &+ \frac{1}{N_1}E\left[\frac{1}{N_1}\sum_{W_i=0}(M-1)K_M(i)(K_M(i)-1)\bar{\sigma}^4\right] = o(1). \end{aligned}$$

The variance of the third term divided by $4/M^2$ is

$$\begin{aligned} E\left[\left(\frac{1}{N_1}\sum_{W_i=1}\sum_{m=1}^{M}\varepsilon_i\varepsilon_{j_m(i)}\right)^2\right] &= \frac{1}{N_1}E\left[\frac{1}{N_1}\sum_{W_i=1}\left(\sum_{m=1}^{M}\varepsilon_i\varepsilon_{j_m(i)}\right)^2\right] \\ &+ \frac{2}{N_1}E\left[\frac{1}{N_1}\sum_{W_i=1}\sum_{W_i=1,j>i}\left(\sum_{m=1}^{M}\varepsilon_i\varepsilon_{j_m(i)}\right)\left(\sum_{m=1}^{M}\varepsilon_j\varepsilon_{j_m(j)}\right)\right] \\ &\le \frac{1}{N_1}E\left[\frac{1}{N_1}\sum_{W_i=1}M\bar{\sigma}^4\right] = o(1). \end{aligned}$$

As a result, we obtain

$$\frac{1}{N_1}\sum_{W_i=1}\left(\varepsilon_i - \frac{1}{M}\sum_{m=1}^{M}\varepsilon_{j_m(i)}\right)^2 - \frac{1}{N_1}\sum_{W_i=1}\left(W_i + (1-W_i)\frac{K_M(i)}{M^2}\right)\sigma^2(X_i, W_i) = o_p(1).$$

Finally consider the last term on the right hand side of equation (A.39). Let

$$\Psi_{M,i}^t = \left(\frac{1}{M}\sum_{m=1}^{M}\mu_1(X_i) - \mu_0(X_{j_m(i)}) - \tau^t\right).$$

Notice that there is a finite bound $\overline{\Psi^t}$, such that $|\Psi_{M,i}^t| \le \overline{\Psi^t}$ for all i. The conditional expectation of the last term of equation (A.39) is zero, so the unconditional expectation is also zero. The conditional variance of this term (divided by 4) is:

$$
\begin{aligned}
&E\left[\frac{1}{N_1^2}\sum_{W_i=1}\sum_{W_j=1}\Psi^t_{M,i}\Psi^t_{M,j}\left(\varepsilon_i-\frac{1}{M}\sum_{m=1}^{M}\varepsilon_{j_m(i)}\right)\left(\varepsilon_j-\frac{1}{M}\sum_{m=1}^{M}\varepsilon_{j_m(j)}\right)\right]\\
&\leq\left|E\left[\frac{1}{N_1^2}\sum_{W_i=1}\sum_{W_j=1}\Psi^t_{M,i}\Psi^t_{M,j}\varepsilon_i\varepsilon_j\right]\right|+2\left|E\left[\frac{1}{MN_1^2}\sum_{W_i=1}\sum_{W_j=1}\Psi^t_{M,i}\Psi^t_{M,j}\varepsilon_i\sum_{m=1}^{M}\varepsilon_{j_m(j)}\right]\right|\\
&\qquad+\left|E\left[\frac{1}{M^2N_1^2}\sum_{W_i=1}\sum_{W_j=1}\Psi^t_{M,i}\Psi^t_{M,j}\sum_{m=1}^{M}\varepsilon_{j_m(i)}\sum_{m=1}^{M}\varepsilon_{j_m(j)}\right]\right|.
\end{aligned}
$$

Notice that:

$$
\left|E\left[\frac{1}{N_1^2}\sum_{W_i=1}\sum_{W_j=1}\Psi^t_{M,i}\Psi^t_{M,j}\varepsilon_i\varepsilon_j\right]\right|=\left|E\left[\frac{1}{N_1^2}\sum_{W_i=1}(\Psi^t_{M,i})^2\varepsilon_i^2\right]\right|\leq\frac{\overline{\Psi^t}^2}{N_1}E\left[\frac{1}{N_1}\sum_{W_i=1}\varepsilon_i^2\right]=o(1),
$$

$$
\left|E\left[\frac{1}{MN_1^2}\sum_{W_i=1}\sum_{W_j=1}\Psi^t_{M,i}\Psi^t_{M,j}\varepsilon_i\sum_{m=1}^{M}\varepsilon_{j_m(j)}\right]\right|=0,
$$

$$
\begin{aligned}
&\left|E\left[\frac{1}{M^2N_1^2}\sum_{W_i=1}\sum_{W_j=1}\Psi^t_{M,i}\Psi^t_{M,j}\sum_{m=1}^{M}\varepsilon_{j_m(i)}\sum_{m=1}^{M}\varepsilon_{j_m(j)}\right]\right|\leq\left|E\left[\frac{1}{M^2N_1^2}\sum_{W_i=1}(\Psi^t_{M,i})^2\left(\sum_{m=1}^{M}\varepsilon_{j_m(i)}\right)^2\right]\right|\\
&+2\left|E\left[\frac{1}{M^2N_1^2}\sum_{W_i=1}\sum_{W_j=1,j>i}\Psi^t_{M,i}\Psi^t_{M,j}\sum_{m=1}^{M}\varepsilon_{j_m(i)}\sum_{m=1}^{M}\varepsilon_{j_m(j)}\right]\right|\leq\frac{\overline{\Psi^t}^2}{N_1}E\left[\frac{1}{MN_1}\sum_{W_i=1}\bar{\sigma}^2\right]\\
&\qquad+\frac{\overline{\Psi^t}^2}{N_1}E\left[\frac{1}{M^2N_1}\sum_{W_i=0}K_M(i)(K_M(i)-1)\varepsilon_i^2\right]=o(1).
\end{aligned}
$$

As a result, we obtain:

$$
\begin{aligned}
\frac{1}{N_1}\sum_{W_i=1}\left(\widehat{Y}_i(1)-\widehat{Y}_i(0)-\widehat{\tau}^{sm,t}_M\right)^2 &= \frac{1}{N_1}\sum_{W_i=1}(\mu_1(X_i)-\mu_0(X_i)-\tau^t)^2\\
&+ \frac{1}{N_1}\sum_{i=1}^{N}\left(W_i+(1-W_i)\frac{K_M(i)}{M^2}\right)\sigma^2(X_i,W_i)+o_p(1)
\end{aligned}
$$

Applying the previous lemma:

$$
\left|\frac{1}{N_1}\sum_{i=1}^{N}\left(W_i+(1-W_i)\frac{K_M(i)}{M^2}\right)\sigma^2(X_i,W_i)-\frac{1}{N_1}\sum_{i=1}^{N}\left(W_i+(1-W_i)\frac{K_M(i)}{M^2}\right)\widehat{\sigma}^2(X_i,W_i)\right|=o_p(1).
$$

Therefore,

$$
\widehat{V}^{\tau(X),t}=\frac{1}{N_1}\sum_{W_i=1}\left(Y_i-\widehat{Y}_i(0)-\widehat{\tau}^{sm,t}_M\right)^2-\frac{1}{N_1}\sum_{i=1}^{N}\left(W_i+(1-W_i)\frac{K_M(i)}{M^2}\right)\widehat{\sigma}^2(X_i,W_i)\xrightarrow{p}V^{\tau(X),t}.
$$

□

Additional References

Gradshteyn, I.S. and I.M. Ryzhik, (2000), *Tables of Integrals, Series, and Products*, 6th edition, Academic Press, San Diego.