ESTIMATION OF SINGLE-INDEX MODELS WITH FIXED CENSORED RESPONSES

Hailin Huang[1], Yuanzhang Li[1], Hua Liang[1] and Yanlin Tang[2]

[1]*George Washington University and* [2]*East China Normal University*

Supplementary Material

The Supplementary Material includes the technical proofs of Proposition 1 and Theorems 1-3.

S1. Proof of Proposition 1

Proof. It is easy to verify that $Y_i = I(Y_i^* > 0)Y_i^* = I(Y_i > 0)Y_i$, where $I(\cdot)$ is the indicator function. Then, the proof follows by a routine calculation as

$$
\begin{aligned}
E(Y_i|X_i^\top\beta) &= E\{Y_i I(Y_i > 0)|X_i^\top\beta\} = E\{Y_i^* I(Y_i^* > 0)|X_i^\top\beta\} \\
&= \int_{-\infty}^{m(X_i^\top\beta)} \{m(X_i^\top\beta) - \epsilon_i\} f(\epsilon_i) d\epsilon_i \\
&= m(X_i^\top\beta)F(m(X_i^\top\beta)) - \int_{-\infty}^{m(X_i^\top\beta)} \epsilon_i f(\epsilon_i) d\epsilon_i \\
&= m(X_i^\top\beta)F(m(X_i^\top\beta)) - \left\{\epsilon_i F(\epsilon_i)|_{-\infty}^{m(X_i^\top\beta)} - \int_{-\infty}^{m(X_i^\top\beta)} F(\epsilon_i) d\epsilon_i\right\} \\
&= \int_{-\infty}^{m(X_i^\top\beta)} F(\epsilon_i) d\epsilon_i.
\end{aligned}
$$

S2. Proof of Theorem 1

Liang et al. (2010) considered model

$$Y_i = r(X_i^\top \beta) + Z_i^\top \alpha + e_i, i = 1, 2, \ldots, n, \tag{S2.1}$$

and our model can be expressed as a special case of model (S2.1), with $\alpha = 0$, $e_i = -\epsilon_i'$ and $r(u) = w \circ m(u)$, where $w(t) = \int_{-\infty}^{t} F(\epsilon)d\epsilon$. To prove Theorem 1, we only need to verify the assumptions for their Theorem 1. Under our Assumptions A.1-A.4, we can easily verify Conditions (i)-(v) in Liang et al. (2010), and we don't need their condition (vi) since we use the Moore-Penrose inverse of the matrix W_0.

S3. Proof of Theorem 2

Before presenting the proof of Theorem 2, we prove three lemmas to facilitate the proof. Lemma 1 is used to prove Lemma 2, which is used to prove Lemma 3.

By using the profile least-squares principle and applying Theorem 1, we can obtain a root$-n$ consistent estimator $\hat{\beta}$ of β_0. Thus, all calculations in this section, unless stated otherwise, correspond to $u = x^\top \beta$, $x \in D_X$ and $\beta \in \Theta_{c_0} = \{\beta : ||\beta - \beta_0|| \leq c_0 n^{-1/2}\}$ for some $c_0 > 0$; similar justification can be found in Zhu and Xue (2006) and Wang et al. (2010). We define

$$W_{ni}(t;\beta) = \frac{K_{h_2}(r(X_i^\top\beta)-t)[S_{n,2}(t;\beta,h_2)-\{r(X_i^\top\beta)-t\}S_{n,1}(t;\beta,h_2)]}{S_{n,0}(t;\beta,h_2)S_{n,2}(t;\beta,h_2)-S_{n,1}^2(t;\beta,h_2)},$$

where $S_{n,l}(t;\beta,h_2)=\sum_{j=1}^n\{r(X_j^\top\beta)-t\}^lK_{h_2}(r(X_j^\top\beta)-t)$ for $l=0,1,2$.

Lemma 1. *Suppose Assumptions A.1- A.4 hold, and $r(\cdot)$ is a known function. Then, for $i=1,\ldots,,n$, $x\in D_X$ and $\beta\in\Theta_{c_0}$, we have*

$$E\Big\{q(r(X_i^\top\beta_0))-\sum_{j=1}^n W_{nj}(r(X_i^\top\beta_0);\beta_0)q(r(X_j^\top\beta_0))\Big\}^2=O(h_2^4),$$
$$E\Big\{q(r(x^\top\beta))-\sum_{j=1}^n W_{nj}(r(x^\top\beta);\beta)q(r(X_j^\top\beta))\Big\}^2=O(h_2^4),$$
$$E\Big\{\sum_{j=1}^n W_{nj}^2(r(x^\top\beta);\beta)\Big\}=O((nh_2)^{-1}).$$

Proof. Under Assumptions A.1-A.4, the Conditions $C_1,C_2,C_3(i)$ of Wang et al. (2010) and Conditions 1-3 of Zhu and Xue (2006) are satisfied. Since here $r(\cdot)$ is assumed to be known, proof then follows by simply replacing $T_i=X_i^\top\beta$ with $T_i=r(X_i^\top\beta)$ in Lemmas 1 and 2 of Zhu and Xue (2006). □

Lemma 2. *Suppose Assumptions A.1-A.4 hold, and $r(\cdot)$ is a known function. Then, for $\beta\in\Theta_{c_0}$, we have*

$$E|\hat{q}(r(X_i^\top\beta);r(\cdot),\beta)-q(r(X_i^\top\beta))|=O(h_2^2+(nh_2)^{-1/2}),\ i=1,\ldots,n.$$

Proof. When $r(\cdot)$ is known, by a routine calculation, we have that, $\forall u=$

$x^\top\beta$, $\sum_{j=1}^n W_{nj}(r(u);\beta)\equiv 1$, and

$$\hat{q}(r(X_i^\top\beta);r(\cdot),\beta)=\sum_{j=1}^n W_{nj}(r(X_i^\top\beta);\beta)I(Y_j>0).$$

Let $I(Y_i>0)=q(r(X_i^\top\beta_0))+e_i$, where e_i is the error term in second stage estimation for $q(\cdot)$. By Assumption A.2(iii), e_i^2 is bounded. For notational convenience, we define $u_i=X_i^\top\beta, u_{i,0}=X_i^\top\beta_0$, $i=1,\ldots,n$. Then,

$$\begin{aligned}\hat{q}(r(u_i);r(\cdot),\beta)-q(r(u_i)) &= \sum_{j=1}^n W_{nj}(r(u_i);\beta)I(Y_j>0)-q(r(u_i))\\ &= \sum_{j=1}^n W_{nj}(r(u_i);\beta)\{q(r(u_{j,0}))+e_j\}-q(r(u_i))\\ &= \sum_{j=1}^n W_{nj}(r(u_i);\beta)\{q(r(u_{j,0}))-q(r(u_i))+e_j\}.\end{aligned}$$

It then follows from Lemma 1 that

$$\begin{aligned}&E\{\hat{q}(r(u_i);\beta)-q(r(u_i))\}^2\\ =\;& E\Big[\sum_{j=1}^n W_{nj}(r(u_i);\beta)\{q(r(u_{j,0}))-q(r(u_i))+e_j\}\Big]^2\\ =\;& E\Big[\sum_{j=1}^n W_{nj}(r(u_i);\beta)\{q(r(u_{j,0}))-q(r(u_j))+q(r(u_j))-q(r(u_i))+e_j\}\Big]^2\\ =\;& E\Big[\{-q(r(u_i))+\sum_{j=1}^n W_{nj}(r(u_i);\beta)q(r(u_j))\}\\ &-\sum_{j=1}^n W_{nj}(r(u_i);\beta)\{q(r(u_j))-q(r(u_{j,0}))\}+\sum_{j=1}^n W_{nj}(r(u_i);\beta)e_j)\Big]^2\\ \le\;& 2E\Big\{q(r(u_i))-\sum_{j=1}^n W_{nj}(r(u_i);\beta)q(r(u_j))\Big\}^2+2E\Big\{\sum_{j=1}^n W_{nj}(r(u_i);\beta)e_j\Big\}^2+O(n^{-1})\\ \le\;& d_1h_2^4+2\sum_{j=1}^n E\Big\{W_{nj}^2(r(u_i);\beta)e_j^2\Big\}\le d_1h_2^4+d_2(nh_2)^{-1},\end{aligned}$$

where d_1, d_2 are some positive constants. The last second inequality holds due to the fact that $\{W_{nj}(r(u_i);\beta)e_j, j = 1, \ldots, n\}$ are independent mean zero random variables given u_i, and the last inequality holds because e_i^2 is bounded. Using Cauchy-Schwarz inequality, Lemma 2 is proved. □

Lemma 3. *Under the assumptions of Lemma 2, we have*

$$E\Big|\int_{r(u)}^{\lambda_r} \frac{q(s) - \hat{q}(s; r, \beta)}{q^2(s)} ds\Big| = O(h_2^2 + (nh_2)^{-1/2}).$$

Proof. Noticing that $\inf_{u\in\Omega} q(r(u)) > 0$ and using Lemma 2, we have

$$\begin{aligned}
E\Big|\int_{r(u)}^{\lambda_r} \frac{q(s) - \hat{q}(s; r, \beta)}{q^2(s)} ds\Big| &\leq \int_{x\in D_X} \int_{r(u)}^{\lambda_r} \Big|\frac{q(s) - \hat{q}(s; r, \beta)}{q^2(s)}\Big| ds f_X(x) dx \\
&= \int_{r(u)}^{\lambda_r} \int_{x\in D_X} \Big|\frac{q(s) - \hat{q}(s; r, \beta)}{q^2(s)}\Big| f_X(x) dx ds \\
&\leq \frac{1}{\inf_s q^2(s)} \int_{r(u)}^{\lambda_r} \int_{x\in D_X} |q(s) - \hat{q}(s; r, \beta)| f_X(x) dx ds \\
&\leq \frac{1}{\inf_s q^2(s)} \int_{r(u)}^{\lambda_r} E|q(s) - \hat{q}(s; r, \beta)| ds \\
&\leq \frac{1}{\inf_s q^2(s)} O(h_2^2 + (nh_2)^{-1/2}),
\end{aligned}$$

where $\inf_s q^2(s)$ is taken over $s = r(u)$ for $u \in \Omega$. □

Proof of Theorem 2. By Theorems 2 and 3 of Lewbel and Linton (2002), if $\sup_{\epsilon\in\Omega_\epsilon}(\epsilon) \leq \sup_u r(u) = \lambda_r$, where the superium is taken over $\{u : u = x^\top \beta_0, x \in D_X\}$, $m(u)$ can be written as

$$m(u) = \lambda_r - \int_{r(u)}^{\lambda_r} \frac{1}{q(s)} ds.$$

In general, without the assumption above, $m(u)$ can be written as

$$m(u) = k_0 + \lambda_r - \int_{r(u)}^{\lambda_r} \frac{1}{q(s)} ds.$$

We can see that the only difference is the constant k_0. In what follows, we use the form without k_0 as Lewbel and Linton (2002).

Note that

$$\begin{aligned} \hat{m}(u) - m(u) &= \hat{\lambda}_r - \lambda_r + \int_{r(u)}^{\lambda_r} \frac{1}{q(s)} ds - \int_{\hat{r}(u)}^{\hat{\lambda}_r} \frac{1}{\hat{q}(s)} ds \\ &= \hat{\lambda}_r - \lambda_r + \int_{r(u)}^{\hat{r}(u)} \frac{1}{q(s)} ds + \int_{\hat{r}(u)}^{\lambda_r} \frac{1}{q(s)} ds - \int_{\hat{r}(u)}^{\hat{\lambda}_r} \frac{1}{\hat{q}(s)} ds. \quad \text{(S3.2)} \end{aligned}$$

By Taylor expansion of $q(\cdot)$ around $r(u)$, we have

$$\begin{aligned} \int_{r(u)}^{\hat{r}(u)} \frac{1}{q(s)} ds &= \int_{r(u)}^{\hat{r}(u)} \Big[\frac{1}{q(r(u))} - \frac{q'(r(u))}{q^2(r(u))} \{s - r(u)\} + O\{s - r(u)\}^2 \Big] ds \\ &= \frac{1}{q(r(u))} \{\hat{r}(u) - r(u)\} - \frac{q'(r(u))}{2q^2(r(u))} \{\hat{r}(u) - r(u)\}^2 + O\{\hat{r}(u) - r(u)\}^3. \text{(S3.3)} \end{aligned}$$

By Taylor expansion of $\hat{q}(\cdot)$ around λ_r, we have

$$\begin{aligned} \int_{\lambda_r}^{\hat{\lambda}_r} \frac{1}{\hat{q}(s)} ds &= \int_{\lambda_r}^{\hat{\lambda}_r} \Big\{ \frac{1}{\hat{q}(\lambda_r)} - \frac{\hat{q}'(\lambda_r)}{\hat{q}^2(\lambda_r)} (s - \lambda_r) + O(s - \lambda_r)^2 \Big\} ds \\ &= \frac{1}{\hat{q}(\lambda_r)} (\hat{\lambda}_r - \lambda_r) - \frac{\hat{q}'(\lambda_r)}{2\hat{q}^2(\lambda_r)} (\hat{\lambda}_r - \lambda_r)^2 + O(\hat{\lambda}_r - \lambda_r)^3 \\ &= \frac{1}{q(\lambda_r)} (\hat{\lambda}_r - \lambda_r) + \frac{q(\lambda_r) - \hat{q}(\lambda_r)}{\hat{q}(\lambda_r) q(\lambda_r)} (\hat{\lambda}_r - \lambda_r) \\ &\quad - \frac{\hat{q}'(\lambda_r)}{2\hat{q}^2(\lambda_r)} (\hat{\lambda}_r - \lambda_r)^2 + O(\hat{\lambda}_r - \lambda_r)^3. \quad \text{(S3.4)} \end{aligned}$$

Combining (S3.2)-(S3.4), and by the mean value theorem, we have

$$\hat{m}(u) - m(u)$$

$$
\begin{aligned}
&= (1-\frac{1}{q(\lambda_r)})(\hat{\lambda}_r-\lambda_r)+\frac{1}{q(r(u))}\{\hat{r}(u)-r(u)\}-\frac{q'(r(u))}{2q^2(r(u))}\{\hat{r}(u)-r(u)\}^2 \\
&\quad +O\{\hat{r}(u)-r(u)\}^3-\frac{\{\hat{q}(\bar{r}(u))-q(\bar{r}(u))\}\{\hat{r}(u)-r(u)\}}{q(\bar{r}(u))\hat{q}(\bar{r}(u))}+\int_{r(u)}^{\lambda_r}\frac{\hat{q}(s)-q(s)}{q^2(s)}ds \\
&\quad +\int_{r(u)}^{\lambda_r}\frac{\{\hat{q}(s)-q(s)\}^2}{q^2(s)\hat{q}(s)}ds-\frac{q(\lambda_r)-\hat{q}(\lambda_r)}{\hat{q}(\lambda_r)q(\lambda_r)}(\hat{\lambda}_r-\lambda_r)+\frac{\hat{q}'(\bar{\lambda})}{2\hat{q}^2(\bar{\lambda})}(\hat{\lambda}_r-\lambda_r)^2 \\
&= T_1+T_2+T_3+T_4+T_5+T_6+T_7+T_8+T_9,
\end{aligned}
$$

where $\bar{r}(u)$ is some value between $\hat{r}(u)$ and $r(u)$, and $\bar{\lambda}$ is some value between $\hat{\lambda}_r$ and λ_r,

We first consider $T_k, k=1,\ldots,4$. Since $q(\lambda_r)=1$ by Assumption A.1, we have $T_1=0$. For T_2, similar to Theorem 1 in Carroll et al. (1997), we have

$$
\sqrt{nh_1}\left\{T_2-\frac{1}{2}k_2r''(u)h_1^2\right\}\xrightarrow{\mathcal{D}}N\left\{0,\sigma_u^2/s_0^2(u)\right\}, \tag{S3.5}
$$

where $s_0(u)=q(r(u))$ and k_2 is some constant. Since T_3 and T_4 are higher orders of $\hat{r}(u)-r(u)$, thus both of them are $o_p(h_1^2+(nh_1)^{-1/2})$.

Now we turn to $T_k, k=5,\ldots,9$. Define $\hat{q}(s)=\hat{q}(s;\hat{r},\beta)$ as the estimator of $q(\cdot)$ evaluated at s, $\hat{q}(s;r,\beta)$ as the estimator given $r(\cdot)$, and $\hat{q}(s;r,\beta_0)$ as the estimator given $r(\cdot)$ and β_0. We decompose $\hat{q}(s;\hat{r},\beta)-q(s)$ as

$$
\hat{q}(s;\hat{r},\beta)-q(s)=\hat{q}(s;\hat{r},\beta)-\hat{q}(s;r,\beta)+\hat{q}(s;r,\beta)-q(s).
$$

By Markov inequality, Lemma 3 indicates that $\int_{r(u)}^{\lambda_r}\frac{q(s)-\hat{q}(s;r,\beta)}{q^2(s)}ds=O_p((nh_2)^{-1/2}+$

h_2^2), and thus

$$\begin{aligned} T_6 &= \int_{r(u)}^{\lambda_r} \frac{q(s) - \hat{q}(s; r, \beta)}{q^2(s)} ds + \int_{r(u)}^{\lambda_r} \frac{\hat{q}(s; r, \beta) - \hat{q}(s; \hat{r}, \beta)}{q^2(s)} ds \\ &= O_p(h_2^2 + (nh_2)^{-1/2}) + O_p(h_2^2) = O_p(h_2^2 + (nh_2)^{-1/2}), \quad \text{(S3.6)} \end{aligned}$$

where the order of $\int_{r(u)}^{\lambda_r} \frac{\hat{q}(s;r,\beta) - \hat{q}(s;\hat{r},\beta)}{q^2(s)} ds$ follows from Lemmas 2 and 3 and Theorem 5 in Lewbel and Linton (2002). By Lemma 2 of Lewbel and Linton (2002), $Var(T_6) = O_p(1/n)$.

By Lemma 1 of Ichimura (1993), we have that $\sup_u |q(r(u)) - \hat{q}(r(u); r, \beta)| = o_p(1)$. By (26) and Theorem 5 of Lewbel and Linton (2002), we have $\sup_u |\hat{q}(r(u); r, \beta) - \hat{q}(r(u); \hat{r}, \beta)| = o_p(1)$. Thus, $\sup_u |q(r(u)) - \hat{q}(\hat{r}(u); \hat{r}, \beta)| = o_p(1)$, where the supreme is taken over $u = x^\top \beta, x \in D_X$ and $\beta \in \Theta_{c_0}$. Then,

$$\begin{aligned} T_5 &= \frac{\{\hat{q}(\bar{r}(u)) - q(\bar{r}(u))\}\{\hat{r}(u) - r(u)\}}{q(\bar{r}(u))\hat{q}(\bar{r}(u))} = \frac{o_p(1)O_p(h_1^2 + (nh_1)^{-1/2})}{q^2(\bar{r}(u))(1 + o_p(1))} \\ &= o_p(h_1^2 + (nh_1)^{-1/2}), \\ T_7 &= \int_{r(u)}^{\lambda_r} \frac{\{\hat{q}(s) - q(s)\}^2}{q^2(s)\hat{q}(s)} ds \le \int_{r(u)}^{\lambda_r} \frac{\{\hat{q}(s) - q(s)\}^2}{q^3(s)(\hat{q}(s)/q(s))} ds \\ &\le \frac{1}{1 + o_p(1)} \int_{r(u)}^{\lambda_r} \frac{\{\hat{q}(s) - q(s)\}^2}{q^3(s)} ds \\ &\le \frac{1}{1 + o_p(1)} o_p(h_2^2 + (nh_2)^{-1/2}) = o_p(h_2^2 + (nh_2)^{-1/2}), \\ T_8 &= \frac{q(\lambda_r) - \hat{q}(\lambda_r)}{\hat{q}(\lambda_r)q(\lambda_r)}(\hat{\lambda}_r - \lambda_r) \frac{o_p(1)O_p(h_1^2 + (nh_1)^{-1/2})}{q^2(\lambda_r)(1 + o_p(1))} = o_p(h_1^2 + (nh_1)^{-1/2}), \\ T_9 &= \frac{\hat{q}'(\bar{\lambda})}{2\hat{q}^2(\bar{\lambda})}(\hat{\lambda}_r - \lambda_r)^2 = o_p(h_1^2 + (nh_1)^{-1/2}). \end{aligned}$$

In summary, $T_1+T_3+T_4+T_5+T_7+T_8+T_9 = o_p(h_1^2+(nh_1)^{-1/2})$ under Assumption A.3 (ii), and together with (S3.5)-(S3.6), the proof of Theorem 2 is completed, and the bounded function $b_m(\cdot)$ is determined by T_2 and T_6.

S4. Proof of Theorem 3

We first present two lemmas for proving Theorem 3.

Lemma 4. *Under Assumptions A.1-A.4, suppose that* $\beta \in \Theta_{c_0}$*, then* $K(\frac{X_i^\top \beta_0 - X_j^\top \beta_0}{h})/h - K(\frac{X_i^\top \beta - X_j^\top \beta}{h})/h = O(n^{-1/2})$.

Proof. By the mean value theorem, we have

$$
\begin{aligned}
& \frac{1}{h}K\Big(\frac{X_i^\top \beta_0 - X_j^\top \beta_0}{h}\Big) - \frac{1}{h}K\Big(\frac{X_i^\top \beta - X_j^\top \beta}{h}\Big) \\
= & \frac{1}{h}K^{'}\Big(\frac{X_i^\top \beta^{**} - X_j^\top \beta^{**}}{h}\Big)\Big(\frac{X_i^\top \beta_0 - X_j^\top \beta_0}{h} - \frac{X_i^\top \beta - X_j^\top \beta}{h}\Big),
\end{aligned}
$$

where β^{**} is some value between β_0 and β. By Assumption A.4, there exists a constant L, such that $K^{'}(s) \le C_4|s|^{-2}$ for some constant C_4 when $s > L$. Then we bound the difference in two cases.

Case 1. When $|\frac{X_i^\top \beta^{**} - X_j^\top \beta^{**}}{h}| > L$,

$$
\begin{aligned}
& \Big|\frac{1}{h}K\Big(\frac{X_i^\top \beta_0 - X_j^\top \beta_0}{h}\Big) - \frac{1}{h}K\Big(\frac{X_i^\top \beta - X_j^\top \beta}{h}\Big)\Big| \\
= & \frac{1}{h}K^{'}\Big(\frac{X_i^\top \beta^{**} - X_j^\top \beta^{**}}{h}\Big)\Big|\frac{X_i^\top \beta_0 - X_j^\top \beta_0}{h} - \frac{X_i^\top \beta - X_j^\top \beta}{h}\Big| \\
\le & \frac{1}{h^2}|C_4|\Big|\frac{X_i^\top \beta^{**} - X_j^\top \beta^{**}}{h}\Big|^{-2} O(n^{-1/2}) \\
\le & O(n^{-1/2}).
\end{aligned}
$$

Case 2. When $|\frac{X_i^\top\beta^{**}-X_j^\top\beta^{**}}{h}| < L < \infty$, since the support of X is bounded, it implies that $\frac{1}{h} < C < \infty$ or $|X_i^\top\beta^{**} - X_j^\top\beta^{**}| = 0$ (this implies $K^{'}(0) = 0$). Thus,

$$\begin{aligned}
& \Big|\frac{1}{h}K\Big(\frac{X_i^\top\beta_0 - X_j^\top\beta_0}{h}\Big) - \frac{1}{h}K\Big(\frac{X_i^\top\beta - X_j^\top\beta}{h}\Big)\Big| \\
= \ & \frac{1}{h}\Big|K^{'}\Big(\frac{X_i^\top\beta^{**} - X_j^\top\beta^{**}}{h}\Big)\Big|\Big|\frac{X_i^\top\beta_0 - X_j^\top\beta_0}{h} - \frac{X_i^\top\beta - X_j^\top\beta}{h}\Big| \\
= \ & O(n^{-1/2}).
\end{aligned}$$

Result follows from Cases 1 and 2. □

Lemma 5. (Theorem 1; Hall, 1984) *Let $Z_i, i = 1, 2, \ldots, , n$ be i.i.d random vectors, for each n, and let $U_n = \sum_{1\le i\le j\le n} H_n(Z_i, Z_j)$, $M_n(x, y) = E\{H_n(Z_1, x)H_n(Z_1, y)\}$, where H_n is a sequence of measurable functions symmetric under permutation, with $E\{H_n(Z_1, Z_2|Z_1)\} = 0$ a.s. and $E(H_n^2(Z_1, Z_2)) < \infty$, for each $n \ge 1$. If $\{EM_n^2(Z_1, Z_2) + n^{-1}H_n^4(Z_1, Z_2)\}/EH_n^2(Z_1, Z_2) \to 0$, then U_n/n is asymptotically normal distributed with mean zero and variance $EH_n^2(Z_1, Z_2)$.*

Proof of Theorem 3. We first investigate the asymptotic property of V_n.

Recall the definition $r_0(u) = \int_{-\infty}^{\zeta_0+\zeta_1 u} \Phi(\epsilon/\sigma)d\epsilon$ in Section 2.4 of the main part. To reflect the dependence of r on σ, we rewrite the definition as $r_0(u, \sigma) = \int_{-\infty}^{\zeta_0+\zeta_1 u} \Phi(\epsilon/\sigma)d\epsilon$. Let $\epsilon_i^{'} = Y_i - r_0(X_i^\top\beta_0, \sigma)$. Then

$$\hat{\epsilon}_i^{'} = \epsilon_i^{'} - \{r_0(X_i^\top\beta, \hat{\sigma}) - r_0(X_i^\top\beta_0, \sigma)\}. \tag{S4.7}$$

Thus, we can write V_n as the sum of the following three terms.

$$
\begin{aligned}
V_{1n} &= \frac{1}{n(n-1)h}\sum_{i\neq j} K(\frac{X_i^\top\beta - X_j^\top\beta}{h})\epsilon_i^{'}\epsilon_j^{'}. \\
V_{2n} &= \frac{1}{n(n-1)h}\sum_{i\neq j} K(\frac{X_i^\top\beta - X_j^\top\beta}{h})\epsilon_i^{'}\{r_0(X_j^\top\beta,\hat\sigma) - r_0(X_j^\top\beta_0,\sigma)\}. \\
V_{3n} &= \frac{1}{n(n-1)h}\sum_{i\neq j} K(\frac{X_i^\top\beta - X_j^\top\beta}{h})\{r_0(X_i^\top\beta,\hat\sigma) - r_0(X_j^\top\beta_0,\sigma)\} \\
&\quad \{r_0(X_i^\top\beta,\hat\sigma) - r_0(X_j^\top\beta_0,\sigma)\}.
\end{aligned}
$$

We establish the asymptotic property of V_{1n}, V_{2n} and V_{3n} separately. For convenience, we define terms V_{2n}^* and V_{3n}^* by replacing β by β_0 in V_{2n} and V_{3n}.

Define $Z_i = (X_i^\top\beta, \epsilon_i')$, and let $H_n(Z_i, Z_j) = \frac{1}{h}K\left(\frac{X_i^\top\beta - X_j^\top\beta}{h}\right)\epsilon_i^{'}\epsilon_j^{'}$. It can be seen that $E\{H_n(Z_1, Z_2|Z_1)\} = 0$ (first take expectation conditional on Z_2 and then on ϵ') and $E\{H_n^2(Z_1, Z_2|Z_1)\} < \infty$ for each n, as $\epsilon_i^{'}$ has finite second moment. Define $M_n(x, y) = E\{H_n(Z_1, x)H_n(Z_1, y)\}$. Letting $U = X^\top\beta$, and following the results of Theorem 3.1 in Koul et al. (2014), we have

$$
\begin{aligned}
&E\{M_n^2(Z_1, Z_2)\} \\
= \; & E[E\{H_n(Z_3, Z_1)H_n(Z_3, Z_2)\}|(Z_1, Z_2)]^2 \\
= \; & E\Big[E\Big\{\frac{1}{h^2}K\Big(\frac{X_3^\top\beta - X_1^\top\beta}{h}\Big)K\Big(\frac{X_3^\top\beta - X_2^\top\beta}{h}\Big)\epsilon_1'\epsilon_2'\epsilon_3'^2|Z_1, Z_2\Big\}\Big]^2 \\
= \; & \frac{1}{h^4}E\Big\{\epsilon_1'\epsilon_2'\int K\Big(\frac{X_3^\top\beta - X_1^\top\beta}{h}\Big)K\Big(\frac{X_3^\top\beta - X_2^\top\beta}{h}\Big)\tau^2(X^\top\beta)f_U(X_3^\top\beta)dU\Big\}^2
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{h^4} E\Big[\epsilon_1^{\prime 2}\epsilon_2^{\prime 2}\Big\{\int K\Big(\frac{X_3^\top\beta - X_1^\top\beta}{h}\Big)K\Big(\frac{X_3^\top\beta - X_2^\top\beta}{h}\Big)\tau^2(X^\top\beta)f_U(X_3^\top\beta)dU\Big\}^2\Big] \\
&\le \frac{1}{h^4} E\Big[\epsilon_1^{\prime 2}\epsilon_2^{\prime 2}\Big\{\int K\Big(\frac{X_3^\top\beta_0 - X_1^\top\beta_0}{h}\Big)K\Big(\frac{X_3^\top\beta_0 - X_2^\top\beta_0}{h}\Big)\tau^2(X^\top\beta_0)f_U(X_3^\top\beta_0)dU\Big\}^2\Big] \\
&\quad + O(\frac{1}{\sqrt{n}h}) \\
&= \frac{c}{h^4} E\Big[\epsilon_1^{\prime 2}\epsilon_2^{\prime 2}\Big\{\int K\Big(\frac{X_3^\top\beta_0 - X_1^\top\beta_0}{h}\Big)K\Big(\frac{X_3^\top\beta_0 - X_2^\top\beta_0}{h}\Big)\tau^2(X^\top\beta_0)f_U(X_3^\top\beta_0)dU\Big\}^2\Big] \\
&\quad + O(\frac{1}{\sqrt{n}h}) \\
&= O(1/h) + O(\frac{1}{\sqrt{n}h}),
\end{aligned}
$$

$$
\begin{aligned}
&E\{H_n^2(Z_1, Z_2)\} \\
&= E_\beta\Big\{\frac{1}{h}K\Big(\frac{X_1^\top\beta - X_2^\top\beta}{h}\Big)\epsilon_1^{'}\epsilon_2^{'}\Big\}^2 = E_\beta\Big\{\frac{1}{h^2}K^2\Big(\frac{X_1^\top\beta - X_2^\top\beta}{h}\Big)\epsilon_1^{2'}\epsilon_2^{2'}\Big\} \\
&= E_{\beta_0}\Big\{\frac{1}{h^2}K^2\Big(\frac{X_1^\top\beta_0 - X_2^\top\beta_0}{h}\Big)\epsilon_1^{2'}\epsilon_2^{2'}\Big\} + O(\frac{1}{\sqrt{n}h}) = O(1/h) + O(\frac{1}{\sqrt{n}h}).
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
E\{H_n^4(Z_1, Z_2)\} &= E_\beta\Big\{\frac{1}{h}K\Big(\frac{X_1^\top\beta - X_2^\top\beta}{h}\Big)\epsilon_1^{'}\epsilon_2^{'}\Big\}^4 \\
&= E_{\beta_0}\Big\{\frac{1}{h^4}K^4\Big(\frac{X_1^\top\beta_0 - X_2^\top\beta_0}{h}\Big)\epsilon_1^{4'}\epsilon_2^{4'}\Big\} + O(\frac{1}{\sqrt{n}h^3}) \\
&= O(\frac{1}{h^3}) + O(\frac{1}{\sqrt{n}h^3}).
\end{aligned}
$$

Thus,

$$
\frac{EM_n^2(Z_1, Z_2) + n^{-1}H_n^4(Z_1, Z_2)}{E\{H_n^2(Z_1, Z_2)\}} = \frac{O(1/h) + O(1/nh^3)}{O(1/h^2)} \to 0.
$$

These results indicate that $nh^{1/2}V_{1n} \to N(0, \gamma^2)$. Furthermore,

$$
|V_{2n}| \le \frac{1}{n(n-1)h}\sum_{i\neq j} K\Big(\frac{X_i^\top\beta - X_j^\top\beta}{h}\Big)|\epsilon_i^{'}||r_0(X_j^\top\beta, \hat\sigma) - r_0(X_j^\top\beta_0, \sigma)|
$$

$$
\begin{aligned}
&= \frac{1}{n(n-1)h}\sum_{i\neq j}\Big\{K\Big(\frac{X_i^\top\beta_0 - X_j^\top\beta_0}{h}\Big) + K\Big(\frac{X_i^\top\beta - X_j^\top\beta}{h}\Big) \\
&\quad -K\Big(\frac{X_i^\top\beta_0 - X_j^\top\beta_0}{h}\Big)\Big\}|\epsilon_i^{'}||r_0(X_j^\top\beta_0,\hat\sigma) - r_0(X_j^\top\beta_0,\sigma)|\{1+O_p(n^{-1/2})\} \\
&\leq \frac{1}{n(n-1)h}\sum_{i\neq j}K\Big(\frac{X_i^\top\beta_0 - X_j^\top\beta_0}{h}\Big)|r_0(X_j^\top\beta_0,\hat\sigma) - r_0(X_j^\top\beta_0,\sigma)||\epsilon_i^{'}|\{1+O_p(n^{-1/2})\} \\
&\quad +\frac{1}{n(n-1)h}\sum_{i\neq j}\Big|K\Big(\frac{X_i^\top\beta - X_j^\top\beta}{h}\Big) - K\Big(\frac{X_i^\top\beta_0 - X_j^\top\beta_0}{h}\Big)\Big||\epsilon_i^{'}| \\
&\quad \times|r_0(X_j^\top\beta_0,\hat\sigma) - r_0(X_j^\top\beta_0,\sigma)|\{1+O_p(n^{-1/2})\} \\
&= V_{2n}^*\{1+O_p(n^{-1/2})\} + \frac{1}{n(n-1)}\sum_{i\neq j}O_p(n^{-1/2})o_p(n^{-1/2})|\epsilon_i^{'}|\{1+O_p(n^{-1/2})\} \\
&= O_p(1/n),
\end{aligned}
$$

$$
\begin{aligned}
|V_{3n}| &\leq \frac{1}{n(n-1)h}\sum_{i\neq j}K\Big(\frac{X_i^\top\beta - X_j^\top\beta}{h}\Big)|r_0(X_j^\top\beta_0,\hat\sigma) - r_0(X_j^\top\beta_0,\sigma)| \\
&\quad \times|r_0(X_j^\top\beta,\hat\sigma) - r_0(X_j^\top\beta_0,\sigma)| \\
&= \frac{1}{n(n-1)h}\sum_{i\neq j}\Big\{K\Big(\frac{X_i^\top\beta_0 - X_j^\top\beta_0}{h}\Big) + K\Big(\frac{X_i^\top\beta - X_j^\top\beta}{h}\Big) - K\Big(\frac{X_i^\top\beta_0 - X_j^\top\beta_0}{h}\Big)\Big\} \\
&\quad \times|r_0(X_j^\top\beta_0,\hat\sigma) - r_0(X_j^\top\beta_0,\sigma)||r_0(X_j^\top\beta_0,\hat\sigma) - r_0(X_j^\top\beta_0,\sigma)|\{1+O_p(n^{-1/2})\}^2 \\
&\leq \frac{1}{n(n-1)h}\sum_{i\neq j}K\Big(\frac{X_i^\top\beta_0 - X_j^\top\beta_0}{h}\Big)|r_0(X_j^\top\beta_0,\hat\sigma) \\
&\quad -r_0(X_j^\top\beta_0,\sigma)||r_0(X_i^\top\beta_0,\hat\sigma) - r_0(X_i^\top\beta_0,\sigma)|\{1+O_p(n^{-1/2})\}^2 \\
&\quad +\frac{1}{n(n-1)h}\sum_{i\neq j}\Big|K\Big(\frac{X_i^\top\beta - X_j^\top\beta}{h}\Big) - K\Big(\frac{X_i^\top\beta_0 - X_j^\top\beta_0}{h}\Big)\Big| \\
&\quad \times|r_0(X_j^\top\beta_0,\hat\sigma) - r_0(X_j^\top\beta_0,\sigma)|\{1+O_p(n^{-1/2})\}^2 \\
&= V_{3n}^*\{1+O_p(n^{-1/2})\}^2 + \frac{1}{n(n-1)}\sum_{i\neq j}o_p(n^{-1})O_p(n^{-1/2})\{1+O_p(n^{-1/2})\}^2
\end{aligned}
$$

$$= O_p(1/n).$$

These calculations imply that $nh^{1/2}V_{2n} = o_p(1)$ and $nh^{1/2}V_{3n} = o_p(1)$. It then follows that

$$nh^{1/2}V_n \to N(0, \gamma^2). \tag{S4.8}$$

Now we discuss the asymptotic property of our estimator $\hat{\gamma}^2$. It follows similarly to the proof to derive the asymptotic property of V_n that

$$\begin{aligned}
\hat{\gamma}^2 &= \frac{2}{n(n-1)h}\sum_{i\neq j}\Big\{K\Big(\frac{X_i^\top\beta_0 - X_j^\top\beta_0}{h}\Big) + K\Big(\frac{X_i^\top\beta - X_j^\top\beta}{h}\Big) - K\Big(\frac{X_i^\top\beta_0 - X_j^\top\beta_0}{h}\Big)\Big\}^2 \\
&\quad \times \hat{\epsilon}_i^{\prime 2*}\hat{\epsilon}_j^{\prime 2*} \\
&= \frac{2}{n(n-1)h}\sum_{i\neq j} K^2\Big(\frac{X_i^\top\beta_0 - X_j^\top\beta_0}{h}\Big)\epsilon_i^{\prime 2}\epsilon_j^{\prime 2}\{1 + O_p(n^{-1/2})\}^2 \\
&+ \frac{2}{n(n-1)h}\sum_{i\neq j}\Big\{K\Big(\frac{X_i^\top\beta - X_j^\top\beta}{h}\Big) - K\Big(\frac{X_i^\top\beta_0 - X_j^\top\beta_0}{h}\Big)\Big\}K\Big(\frac{X_i^\top\beta_0 - X_j^\top\beta_0}{h}\Big) \\
&\quad \times \epsilon_i^{\prime 2}\epsilon_j^{\prime 2}\{1 + O_p(n^{-1/2})\}^2 \\
&+ \frac{2}{n(n-1)h}\sum_{i\neq j}\Big\{K\Big(\frac{X_i^\top\beta - X_j^\top\beta}{h}\Big) - K\Big(\frac{X_i^\top\beta_0 - X_j^\top\beta_0}{h}\Big)\Big\}^2 \\
&\quad \times \epsilon_i^{\prime 2}\epsilon_j^{\prime 2}\{1 + O_p(n^{-1/2})\}^2 \\
&= \gamma^2\{1 + O_p(n^{-1/2})\}^2 + o_p(n^{-1/2}).
\end{aligned}$$

Thus, $\hat{\gamma}^2 = \gamma^2 + O_p(n^{-1/2})$. We have a consistent estimator of γ^2. Theorem 3 then follows from (S4.8).

Bibliography

Carroll, R. J., Fan, J., Gijbels, I., and Wand, M. P. (1997). Generalized partially linear single-index models. *Journal of the American Statistical Association*, 92(438):477–489.

Hall, P. (1984). Central limit theorem for integrated square error of multivariate nonparametric density estimators. *Journal of Multivariate Analysis*, 14(1):1–16.

Ichimura, H. (1993). Semiparametric least squares (SLS) and weighted SLS estimation of single-index models. *Journal of Econometrics*, 58:71–120.

Koul, H. L., Song, W., and Liu, S. (2014). Model checking in Tobit regression via nonparametric smoothing. *Journal of Multivariate Analysis*, 125:36–49.

Lewbel, A. and Linton, O. (2002). Nonparametric censored and truncated regression. *Econometrica*, 70(2):765–779.

Liang, H., Liu, X., Li, R., and Tsai, C. L. (2010). Estimation and testing for partially linear single-index models. *The Annals of Statistics*, 38:3811–3836.

Wang, J.-L., Xue, L., Zhu, L., and Chong, Y. S. (2010). Estimation for a partial-linear single-index model. *The Annals of Statistics*, 38:246–274.

Zhu, L. and Xue, L. (2006). Empirical likelihood confidence regions in a partially linear single-index model. *Journal of the Royal Statistical Society, Series B*, 68:549–570.