Supplement: Proofs

Lu Lin[1,5], Feng Li[2], Kangning Wang[3] and Lixing Zhu[4,6]

[1] *Shandong University, Jinan, China*

[2] *Zhengzhou University, Zhengzhou, China*

[3] *Shandong Technology and Business University, Yantai, China*

[4] *Hong Kong Baptist University, Hong Kong, Cnina*

[5] *Qufu Normal University, Qufu, Cnina*

[6] *Beijing Normal University, Beijing, Cnina*

Proof of Theorem 2.1. We only prove the theorem for the AWLS estimator (2.4). For (2.3), the proof is similar and thus is omitted here. It can be easily verified that $\tilde{\theta} = \sum\limits_{k=1}^{m} \tilde{w}_k \hat{\theta}_{\tau_k}$,

$$\sum_{k=1}^{m} \Big(w_k - \bar{\xi} \frac{w_k(\xi(\tau_k) - \bar{\xi})}{\sum_{k=1}^{m} w_k(\xi(\tau_k) - \bar{\xi})^2} \Big) \theta = \theta,$$

$$\sum_{k=1}^{m} \Big(w_k - \bar{\xi} \frac{w_k(\xi(\tau_k) - \bar{\xi})}{\sum_{k=1}^{m} w_k(\xi(\tau_k) - \bar{\xi})^2} \Big) \xi(\tau_k) = 0.$$

Then

$$\begin{aligned}
\tilde{\theta} &= \sum_{k=1}^{m} \Big(w_k - \bar{\xi} \tfrac{w_k(\xi(\tau_k) - \bar{\xi})}{\sum_{k=1}^{m} w_k(\xi(\tau_k) - \bar{\xi})^2} \Big) \hat{\theta}_{\tau_k} \\
&= \sum_{k=1}^{m} \Big(w_k - \bar{\xi} \tfrac{w_k(\xi(\tau_k) - \bar{\xi})}{\sum_{k=1}^{m} w_k(\xi(\tau_k) - \bar{\xi})^2} \Big) (\theta + \xi(\tau_k)\varphi_n + \epsilon_n(\tau_k)) \\
&= \theta + \sum_{k=1}^{m} \Big(w_k - \bar{\xi} \tfrac{w_k(\xi(\tau_k) - \bar{\xi})}{\sum_{k=1}^{m} w_k(\xi(\tau_k) - \bar{\xi})^2} \Big) \epsilon_n(\tau_k),
\end{aligned}$$

which completes the proof. □

Proof of Theorem 2.2. (1) We first give the proof for the case when b_n is known. By the definition of $\tilde{\theta}$ and the second inequality in (C1), we have

$$\begin{aligned}\tilde{\theta} &= \sum_{k=1}^{m} w_k \left(\hat{\theta}_{\tau_k} - b_n \xi_n(\tau_k, \hat{\theta}_{\tau_k})\right) \\ &= \sum_{k=1}^{m} w_k \left(\hat{\theta}_{\tau_k} - b_n[\xi_n(\tau_k, \theta) + \xi_n'(\tau_k, \theta)(\hat{\theta}_{\tau_k} - \theta)] + o_p(\hat{\theta}_{\tau_k} - \theta)\right) \\ &= \theta + \sum_{k=1}^{m} w_k \epsilon_n(\tau_k) + \sum_{k=1}^{m} w_k \left(-b_n \xi_n'(\tau_k, \theta)(\hat{\theta}_{\tau_k} - \theta) + o_p(\hat{\theta}_{\tau_k} - \theta)\right).\end{aligned}$$

The first inequality in (C1) and above result leads to the conclusion of the theorem.

(2) When b_n is unknown, we consider its estimator as

$$\begin{aligned}\hat{b}_n &= \frac{\sum_{k=1}^{m} w_k \hat{\theta}_{\tau_k} \left(\hat{\xi}_n(\tau_k) - \bar{\hat{\xi}}_n\right)}{\sum_{k=1}^{m} w_k \left(\hat{\xi}_n(\tau_k) - \bar{\hat{\xi}}_n\right)^2} \\ &= \frac{\sum_{k=1}^{m} w_k \left(\theta + b_n \xi_n(\tau_k) + \epsilon_n(\tau_k)\right) \left(\hat{\xi}_n(\tau_k) - \bar{\hat{\xi}}_n\right)}{\sum_{k=1}^{m} w_k \left(\hat{\xi}_n(\tau_k) - \bar{\hat{\xi}}_n\right)^2} \\ &= \frac{\sum_{k=1}^{m} w_k \left(b_n \xi_n(\tau_k) + \epsilon_n(\tau_k)\right) \left(\hat{\xi}_n(\tau_k) - \bar{\hat{\xi}}_n\right)}{\sum_{k=1}^{m} w_k \left(\hat{\xi}_n(\tau_k) - \bar{\hat{\xi}}_n\right)^2}.\end{aligned}$$

Note that

$$\xi_n(\tau_k) = \xi_n(\tau_k, \theta) = \hat{\xi}_n(\tau_k) + \hat{\xi}_n'(\tau_k)(\theta - \hat{\theta}_{\tau_k}) + o_p(\theta - \hat{\theta}_{\tau_k}).$$

Then

$$\begin{aligned}\hat{b}_n &= \frac{\sum_{k=1}^m w_k \left[b_n \left(\hat{\xi}_n(\tau_k) + \hat{\xi}_n'(\tau_k)(\theta - \hat{\theta}_{\tau_k}) + o_p(\theta - \hat{\theta}_{\tau_k})\right) + \epsilon_n(\tau_k)\right] \left(\hat{\xi}_n(\tau_k) - \bar{\hat{\xi}}_n\right)}{\sum_{k=1}^m w_k \left(\hat{\xi}_n(\tau_k) - \bar{\hat{\xi}}_n\right)^2} \\ &= b_n + \frac{\sum_{k=1}^m w_k \left[b_n \left(\hat{\xi}_n'(\tau_k)(\theta - \hat{\theta}_{\tau_k}) + o_p(\theta - \hat{\theta}_{\tau_k})\right) + \epsilon_n(\tau_k)\right] \left(\hat{\xi}_n(\tau_k) - \bar{\hat{\xi}}_n\right)}{\sum_{k=1}^m w_k \left(\hat{\xi}_n(\tau_k) - \bar{\hat{\xi}}_n\right)^2} \\ &= b_n + \frac{\sum_{k=1}^m w_k \left(o_p(\theta - \hat{\theta}_{\tau_k}) + \epsilon_n(\tau_k)\right) \left(\hat{\xi}_n(\tau_k) - \bar{\hat{\xi}}_n\right)}{\sum_{k=1}^m w_k \left(\hat{\xi}_n(\tau_k) - \bar{\hat{\xi}}_n\right)^2} \\ &= b_n + o_p(\theta - \hat{\theta}_{\tau_k}).\end{aligned}$$

Thus, by the above result and the method given in the proof (1), we can prove the conclusion of the theorem. □

Proof of Theorem 2.3. We only prove the conclusion for the case when b_n is unknown. For the other case, the proof is more simple.

When b_n is unknown and $\xi_n(\tau)$ depends on θ, the AWLS estimator is given in (2.3), i.e., the estimator has the form as $\tilde{\theta} = \sum_{k=1}^m w_k \hat{\theta}_{\tau_k} - \hat{b}_n \bar{\hat{\xi}}_n$. As shown in the proof of Theorem 2.2, $\hat{b}_n = b_n + o_p(\theta - \hat{\theta}_{\tau_k})$. Similarly,

$$\bar{\hat{\xi}}_n = \sum_{k=1}^m w_k \hat{\xi}_n(\tau_k) = \sum_{k=1}^m w_k \left(\xi_n(\tau_k) + \xi_n'(\tau_k)(\hat{\theta}_{\tau_k} - \theta) + o_p(\hat{\theta}_{\tau_k} - \theta)\right).$$

Then,

$$\begin{aligned}\tilde{\theta} &= \sum_{k=1}^m w_k \hat{\theta}_{\tau_k} - b_n \sum_{k=1}^m w_k \xi_n(\tau_k) - b_n \sum_{k=1}^m w_k \xi_n'(\tau_k)(\hat{\theta}_{\tau_k} - \theta) + b_n o_p(\hat{\theta}_{\tau_k} - \theta) \\ &= -b_n \sum_{k=1}^m w_k \xi_n'(\tau_k)(\hat{\theta}_{\tau_k} - \theta) + b_n o_p(\hat{\theta}_{\tau_k} - \theta) + \epsilon_n(\tau_k) \\ &= -\sum_{k=1}^m w_k g(\tau_k)(\hat{\theta}_{\tau_k} - \theta) + b_n O_p(\hat{\theta}_{\tau_k} - \theta) + b_n o_p(\hat{\theta}_{\tau_k} - \theta) + \epsilon_n(\tau_k).\end{aligned}$$

This completes the proof for the first conclusion.

Note that

$$\mathbf{w}^* = \min_{\mathbf{1}^T\mathbf{w}_g=1} \mathbf{w}_g^T \lim \Sigma_{\hat{\boldsymbol{\theta}}} \mathbf{w}_g.$$

By Lagrange multipliers, we see that the optimal weight vector has the following closed representation:

$$\mathbf{w}^* = \left(\mathbf{1}^T(\lim \Sigma_{\hat{\boldsymbol{\theta}}})^{-1}\mathbf{1}\right)^{-1} (\lim \Sigma_{\hat{\boldsymbol{\theta}}})^{-1}\mathbf{1},$$

which completes the proof for the second result of the theorem. □

Proof of Theorem 3.1. Clearly,

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^n X_i(\tau - I(Y_i - X_i^T\hat{\beta}_\tau \le \hat{b}_\tau)) \\
&= \frac{1}{n}\sum_{i=1}^n X_i\{(Z_i - E(Z_i|X_i, \hat{\beta}_\tau, \hat{b}_\tau)) + E(Z_i|X_i, \hat{\beta}_\tau, \hat{b}_\tau) + (\tau - I(Y_i - X_i^T\beta \le b_\tau))\},
\end{aligned}$$

where $Z_i = \tau - I(e_i \le \hat{b}_\tau + X_i^T(\hat{\beta}_\tau - \beta)) - (\tau - I(e \le b_\tau))$ and $E(Z_i|X_i, \hat{\beta}_\tau, \hat{b}_\tau) = \tau - F_e(\hat{b}_\tau + X_i^T(\hat{\beta}_\tau - \beta)) - (\tau - F_e(b_\tau))$. To apply the standard empirical process to prove the asymptotic properties, we only need to get that for any $\tilde{\beta}$ and $\tilde{b}_\tau$ such that $\|\tilde{\beta} - \beta\| = O(1/\sqrt{n})$ and $|\tilde{b}_\tau - b_\tau| = O(1/\sqrt{n})$, the variance of $X_i(Z_i - E(Z_i|X_i, \tilde{\beta}_\tau, \tilde{b}_\tau))$ is of order n^{-1}. This result can be easily computed. Thus, by the standard empirical process theory (see Chapter II. Theorem 37 of Pollard 1984), the first term $\frac{1}{n}\sum_{i=1}^n X_i(Z_{1i} - E(Z_i|X_i, \hat{\beta}_\tau, \hat{b}_\tau))$ has the convergence rate of order $O_p(n^{-3/4})$. Further, by (C4), we have

$$E(Z_i|X_i, \hat{\beta}_\tau, \hat{b}_\tau) = -(\hat{b}_\tau - b_\tau)f_e(b_\tau) - (\hat{\beta}_\tau - \beta)^T X_i f_e(b_\tau) + O_p(n^{-1}).$$

Since X is centralized and $\hat{b}_\tau - b_\tau = O_p(n^{-1/2})$,

$$\frac{1}{n}\sum_{i=1}^n X_i(\hat{b}_\tau - b_\tau) = O_p(n^{-1}).$$

By the definition of D_n,

$$D_n^{-1}\frac{1}{n}\sum_{i=1}^n X_i X_i^T(\hat{\beta}_\tau - \beta) = \hat{\beta}_\tau - \beta.$$

Consequently,

$$\frac{1}{n}D_n^{-1}\sum_{i=1}^{n} X_i E(Z_i|X_i,\hat{\beta}_\tau,\hat{b}_\tau) = -(\hat{\beta}_\tau-\beta)f_e(b_\tau)+O_p(n^{-1}).$$

Combing the results above leads to

$$\begin{aligned}&\frac{1}{n}D_n^{-1}\sum_{i=1}^{n} X_i(\tau - I(Y_i - X_i^T\hat{\beta}_\tau \le \hat{b}_\tau))\\ &= \frac{1}{n}D_n^{-1}\sum_{i=1}^{n} X_i(\tau - I(Y_i - X_i^T\beta \le b_\tau)) - (\hat{\beta}_\tau-\beta)f_e(b_\tau)+o_p(n^{-3/4}).\end{aligned}$$

It follows from the result above and the Bahadur representation given in Example 1 that

$$\begin{aligned}\tilde{\beta} &= \sum_{k=1}^{m} w_k\Big\{\hat{\beta}_{\tau_k} - \frac{1}{f_e(b_{\tau_k})n}D_n^{-1}\sum_{i=1}^{n} X_i(\tau_k - I(Y_i \le \hat{b}_{\tau_k}+\hat{\beta}_{\tau_k}^T X_i))(1+o_p(n^{-1/2}))\Big\}\\ &= \sum_{k=1}^{m} w_k\Big\{\hat{\beta}_{\tau_k} - \frac{1}{f_e(b_{\tau_k})n}D_n^{-1}\sum_{i=1}^{n} X_i(\tau_k - I(Y_i \le b_{\tau_k}+\beta^T X_i)) + \hat{\beta}_{\tau_k} - \beta\Big\}\\ &\quad +o_p(n^{-3/4})\\ &= \sum_{k=1}^{m} w_k\Big\{\hat{\beta}_{\tau_k} - \frac{1}{f_e(b_{\tau_k})n}D_n^{-1}\sum_{i=1}^{n} X_i(\tau_k - I(Y_i \le b_{\tau_k}+\beta^T X_i)) + \hat{\beta}_{\tau_k} - \beta\Big\}\\ &\quad +o_p(n^{-3/4})\\ &= \beta + \sum_{k=1}^{m} w_k(\hat{\beta}_{\tau_k}-\beta) + \sum_{k=1}^{m} w_k\epsilon_n(\tau_k) + o_p(n^{-3/4})\\ &\quad +o_p(n^{-3/4})\\ &= \beta + \sum_{k=1}^{m} w_k(\hat{\beta}_{\tau_k}-\beta) + O_p(n^{-3/4}).\end{aligned}$$

Note that $\hat{\beta}_{\tau_k}$ has the same Bahadur representation as given in Example 1. By the representation, we finish the proof. □

Proof of Theorem 3.2. We first prove the result in (1). The estimator $\tilde{r}_1(x)$

in (3.5) can be rewritten as

$$\begin{aligned}\tilde{r}_1(x) &= \frac{1}{n}\sum_{i=1}^{n}\tilde{W}_i Y_i \\ &= r(x) + \frac{1}{n}\sum_{i=1}^{n}\tilde{W}_i e_i + \frac{1}{n}\sum_{i=1}^{n}\tilde{W}_i(r(X_i)-r(x)) \\ &=: r(x) + \tilde{r}_{11}(x) + \tilde{r}_{12}(x),\end{aligned}$$

where $\tilde{W}_i = \sum\limits_{k=1}^{m} g_k \frac{K_{\tau_k}(X_i-x)}{\hat{f}_X(x)}$, $\hat{f}_X(x) = \frac{1}{n}\sum\limits_{i=1}^{n} K_{\tau_k}(X_i - x)$ and $g_k = w_k - \overline{\tau^2}\frac{w_k(\tau_k^2-\overline{\tau^2})}{\sum\limits_{k=1}^{m} w_k(\tau_k^2-\overline{\tau^2})^2}$, with $\sum_{k=1}^{m} g_k = 1$ due to the constraint that $\sum_{k=1}^{m} w_k = 1$. Note that $\hat{f}_X(x) \xrightarrow{p} f_X(x)$ at the rate of $n^{-2\eta}$ (or $n^{-4\eta}$) under Condition (C5) (or Condition (C6)). We only need to consider the property of $\tilde{r}_{11}^*(x) = \frac{1}{n}\sum\limits_{i=1}^{n} W_i e_i$ and $\tilde{r}_{12}^*(x) = \frac{1}{n}\sum\limits_{i=1}^{n} W_i(r(X_i) - r(x))$, where $W_i = \sum\limits_{k=1}^{m} g_k K_{\tau_k}(X_i - x)$. From the standard technique (see e.g. Härdle 1990), it is to see that $\tilde{r}_{11}^*(x)$ has zero mean and variance

$$\begin{aligned}&\frac{\sigma^2}{n}E\Big(\sum_{k=1}^{m} g_k K_{\tau_k}(X-x)\Big)^2 \\ &= \frac{\sigma^2 f_X(x)}{n^{1-\eta}}\sum_{k=1}^{m}\sum_{j=1}^{m} g_k V_j\,\tau_k^{-1}\tau_j^{-1}\int K(u/\tau_k)K(u/\tau_j)du + o(1/n^{1-\eta}) \\ &= \frac{\sigma^2 f_X(x)}{n^{1-\eta}}\mathbf{w}^T A_1(\mathbf{w})\mathbf{w} + o(1/n^{1-\eta}).\end{aligned}$$

Then $\sqrt{n^{1-\eta}}\,\tilde{r}_{11}^*(x) \xrightarrow{\mathcal{D}} N\Big(0, \sigma^2\mathbf{w}^T A_1(\mathbf{w})\mathbf{w} f_X(x)\Big)$ and consequently

$$\sqrt{n^{1-\eta}}\,\tilde{r}_{11}(x) \xrightarrow{\mathcal{D}} N\Big(0, \sigma^2\mathbf{w}^T A_1(\mathbf{w})\mathbf{w}/f_X(x)\Big).$$

Further, from the definition of g_k above, it is easy to see that $\sum\limits_{k=1}^{m} g_k\tau_k^2 = 0$. Because of this, when Condition (C6) holds (noticing the symmetry of the kernel function), the application of Taylor expansion yields $\tilde{r}^*_{12}(x)$ has mean

$$\begin{aligned} E(\tilde{r}^*_{12}(x)) &= \sum_{k=1}^{m} g_k E\Big(K_{\tau_k}(X-x)(r(X)-r(x))\Big) \\ &= \sum_{k=1}^{m} g_k \tau_k^2 n^{-2\eta} b(x) f_X(x) + \sum_{k=1}^{m} g_k \tau_k^4 n^{-4\eta} c(x) f_X(x) + o(n^{-4\eta}) \\ &= \sum_{k=1}^{m} g_k \tau_k^4 n^{-4\eta} c(x) f_X(x) + o(n^{-4\eta}), \end{aligned}$$

where $b(x)$ and $c(x)$ are known functions. When the weaker Condition (C5) is assumed, the term $c_n(x) = \sum_{k=1}^{m} g_k \tau_k^4 n^{-4\eta} c(x) f_X(x) + o(n^{-4\eta})$ should be a term $o(n^{-2\eta})$ instead. Namely, the mean becomes $o(n^{-2\eta})$. Furthermore, under only Condition (C5), $\tilde{r}^*_{12}(x)$ has the variance

$$\begin{aligned} &\frac{1}{n} Var\Big(\sum_{k=1}^{m} g_k K_{\tau_k}(X-x)(r(X)-r(x))\Big) \\ &= \frac{1}{n} O\left(E\Big(\sum_{k=1}^{m} g_k K_{\tau_k}(X-x)(r(X)-r(x))\Big)^2 \right) \\ &= O\Big(\frac{1}{n^{1+3\eta}}\Big), \end{aligned}$$

which is of smaller order than $O\Big(\frac{1}{n^{1-\eta}}\Big)$. Thus for $c_n(x) = o(n^{-2\eta})$ or $c_n(x) = n^{-4\eta} c(x) \sum_{k=1}^{m} g_k \tau_k^4$ according to Condition (C5) or Condition (C6)

$$\sqrt{n^{1-\eta}}\Big(\tilde{r}^*_{12}(x) - c_n(x) f_X(x)\Big) \xrightarrow{p} 0,$$

implying

$$\sqrt{n^{1-\eta}}\Big(\tilde{r}_{12}(x) - c_n(x)\Big) \xrightarrow{p} 0.$$

Therefore, combining the above results leads to the asymptotic normality of the first estimator $\tilde{r}_1(x)$.

Now we prove the asymptotic normality for the AWLS estimator $\tilde{r}_2(x)$ defined in (3.7) under Condition (C6). By the Bahadur representation, we obtain

$$\begin{aligned}
\tilde{r}_2(x) &= \sum_{k=1}^{m} w_k \Big\{ \hat{r}_{\tau_k}(x) - n^{-1} v_{\tau_k}^{-1}(x) \sum_{i=1}^{n} K_{\tau_k}(X_i - x)(Y_i - \hat{r}_{\tau_k}(x)) \Big\} \\
&= \sum_{k=1}^{m} w_k \Big\{ \hat{r}_{\tau_k}(x) - n^{-1} v_{\tau_k}^{-1}(x) \sum_{i=1}^{n} K_{\tau_k}(X_i - x)(Y_i - r_{\tau_k}(x)) \Big\} \\
&\quad + n^{-1} \sum_{k=1}^{m} w_k v_{\tau_k}^{-1}(x) \sum_{i=1}^{n} K_{\tau_k}(X_i - x)(\hat{r}_{\tau_k}(x) - r(x)) \\
&= r(x) + n^{-2} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{m} \frac{w_k K_{\tau_k}(X_i - x) K_{\tau_k}(X_j - x) e_j}{v_{\tau_k}(x) \hat{f}_X(x)} \\
&\quad + n^{-2} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{m} \frac{w_k K_{\tau_k}(X_i - x) K_{\tau_k}(X_j - x)(r(X_j) - r(x))}{v_{\tau_k}(x) \hat{f}_X(x)} \\
&\quad + O_p(n^{-3(1-\eta)/4}) \\
&=: r(x) + \tilde{r}_{21}(x) + \tilde{r}_{22}(x) + O_p(n^{-3(1-\eta)/4}). \qquad (0.1)
\end{aligned}$$

Here the last term in (0.1) is of smaller order than $o_p(n^{-(1-\eta)/2})$, i.e., $O_p(n^{-3(1-\eta)/4}) = o_p(n^{-(1-\eta)/2})$. Note that $v_\tau \to f_X(x)$, and $\hat{f}_X(x) \xrightarrow{P} f_X(x)$ at the rate of $n^{-2\eta}$. Similar to the above proof for $\tilde{r}_1(x)$, we also only need to study the property of

$$\begin{aligned}
\tilde{r}_{21}^*(x) &= n^{-2} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{m} w_k K_{\tau_k}(X_i - x) K_{\tau_k}(X_j - x) e_j \quad \text{and} \\
\tilde{r}_{22}^*(x) &= n^{-2} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{m} w_k K_{\tau_k}(X_i - x) K_{\tau_k}(X_j - x)(r(X_j) - r(x)).
\end{aligned}$$

Consider $\tilde{r}_{21}^*(x)$ first. It can be rewritten as

$$\begin{aligned}
\tilde{r}_{21}^*(x) &= \frac{1}{2n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{m} w_k K_{\tau_k}(X_i - x) K_{\tau_k}(X_j - x)(e_j + e_i) \\
&= \frac{1}{n(n-1)} \sum\sum_{1 \le i < j \le n} \sum_{k=1}^{m} w_k K_{\tau_k}(X_i - x) K_{\tau_k}(X_j - x)(e_j + e_i) + O_p(1/n) \\
&=: \frac{1}{2} \tilde{r}_{21}^0(x) + O_p(1/n),
\end{aligned}$$

where

$$\tilde{r}_{21}^0(x) = \frac{2}{n(n-1)} \sum\sum_{1\le i<j\le n} \sum_{k=1}^m w_k K_{\tau_k}(X_i - x) K_{\tau_k}(X_j - x)(e_j + e_i).$$

$\tilde{r}_{21}^0(x)$ is an U-statistic with the kernel

$$h(X_1, e_1; X_2, e_2) = \sum_{k=1}^m w_k K_{\tau_k}(X_1 - x) K_{\tau_k}(X_2 - x)(e_2 + e_1).$$

The kernel has zero mean, its projection onto (X_1, e_1) is

$$E(h(X_1, e_1; X_2, e_2)|X_1, e_1) = f_X(x) \sum_{k=1}^m w_k K_{\tau_k}(X_1 - x)e_1 + o(n^{-\eta})$$

and the variance of the projection is

$$\begin{aligned}&\sigma^2 f_X^3(x) n^{-\eta} \sum_{k=1}^m \sum_{j=1}^m w_k w_j\, \tau_k^{-1} \tau_j^{-1} \int K\Big(\frac{u}{\tau_k}\Big) K\Big(\frac{u}{\tau_j}\Big) du + o(n^{-\eta}) \\ &= \sigma^2 f_X^3(x) n^{-\eta} \mathbf{w}^T A_2(\mathbf{w})\mathbf{w} + o(n^{-\eta}).\end{aligned}$$

Then, from the asymptotic normality of U-statistic (see, e.g.. Serfling (1980), Theorem 5.5.1 A), it follows that $\sqrt{n^{1-\eta}}\, \tilde{r}_{21}^0(x) \xrightarrow{\mathcal{D}} N\Big(0, 4\sigma^2 \mathbf{w}^T A_2(\mathbf{w})\mathbf{w} f_X^3(x)\Big)$. Thus, $\sqrt{n^{1-\eta}}\, \tilde{r}_{21}^*(x) \xrightarrow{\mathcal{D}} N\Big(0, \sigma^2 \mathbf{w}^T A_2(\mathbf{w})\mathbf{w} f_X^3(x)\Big)$ and consequently

$$\sqrt{n^{1-\eta}}\, \tilde{r}_{21}(x) \xrightarrow{\mathcal{D}} N\Big(0, \sigma^2 \mathbf{w}^T A_2(\mathbf{w})\mathbf{w}/f_X(x)\Big).$$

We now consider $\tilde{r}_{22}^*(x)$. It is easy to see that it mean is the following:

$$\begin{aligned}&\sum_{k=1}^m w_k E\Big(K_{\tau_k}(X_1 - x) K_{\tau_k}(X_2 - x)(r(X_2) - r(x))\Big) + O(1/n) \\ &= \sum_{k=1}^m w_k \tau_k^2 d(x) f_X^2(x) n^{-2\eta} + o(n^{-2\eta}),\end{aligned}$$

for a given function $d(x)$. The variance of $\tilde{r}_{22}^*(x)$ can be written as

$$\begin{aligned}&n^{-4} E\left(\sum_{i=1}^n \sum_{j=1}^n a_{ij}\right)^2 - E^2(\tilde{r}_{22}^*(x)) \\ &= n^{-4} \sum_{i=1}^n \sum_{j=1}^n E(a_{ij}^2) + n^{-4} \sum\sum\sum\sum_{i\neq j, i\neq i', i\neq j', j\neq j'} E\left(a_{ii'} a_{jj'}\right) - E^2(\tilde{r}_{22}(x)),\end{aligned}$$

where

$$a_{ij} = \sum_{k=1}^{m} w_k K_{\tau_k}(X_i - x) K_{\tau_k}(X_j - x)(r(X_j) - r(x)).$$

We can easily verify that, under Condition (C5), $n^{-4} \sum_{i=1}^{n} \sum_{j=1}^{n} E a_{ij}^2$ is of order $O\left(\frac{1}{n^{2(1+\eta)}}\right)$. Moreover, for $i \neq j, i \neq i', i \neq j'$ and $j \neq j'$, we have $E(a_{ii'})E(a_{jj'}) = E^2(\tilde{r}_{22}(x)) + O\left(\frac{1}{n}\right)$. Thus the variance of $\tilde{r}_{22}^*(x)$ is of order $O\left(\frac{1}{n^{2(1+\eta)}}\right)$, which is of smaller order than $O\left(\frac{1}{n^{1-\eta}}\right)$. Consequently,

$$\sqrt{n^{1-\eta}}\Big(\tilde{r}_{22}^*(x) - \sum_{k=1}^{m} w_k \tau_k^2 d(x) f_X^2(x) n^{-2\eta}\Big) \xrightarrow{p} 0,$$

implying

$$\sqrt{n^{1-\eta}}\Big(\tilde{r}_{22}(x) - \sum_{k=1}^{m} w_k \tau_k^2 d(x) n^{-2\eta}\Big) \xrightarrow{p} 0.$$

The above results together yields the asymptotic normality of the second AWLS estimator $\tilde{r}_2(x)$. □

References

Pollard, D. (1984). *Convergence of stochastic processes.* New York, Berlin Heidelberg, Tokyo.

Serfling, R. J. (1980). *Approximation Theorem of Mathematical Statistics.* John Wiley & Sons, Inc.

Zhu, L.-X. (1993). Convergence rates of empirical processes indexed by classes of functions and their applications. *J. Syst. Sci. Math. Sci.*, **13**, 33 - 41.