Supplement to "Extremal Quantile Regressions for Selection Models and the Black-White Wage Gap"

Xavier D'Haultfœuille* Arnaud Maurel† Yichong Zhang‡

July 2017

Abstract

This paper gathers the supplementary material to D'Haultfoeuille *et al.* (2017). First, we provide in Section 1 a set of sufficient conditions for the independence at infinity assumption (Assumption 3) and the admissible rates of convergence of the quantile index (Assumption 6). We consider in Section 2 two situations not covered by the assumptions in D'Haultfoeuille *et al.* (2017), namely measurement errors on the outcome and exogenous selection, and provide conditions for consistency in such cases. Section 3 outlines the theoretical arguments underlying the construction of our data-driven method for choosing the quantile index τ_n. The proofs of the results in this supplement are collected in Section 4. Finally, Section 5 gathers technical lemmas used in the proofs of the main results in the paper.

Notation

We use the same notation as in D'Haultfoeuille *et al.* (2017). We also use some notation related to extreme value theory. A function F is regularly varying at $x \in \{0, +\infty\}$ with index $\alpha \in [-\infty, +\infty]$, and we write $F \in RV_\alpha(x)$, if for any $t > 0$, $\lim_{u \to x} F(tu)/F(u) = t^\alpha$, with the understanding that $t^\infty = \infty$ if $t > 1$ and $= 0$ if $1 > t > 0$ (and similarly for $\alpha = -\infty$). F is slowly varying (resp. rapidly varying) at x if $F \in RV_0(x)$ (resp. $F \in RV_{+\infty}(x)$).

*CREST. E-mail address: xavier.dhaultfoeuille@ensae.fr.

†Duke University, NBER and IZA. E-mail address: apm16@duke.edu.

‡Singapore Management University. E-mail address: yczhang@smu.edu.sg.

1 Identification and convergence rates: sufficient conditions

Throughout this section, we assume the location-scale model:

$$Y^* = \psi(X) + \sigma(X)\eta,$$

where $\eta \perp\!\!\!\perp X$. Under this model specification, Assumption 1 implies that $\sigma(\cdot)$ is a function of X_{-1}. Assumption 5(ii) holds automatically. In the following, we maintain Assumptions 2, 4, 5(i), and 5(iii). We then verify the key assumption concerning the identification, i.e. Assumption 3, under primitive conditions. We also characterize sequences τ_n satisfying Assumption 6 under such conditions, thus allowing us to derive rates of convergence of $\widehat{\beta}_1$.

1.1 Independence at infinity

To get a better sense of Assumption 3, we discuss it below in the context of a threshold crossing selection model described below.[1]

Assumption 7. *(i)* $D = \mathbb{1}\{\phi(X) - \nu \geq 0\}$ *with* $(\eta, \nu) \perp\!\!\!\perp X$*, (ii)* $\inf_{x\in Supp(X)} F_\nu(\phi(x)) = \underline{v} > 0$*, (iii)* F_η *and* F_ν *are continuous and strictly increasing and the copula* C *of* $(-\eta, \nu)$ *is differentiable with respect to its first argument.*

The first condition defines the selection model as a standard threshold crossing model. Importantly however, we do not add any instrument in this selection equation. The second condition ensures that $x \mapsto P(D = 1|X = x)$ is bounded below by a positive number. Note that this condition will typically hold if none of the covariates has a large support, which is precisely the type of situation we are interested in. In this context, Proposition 1.1 provides a restriction on C ensuring that Assumption 3 is satisfied. Hereafter, let $f_C(\tau) = \sup_{u\leq\tau, v\in[\underline{v},1]} |\partial_1 C(u, v) - 1|$.

Proposition 1.1. *Suppose that Assumptions 2 and 7 hold, and*

$$\lim_{\tau\to 0} f_C(\tau) = 0. \tag{1.1}$$

Then Assumption 3 is satisfied.

The key idea is that selection becomes independent of the covariates for large values of the outcome if selection is endogenous enough, in the sense that $(-\eta, \nu)$ satisfies (1.1). To understand this condition better, it is useful to consider two extreme cases. In the perfect dependence case such that $\nu = -\eta$, then $\partial_1 C(u, v) = 1$ for all $u < v$, so that (1.1) actually

[1] Our identification strategy is also natural in the context of the generalized Roy model (Heckman & Vytlacil, 2007). We refer the reader to D'Haultfoeuille *et al.* (2014) for a detailed discussion on this question.

holds exactly for small values of τ. On the other hand, when ν and $-\eta$ are independent, $\partial_1 C(u,v) = v$, and $f_C(\tau) = 1 - \underline{v}$, which is positive except in the degenerate case where $D = 1$ almost surely. In between these two extreme cases, Table 1 provides examples of copulas that satisfy this constraint. It underlines that Assumption 3 may be satisfied even if the dependence between $-\eta$ and ν is very weak. Importantly, it holds for all Gaussian copulas with positive dependence. It also holds for Archimedean copulas under a restriction on the behavior of the generator Ψ around 0. This restriction holds for instance for the Clayton copula, for which $\Psi(u) = (u^{-\theta} - 1)/\theta$, provided that $\theta > 0$. The Gumbel family is another popular Archimedean family of copulas that satisfies the required restriction on Ψ, given a restriction on the parameter θ.

Copula family	Restriction ensuring (1.1)
Gaussian $C(u,v;\rho)$	$\rho > 0$
Archimedean $C(u,v;\Psi) = \Psi^{-1}(\Psi(u) + \Psi(v))$	$\lim_{u\to 0} \Psi(u) = +\infty$,
	Ψ is C^1 and $RV_\alpha(0)$ with $\alpha \in (0, +\infty]$
Gumbel $\Psi(u;\theta) = (-\log(u))^\theta$	$\theta > 1$

Table 1: Examples of copulas satisfying (1.1).

1.2 Convergence rate

The next theorem concerns about the convergence rate of $\widehat{\beta}_1$.

Proposition 1.2. *Suppose that Assumptions 1, 2, 4, 5, and 7 hold. Then there exists τ_n satisfying Assumption 6 such that the rates of convergence of $\widehat{\beta}_1$ is polynomial if for some $a > 0$, $f_C(\tau) = o(\tau^a)$ as $\tau \to 0$.*

We show that if, for some $a > 0$,

$$f_C(\tau) = o(\tau^a), \tag{1.2}$$

then a polynomial rate of convergence, faster than $n^{(a-\alpha)/(2a+1)}$ for any $\alpha \in (0, a)$, is possible. Table 2 below provides examples of copulas of $(\nu, -\eta)$ satisfying the latter condition (see Subsection 4.3 for its verification in each case). It is worth noting that for the last two copulas considered in the table, we actually establish that $f_C(\tau)$ tends to zero exponentially fast in τ. In such situations, (1.2) holds for all a, and it is possible to achieve a rate of convergence for $\widehat{\beta}_1$ that is faster than $n^{1/2-\alpha}$ for any $\alpha \in (0, 1/2)$. In other words, an adequate choice of τ_n can make the rate of convergence arbitrarily close to the standard parametric root-n rate. In all cases, the general idea is that if the tail dependence between η and ν is strong, which can be interpreted as a strong form of endogenous selection, then $\mathcal{B}(.)$ is small. It follows that a large τ_n is admissible, resulting in fast convergence rates.

Copula family	Restriction ensuring (1.2)
Gaussian $C(u,v;\rho)$	$\rho > 0$
Clayton $C(u,v;\theta) = \max\left([u^{-\theta} + v^{-\theta} - 1]^{-1/\theta}, 0\right)$	$\theta > 0$
Rotated Gumbel-Barnett $C(u,v;\theta) = u - u(1-v)\exp(-\theta\log(u)\log(1-v))$	$\theta \in (0,1]$
$C(u,v;\theta) = (1 + [(u^{-1}-1)^{\theta} + (v^{-1}-1)^{\theta}]^{1/\theta})^{-1}$	$\theta > 1$
$C(u,v;\theta) = (1 + [(u^{-1/\theta}-1)^{\theta} + (v^{-1/\theta}-1)^{\theta}]^{1/\theta})^{-\theta}$	$\theta \geq 1$
$C(u,v;\theta) = \theta/\log(\exp(\theta/u) + \exp(\theta/v) - \exp(\theta))$	$\theta > 0$
$C(u,v;\theta) = [\log(\exp(u^{-\theta}) + \exp(v^{-\theta}) - e)]^{-1/\theta}$	$\theta > 0$

Table 2: Examples of copulas leading to a polynomial rate of convergence.

2 Extensions

We address in this section two potentially important concerns, namely measurement errors on the outcome variable and exogenous selection.

2.1 Measurement errors on Y^*

Consider the pure location model:

$$Y^* = X_1'\beta_1 + X_{-1}'\beta_{-1} + \varepsilon,\ \varepsilon \perp\!\!\!\perp X.$$

Suppose that Y^* is measured with errors by $\widetilde{Y} = Y^* + \eta$, with η independent of (D, X, ε). Fix $\nu \in (0,1)$ and let M and y be such that $P(\eta \leq M) \geq 1 - \nu$ and $P(D = 1|X, Y^* = y') \in [h(1-\nu), h(1+\nu)]$ for all $y' \geq y$. Then

$$\begin{aligned} P(D = 1|X, \widetilde{Y} = y + M) &= \int P(D = 1|X, Y^* = y + M - u) f_\eta(u) du \\ &\in [h(1-\nu)^2,\ h(1 + (1 + 1/h)\nu)]. \end{aligned}$$

Because ν was arbitrary, this shows that $\lim_{y\to\infty} P(D = 1|X, \widetilde{Y} = y) = h$. Hence, independence at the limit still holds with $\widetilde{Y}$ instead of Y^*. Second,

$$Q_{\widetilde{Y}|X}(\tau|X) = X_1'\beta_1 + X_{-1}'\beta_{-1} + Q_{\varepsilon+\eta}(\tau).$$

Thus, the model is the same as the one on Y^*, except on the intercept. Hence, at least in a pure location model, our framework is insensitive to classical measurement errors on Y^*.[2]

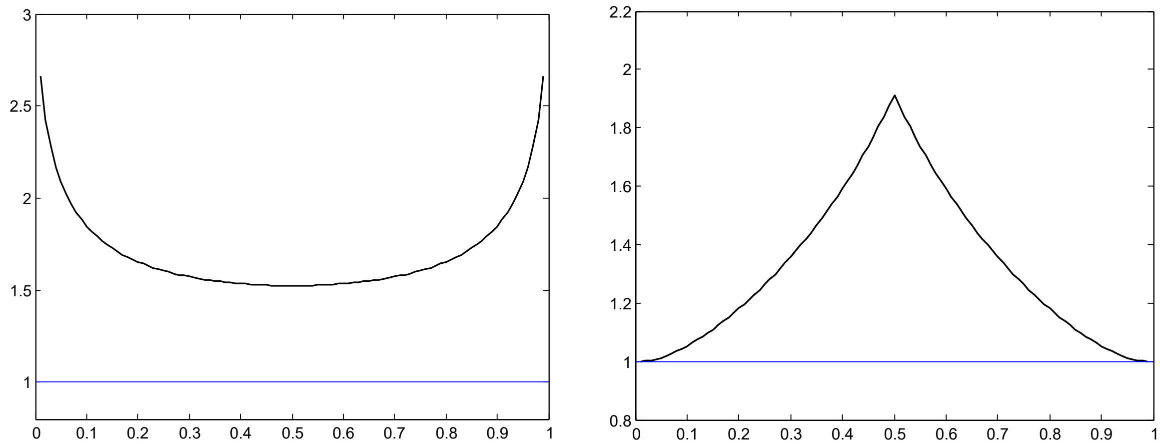
It is also instructive to investigate the effect of measurement errors on the asymptotic variance. Whether the variance of our estimator is more affected than the variance of a traditional quantile regression (QR) estimator really depends on the distributions of ε and η. To see this, consider a simple set-up where $D = 1$ and $Y^* = X_1'\beta_1 + \varepsilon$. We abstract from the selection issue in this example in order to focus on the comparison of the impact of measurement errors on the variance of extremal QR vs. usual QR estimators. Let us denote by $\widehat{\beta}_1^1(\tau)$ the estimator obtained with a QR of order τ of Y^* on X_1, i.e. when the outcome variable is measured without error. The second one, denoted by $\widehat{\beta}_1^2(\tau)$, is obtained with a QR of order τ of $\widetilde{Y}$ on X_1. Finally, we define $R(\tau)$ as the ratio of the asymptotic standard error of $\widehat{\beta}_1^2(\tau)$ to that of $\widehat{\beta}_1^1(\tau)$. It is easy to show that

$$R(\tau) = \frac{f_\varepsilon(Q_\varepsilon(\tau))}{f_{\varepsilon+\eta}(Q_{\varepsilon+\eta}(\tau))}.$$

This result holds for a fixed τ, but a similar statement holds for any intermediate order sequence $\tau_n \to 0$, with the understanding that in this case, $R(\tau_n)$ is the ratio of the normalizing factor of $\widehat{\beta}_1^2(\tau_n)$ to that of $\widehat{\beta}_1^1(\tau_n)$. Specifically, $R(\tau_n) = N_n(\widehat{\beta}_1^2(\tau_n))/N_n(\widehat{\beta}_1^1(\tau_n))$, with $N_n(\widehat{\beta}_1^k(\tau_n))$ such that $(\widehat{\beta}_1^k(\tau_n) - \beta_1)/N_n(\widehat{\beta}_1^k(\tau_n)) \xrightarrow{d} \mathcal{N}(0,1)$ for $k = 1, 2$.

Whether $R(\tau)$ increases as we move away from the middle of the distribution, i.e. as $|\tau - 1/2|$ increases, depends on the distributions of ε and η. For instance, $\tau \mapsto R(\tau)$ is constant when both ε and η are drawn from a standard normal distribution. Next, the left panel in Figure 1 below shows the case where ε follows a standard normal distribution, and η follows a standard Laplace distribution of density $f_\eta(u) = \exp(-|u|)/2$. In this specific case, the ratio $R(\tau)$ increases at the tails. However, the right panel of Figure 1 shows that the situation is reversed when ε follows a standard Laplace distribution, while η follows a standard normal distribution. In this latter case, the effect of the measurement error is large in the middle of the distribution, but becomes negligible in the tails (i.e. for $|1/2 - \tau| \to 1/2$).

[2] In our more general model where $Q_{\varepsilon|X}(\tau|X)$ is linear in X_{-1}, we have, under the same assumptions on the measurement error, $Q_{\widetilde{Y}|X}(\tau|X) = X_1'\beta_1 + q_\tau(X_{-1})$ for some unknown function $q_\tau(.)$. This is a partially linear model. Though probably technically involved, we conjecture that a sieve version of our estimator of β_1, accounting for possible nonlinearities in X_{-1}, would still be consistent and asymptotically normal.



$\varepsilon \sim \mathcal{N}(0,1)$, η standard Laplace. ε standard Laplace, $\eta \sim \mathcal{N}(0,1)$.

Figure 1: $\tau \mapsto R(\tau)$ (dark line) for two pairs of distributions on ε and η.

2.2 Exogenous selection

When selection is conditionally exogenous, $D \perp\!\!\!\perp Y^*|X$, Assumption 2 is violated unless we also have $D \perp\!\!\!\perp X$. However, under some tail conditions on ε, our estimator can still be consistent, because Equation (2.2) in the main paper may still be satisfied. To illustrate this, suppose for simplicity that $X = X_1$ and $\varepsilon \sim \mathcal{N}(\beta_0, \sigma^2)$. Then, for τ small enough, we have

$$\begin{aligned} P(-Y \leq y|X = x) &= P(D = 1, X'\beta_1 + \varepsilon \geq -y|X = x) \\ &= P(D = 1|X = x)P(-\varepsilon \leq y + x'\beta_1). \end{aligned}$$

This implies that

$$Q_{-Y|X}(\tau|x) = -x_1'\beta_1 - \beta_0 + \sigma\Phi^{-1}(\tau/P(D = 1|X = x)).$$

Now, standard normal quantiles satisfy $\Phi^{-1}(\tau) = -(2\ln(1/\tau))^{1/2} + o(1)$ as $\tau \to 0$. Hence, by a Taylor expansion, we get

$$\begin{aligned} Q_{-Y|X}(\tau|x) &= -x_1'\beta_1 - \beta_0 - \sigma[2\ln(1/\tau) + 2\ln(P(D = 1|X = x))]^{1/2} + o(1) \\ &= -x_1'\beta_1 - \beta_0 - \sigma[2\ln(1/\tau)]^{1/2} + o(1). \end{aligned}$$

Thus, Equation (2.2) in the main paper still holds, with $\beta_0(\tau) = -\beta_0 - \sigma[2\ln(1/\tau)]^{1/2}$. The intuition here is that the propensity score $P(D = 1|X = x)$ does not play any role in this context, because it affects the conditional quantile of $-Y|X$ only through the quantile function of the error term, and the latter varies very little in the tail.

We confirm this through simulations, by considering almost the same model as in the simulation section, but with exogenous selection:

$$\begin{aligned} Y^* &= 0.2X_1 + 0.4X_2 + 0.5X_3 + (1 - 0.1X_2 - 0.5X_3)\varepsilon, \\ D &= \mathbb{1}\left\{0.6 + 0.3X_1 + 0.2X_2 + X_3^2 + \eta \geq 0\right\}, \end{aligned} \tag{2.1}$$

where ε and η are two independent standard normal variables. As can be seen in Table 3 below, the performance of our estimator are not far from those of the OLS estimator on the subsample $\{i : D_i = 1\}$, though this latter is the maximum likelihood estimator in this context, and is therefore asymptotically efficient.[3]

The result above is not specific to the normal distribution. The same reasoning applies with for instance Weibull distributions $\mathcal{W}(a, b)$ satisfying $b > 1$. We conjecture that it should hold more generally if $E[\exp(b \max(0, \varepsilon)^\delta)] < +\infty$ for some $b > 0$ and $\delta > 1$. On the other hand, it does not hold if we only impose $E[\exp(b \max(0, \varepsilon)] < +\infty$. To see this, suppose again that $X = X_1$, $P(D = 1|X) = \exp(\alpha + X'\gamma)$ (with $\alpha + X'\gamma \leq 0$) and ε follows a Laplace distribution. Then, for all (x, y) such that $y + x'\beta_1 \leq 0$,

$$P(-Y \leq y|X = x) = P(D = 1, x'\beta_1 + \varepsilon \geq -y|X = x) = \frac{\exp(\alpha + x'\gamma)}{2} \exp(y + x'\beta_1).$$

Thus, for τ small enough, $Q_{-Y|X}(\tau|X) = \ln(2\tau) - \alpha - X'(\beta_1 + \gamma)$, and the coefficient of an extremal quantile regression identifies $\beta_1 + \gamma$ instead of β_1. However, for more usual selection equations, the bias may be small in practice. To illustrate this, we ran a set of simulations based on (2.1), but with ε drawn from a Laplace distribution. It turns out that in such a case, the estimator actually performs very well in finite samples, even better than the OLS for the sample size $n = 1,000$ that we consider here (see again Table 3).

[3] To be precise, the OLS estimator is the MLE of the conditional model of $Y^*|D = 1, X$, not of the full model $(Y^*, D)|X$.

	Normal errors		Laplace errors	
	Extremal	OLS	Extremal	OLS
Bias	0.070	0	0.060	0
Std. dev.	0.113	0.092	0.068	0.094
RMSE	0.133	0.092	0.091	0.094

Note: Results for $\beta_1 = 0.2$, $n = 1,000$ and using 280 simulations. The true bias of the OLS estimator is equal to 0 in this setting.

Table 3: Simulation results under exogenous selection (Extremal and OLS estimator of β_1)

3 Details on the data-driven τ_n

We provide in this section a rationale for the construction of the data-driven τ_n detailed in Section 2.3. We study for that purpose the asymptotic behavior of $\widehat{\beta}_1$ for sequences τ_n that do not satisfy Assumption 6 (iii), but only $\sqrt{n\tau_n}\mathcal{B}(\tau_n) = O(1)$. We show that in this case, $\widehat{\beta}_1$ has an asymptotic bias. Then we relate this bias with the asymptotic behavior of the test statistic $T_J(\tau_n)$, and show how this can be used to select a quantile index for which the asymptotic bias is small.

First, recall that

$$\mu(\tau) = \frac{E\left[(\tau - \mathbb{1}\{\overline{X}'\beta(\tau) \leq Y\})\overline{X}\right]}{\tau}.$$

As shown in the proof of Theorem 2.2, $\mu(\tau)$ is the core component of the bias induced by the fact that (2.2) in the main paper holds only up to a $o(1)$ term. Note that $\mathcal{B}(\tau)$ in Assumption 6 is an upper bound of $||\mu(\tau)||$. Then, under Assumption 6(iii), $\sqrt{\tau_n n}\mu(\tau_n) \to 0$, which implies that the asymptotic bias vanishes. In what follows, we derive the asymptotic bias of our estimator $\widehat{\beta}_1$ as a function of $\mu(\tau)$ and propose a subsampling method to choose τ such that this bias is small.

First, by applying the convexity lemma and the same arguments as in the end of the proof of Theorem 1 in Pollard (1991), we obtain, as in (B.7) in the main paper but more generally, as long as τ_n satisfies Assumption 6(i) and (ii) and $\sqrt{\tau_n n}\mu(\tau_n) = O(1)$,

$$\widehat{Z}_n = -\mathcal{Q}_H^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n} M_{n,i}(\tau_n) - \mathcal{Q}_H^{-1}\sqrt{\tau_n n}\mu(\tau_n) + o_P(1).$$

The first term on the right-hand side converges in distribution to $\mathcal{N}(0, \Omega_0)$. But if $\sqrt{\tau_n n}\mu(\tau_n)$ is not negligible, the second term induces an asymptotic bias on $\widehat{Z}$ and thus on $\widehat{\beta}_1$.

To detect whether such a bias is large for a given τ, and thus select τ appropriately, we consider a very similar minimum distance statistic as the one used in Theorem 2.3. Define $M_{n,i}(\tau)$ as in the proof of Theorem 2.2 and let $\Psi = [I_{d_1}, 0_{d_1 \times (d-d_1)}]$, where $0_{d_1 \times d - d_1}$ denotes the zero matrix of size $d_1 \times d - d_1$. Then define

$$\begin{aligned}
\tilde{Z}_n(\tau) &= -[(1/\ell_1) - 1/\ell_2]^{-1/2} \widehat{\Omega}^{-1} \Psi \mathcal{Q}_H^{-1} \sum_{i=1}^{n} M_{n,i}(\tau)/\sqrt{n} \\
b_n(\tau) &= -[(1/\ell_1) - 1/\ell_2]^{-1/2} \widehat{\Omega}^{-1/2} \Psi \mathcal{Q}_H^{-1} \sqrt{\tau n} \left(\sqrt{\ell_1} \mu(\ell_1 \tau) - \sqrt{\ell_2} \mu(\ell_2 \tau) \right).
\end{aligned}$$

Then, reasoning as in the proof of Theorem 2.3, we obtain

$$[(1/\ell_1) - 1/\ell_2]^{-1/2} \widehat{\Omega}^{-1/2} (\widehat{\beta}_1(\ell_2 \tau) - \widehat{\beta}_1(\ell_1 \tau)) = \tilde{Z}_n(\tau) + b_n(\tau).$$

Hence,

$$T_J(\tau) = \left\| \tilde{Z}_n(\tau) + b_n(\tau) \right\|^2.$$

$\tilde{Z}_n(\tau_n)$ tends to a standard normal multivariate distribution under Assumption 6(i) and (ii). Thus, if $b_n(\tau_n) \xrightarrow{p} b$, the test statistic on the left-hand side converges to a non-central chi-squared distribution, with d_1 degrees of freedom and noncentrality parameter $\|b\|^2$. As a result, if $\sqrt{\tau_n n} \mu(\tau_n) \to 0$, then $b = 0$ and the median of $T_J(\tau_n)$ is asymptotically the median M_{d_1} of a $\chi^2(d_1)$. But if $\sqrt{\tau_n n} \mu(\tau_n) \to c \neq 0$, the asymptotic bias b will not vanish in general, and the difference between the median of $T_J(\tau_n)$ and M_{d_1} will generally be asymptotically different from zero. Following this idea, we estimate the difference between the two medians and use it as a proxy for the asymptotic bias of $\widehat{\beta}_1$. As indicated in the text, we rely for that purpose on subsampling.

4 Proofs of the results in the supplement

4.1 Proof of Proposition 1.1

We verify Assumption 3 with $h = 1$. By Assumption 7 and because $f_C(\tau) \to 0$, we have, as $y \to \infty$,

$$\begin{aligned}
|P(D=1|X=x, Y^*=y) - 1| &= \left| P\left(F_\nu(\nu) \leq F_\nu(\phi(x)) | F_{-\eta}(-\eta) = F_{-\eta}\left(\frac{\psi(x)-y}{\sigma(x)}\right)\right) - 1 \right| \\
&= \left| \partial_1 C\left[F_{-\eta}\left(\frac{\psi(x)-y}{\sigma(x)}\right), F_\nu(\phi(x))\right] - 1 \right| \\
&\leq \sup_{v \in [\underline{v}, 1]} \left| \partial_1 C\left[F_{-\eta}\left(\frac{\psi(x)-y}{\sigma(x)}\right), v\right] - 1 \right| \\
&\longrightarrow 0.
\end{aligned}$$

4.2 Proof of Proposition 1.2

Let $\overline{x} = [1, x']'$, we have

$$\begin{aligned}
\mathcal{B}(\tau) &\leq \sup_{\overline{x} \in \text{Supp}(X)} ||\overline{x}|| \sup_{x \in \text{Supp}(X), t \geq 1-\tau} |P(D=1|X=x, F_{Y^*|X}(Y^*|x) = t) - 1| \\
&= \sup_{\overline{x} \in \text{Supp}(\overline{X})} ||\overline{x}|| \sup_{x \in \text{Supp}(X), u \leq \tau} |P(\nu \leq \phi(x) | F_{-\eta}(-\eta) = u) - 1| \\
&= \sup_{\overline{x} \in \text{Supp}(\overline{X})} ||\overline{x}|| \sup_{x \in \text{Supp}(X), u \leq \tau} |P(F_\nu(\nu) \leq F_\nu(\phi(x)) | F_{-\eta}(-\eta) = u) - 1| \\
&\leq \sup_{\overline{x} \in \text{Supp}(\overline{X})} ||\overline{x}|| \, f_C(\tau),
\end{aligned}$$

where the second inequality follows from $f_C(\tau) = \sup_{u \leq \tau, v \in [\underline{v}, 1]} |\partial_1 C(u, v) - 1|$.

Now, since $f_C(\tau) = o(\tau^a)$ for some $a > 0$, let $\tau_n = n^{-1/(2a+1)}$. Then $\tau_n \to 0$ and $n\tau_n \to \infty$. Because $\mathcal{B}(\tau) = o(\tau^a)$, we also have

$$\sqrt{\tau_n n} \mathcal{B}(\tau_n) = n^{a/(2a+1)} o\left(n^{-a/(2a+1)}\right) = o(1).$$

Hence, this choice of τ_n satisfies Assumption 6. Besides, by Lemma 5.2, $f_{-\eta}(Q_{-\eta}(\tau_n)) \sim \tau_n L(n)$ for n large enough and $L(\cdot)$ some slowly varying function. Then the convergence λ_n is ,

$$\sqrt{n\tau_n} L_n = n^{a/(2a+1)} L(n).$$

This ensures that $\widehat{\beta}_1$ has a polynomial rate of convergence. With such a τ_n, the rate of convergence of $\widehat{\beta}_1$ is faster than $n^{(a-\alpha)/(2a+1)}$, for any $\alpha > 0$, which is also polynomial.

4.3 Verification of (1.1) and (1.2) for several copulas

Case 1: Gaussian copula with $\rho > 0$. We just check (1.2), which is stronger than (1.1). We have, after some algebra,

$$
\begin{aligned}
&1 - \partial_1 C_\rho(u,v) \\
=&1 - \frac{1}{\varphi(\Phi^{-1}(u))} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp(-(\Phi^{-1}(u)^2 - 2\rho\Phi^{-1}(u)s + s^2)/[2(1-\rho^2)])ds \\
=&1 - \frac{e^{-(1-\rho^2)\Phi^{-1}(u)^2/[2(1-\rho^2)]}}{\sqrt{2\pi}\varphi(\Phi^{-1}(u))} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{\sqrt{(2\pi)1-\rho^2}} \exp\left(-(s-\rho\Phi^{-1}(u))^2/[2(1-\rho^2)]\right) ds \\
=&\Phi\left(\frac{\rho\Phi^{-1}(u) - \Phi^{-1}(v)}{\sqrt{1-\rho^2}}\right).
\end{aligned}
$$

Thus, because $\rho > 0$,

$$\sup_{v\in[\underline{v},1], u\leq\tau} 1 - \partial_1 C_\rho(u,v) = \Phi\left(\frac{\rho\Phi^{-1}(\tau) - \Phi^{-1}(\underline{v})}{\sqrt{1-\rho^2}}\right).$$

Now, as $x \to -\infty$, we have $\Phi(x) \sim -\varphi(x)/x$. Because for any $K > 0$, $\exp(-Kx^2) \leq -1/x \leq 1$ for x small enough, we have $\varphi(x/\sigma) \leq \Phi(x) \leq \varphi(x)$ for any $0 < \sigma < 1$. This also implies that $\Phi^{-1}(\tau) \leq \sigma\varphi^{-1}(\tau)$, for τ small enough and with φ^{-1} the inverse of φ on $(-\infty, 0]$. Similarly, for any $m > 0$, there exists $\sigma > 1$ such that for any x small enough, $\varphi(x+m) \leq \varphi(x/\sigma)$. Combining these inequalities, we obtain, for any $K < \rho/\sqrt{1-\rho^2}$,

$$f_C(\tau) \leq \varphi(K\varphi^{-1}(\tau)) = K'\varphi(\Phi^{-1}(\tau))^{K^2} \leq \sqrt{2\pi}^{K^2-1}\tau^{K^2}.$$

The result follows.

Case 2: Archimedean copulas with $\lim_{u\to 0}\Psi(u) = +\infty$ and $\Psi \in RV_\alpha(0)$ with $\alpha \in (0, +\infty]$. Because Ψ is decreasing, we have, by Proposition 0.8 of Resnick (1987), $\Psi^{-1} \in RV_{1/\alpha}(\infty)$. As a result, for all $v \in [\underline{v}, 1]$,

$$u \geq C(u,v) \geq \Psi^{-1}(\Psi(u) + \Psi(\underline{v})) \sim \Psi^{-1}(\Psi(u)) = u \text{ as } u \to 0.$$

In other words,

$$\lim_{u\to 0} \sup_{v\in[\underline{v},1]} |C(u,v)/u - 1| = 0.$$

This implies that

$$\sup_{v\in[\underline{v},1]}\left|\frac{\Psi'(u)}{\Psi'(C(u,v))}-1\right|=\left|\frac{\Psi'(u)}{\Psi'(l(u)u)}-1\right|. \tag{4.1}$$

for some function $l(.)$ tending to one as $u\to 0$. Now, by Proposition 0.7 of Resnick (1987), $\Psi'\in RV_{\alpha-1}(0)$. This implies that the left-hand side of (4.1) tends to 0. (1.1) follows by remarking that $\partial_1 C(u,v)=\Psi'(u)/\Psi'\circ C(u,v)$.

Case 3: Gumbel copulas with $\theta>1$. Some algebra yields

$$\partial_1 C(u,v)=\frac{1}{1+\Psi(v;\theta)/\Psi(u;\theta)}\frac{C(u,v)\log C(u,v)}{u\log u}.$$

Now, by the fact that $x\log(x)$ is decreasing when x is close to 0 and $C(u,v)\le u$, we have $C(u,v)\log C(u,v)\ge u\log(u)$, i.e. $\frac{C(u,v)\log C(u,v)}{u\log u}\le 1$. Because $v\mapsto C(u,v)$ is increasing, $C(u,v)\log C(u,v)\le C(u,\underline{v})\log C(u,\underline{v})$. Furthermore, $0\le\Psi(v,\theta)\le\Psi(\underline{v},\theta)$. Therefore, we have

$$\begin{aligned}
&\sup_{v\in[\underline{v},1]}|\partial_1 C(u,v)-1|\\
\le&\sup_{v\in[\underline{v},1]}\left(\left|\frac{C(u,v)\log C(u,v)}{u\log u}-1\right|+\left|\partial_1 C(u,v)-\frac{C(u,v)\log C(u,v)}{u\log u}\right|\right)\\
\le&\sup_{v\in[\underline{v},1]}\left(1-\frac{C(u,v)\log C(u,v)}{u\log u}\right)+\sup_{v\in[\underline{v},1]}\left(\frac{\Psi(v,\theta)}{\Psi(v,\theta)+\Psi(u,\theta)}\right)\frac{C(u,v)\log C(u,v)}{u\log u}\\
\le&(1-\frac{C(u,\underline{v})\log C(u,\underline{v})}{u\log u})+\frac{\Psi(\underline{v},\theta)}{\Psi(\underline{v},\theta)+\Psi(u,\theta)}
\end{aligned}$$

$\Psi(u,\theta)\to\infty$ as $u\to 0$, so the second term also converges 0. Therefore, to prove (1.1), it suffices to show that $C(u,\underline{v})\sim u$. We have, for $\theta>1$,

$$\begin{aligned}
C(u,\underline{v})&=\exp\left[-\left((-\log u)^\theta+(-\log\underline{v})^\theta\right)^{1/\theta}\right]\\
&=\exp\left[\log u\left(1+\left(\frac{-\log\underline{v}}{-\log u}\right)^\theta\right)^{1/\theta}\right]\\
&=\exp\left[\log u+\frac{(-\log\underline{v})^\theta}{\theta(-\log u)^{\theta-1}}+o\left(\frac{1}{(-\log u)^{\theta-1}}\right)\right]\\
&\sim u.
\end{aligned}$$

Case 4: Clayton copula with $\theta > 0$. We obtain in this case

$$1 - \partial_1 C(u, v; \theta) \leq K u^\theta \left(\frac{1}{v^\theta} - 1 \right)$$

Hence, $f_C(\tau) \leq K' \tau^\theta$, where $K' = K \left(\frac{1}{\underline{v}^\theta} - 1 \right)$. (1.2) follows.

Case 5: Rotated Gumbel-Barnett copula with $\theta \in (0, 1]$. We have

$$1 - \partial_1 C(u, v; \theta) = (1 - v) \exp(-\theta \log(u) \log(1 - v))(1 - \theta \log(1 - P)) \leq O \left(u^{-\theta \log(1 - \underline{v})} \right) \quad (4.2)$$

It follows that (1.2) holds.

Case 6: $C(u, v; \theta) = (1 + [(u^{-1} - 1)^\theta + (v^{-1} - 1)^\theta]^{1/\theta})^{-1}$ with $\theta > 1$. In this case,

$$1 - \partial_1 C(u, v; \theta) = 1 - \left(\frac{1}{u + [(1 - u)^\theta + u^\theta (v^{-1} - 1)^\theta]^{1/\theta}} \right)^2 \left[1 + \left(\frac{v^{-1} - 1}{u^{-1} - 1} \right)^\theta \right]^{1/\theta - 1} \leq K u.$$

(1.2) follows.

Case 7: $C(u, v; \theta) = (1 + [(u^{-1/\theta} - 1)^\theta + (v^{-1/\theta} - 1)^\theta]^{1/\theta})^{-\theta}$ with $\theta \geq 1$. We have

$$\begin{aligned} \partial_1 C(u, v; \theta) &= 1 - \left(u^{1/\theta} + [(1 - u^{1/\theta})^\theta + u (v^{-1/\theta} - 1)^\theta]^{1/\theta} \right)^{-\theta - 1} \left[1 + \left(\frac{v^{-\frac{1}{\theta}} - 1}{u^{-\frac{1}{\theta}} - 1} \right)^\theta \right]^{1/\theta - 1} \\ &\leq K u^{1/\theta}, \end{aligned}$$

which implies (1.2).

Case 8: $C(u, v; \theta) = \theta / \log(\exp(\theta/u) + \exp(\theta/v) - \exp(\theta))$ with $\theta > 0$. We have

$$\begin{aligned} &1 - \partial_1 C(u, v; \theta) \\ &= 1 - 1/(1 + \log(1 + (\exp(\theta/v) - \exp(\theta)) \exp(-\theta/u)))^2 \frac{1}{1 + (\exp(\theta/v) - \exp(\theta)) \exp(-\theta/u)} \\ &\leq K \exp(-\theta/u). \end{aligned}$$

Thus Condition (1.2) is easily satisfied. In this case, any polynomial rate slower than the parametric rate is in fact possible.

Case 9: $C(u,v;\theta) = [\log(\exp(u^{-\theta}) + \exp(v^{-\theta}) - e)]^{-1/\theta}$ with $\theta > 0$. Start from

$$\begin{aligned}
1 - \partial_1 C(u,v;\theta) &= 1 - \left[1 + u^{\theta}\log\left(1 + \frac{\exp(v^{-\theta}) - e}{\exp(u^{-\theta})}\right)\right]^{-1/\theta - 1} \frac{1}{1 + \frac{\exp(v^{-\theta}) - e}{\exp(u^{-\theta})}} \\
&\leq K_1 u^{\theta}\log\Big(1 + [\exp(v^{-\theta}) - e]\exp(-u^{-\theta})\Big) + K_2\exp(-u^{-\theta}) \\
&\leq K\exp(-u^{-\theta}).
\end{aligned}$$

Therefore, Condition (1.2) is easily satisfied and once more, any polynomial rate slower than parametric rate is possible.

5 Technical lemmas used in the main proofs

Lemma 5.1. *For any random variable V such that* $\sup(Supp(V)) = +\infty$*,* $E\exp(b\max(0,V)) < \infty$ *and V is in the attraction domain of generalized extreme value distributions,*

1. *S_V is rapidly varying at its upper tail;*
2. *If, in addition,* $\exp(V)$ *is in the attraction domain of generalized extreme value distributions and F_V has a positive derivative on $[A, +\infty)$ for some B, $Q_{-V}(e\tau) - Q_{-V}(\tau)$ is bounded.*

Proof. We first show that V is rapidly varying at its upper tail. Because $\sup(\text{Supp}(V)) = \infty$, S_V is not in the attraction domain of type III extreme value distributions (see Resnick, 1987, Proposition 1.13). Suppose S_V is not rapidly varying. Then, S_V is not either in the attraction domain of type I extreme value distribution (See Resnick, 1987, Exercise 1.1.9). So S_V is in the attraction domain of type II extreme value distribution, i.e. $S_V \in RV_{-\xi^{-1}}(+\infty)$ with extreme value index $\xi > 0$. We also have

$$\frac{S_{\exp(V)}(tx)}{S_{\exp(V)}(x)} = \frac{S_V(u(x)\log(x))}{S_V(\log(x))} \tag{5.1}$$

where $u(x) = \frac{\log(t)+\log(x)}{\log(x)} \to 1$ as $x \to +\infty$. Because $S_V \in RV_{-\xi^{-1}}(+\infty)$, the right-hand side of Equation (5.1) converges to 1, for any t. However, Lemma 2.1 of D'Haultfoeuille & Maurel (2013) shows that, given $E\exp(b\max(0,V)) < \infty$ for some $b > 0$, $x \mapsto S_{\exp(V)}(tx)/S_{\exp(V)}(x)$ does not tend to one unless $t = 1$. Therefore, we have reached a contradiction, which proves the first point.

Next, since $\exp(V)$ is in the attraction domain of generalized extreme value distributions, it can be only regularly varying or rapidly varying. In both cases, for any $l' > 1$, there exists a constant $\delta > 0$ such that, for all y sufficiently large,

$$\frac{S_{\exp(V)}(y)}{S_{\exp(V)}(l'y)} \geq 1 + \delta.$$

Denote $l' = \exp(l)$ and $y = \exp(z)$, we have

$$S_V(l+z) = S_{\exp(V)}(l'y) \leq \frac{S_{\exp(V)}(y)}{1+\delta} = \frac{S_V(z)}{1+\delta}.$$

This implies, for $l > 0$,

$$F_{-V}(-l-z) \leq \frac{F_{-V}(-z)}{1+\delta}.$$

Let $-z = Q_{-V}(\tau)$ for τ sufficiently small. Because $F_{-V}(-z) = \tau$ since F_V has a positive derivative on $[A, +\infty)$,

$$F_{-V}(-l + Q_{-V}(\tau)) \leq \frac{\tau}{1+\delta}.$$

Thus,

$$Q_{-V}(\tau) - Q_{-V}(\tau/(1+\delta)) \leq l.$$

Since V is rapidly varying at its upper tail, $\frac{Q_{-V}(\tau)-Q_{-V}(\tau/(1+\delta))}{Q_{-V}(e\tau)-Q_{-V}(\tau)} \to \log(1+\delta)$, as $\tau \to 0$. Therefore, for τ sufficiently small, $0 < Q_{-V}(e\tau) - Q_{-V}(\tau) < 2l/(\log(1+\delta))$. □

Lemma 5.2. *Suppose that Assumptions 2 and 5 hold. Then, as* $\tau \to 0$,

$$Q_{-U|X}(e\tau|x) - Q_{-U|X}(\tau|x) \sim H(x_{-1})(Q_{-\eta}(e\tau) - Q_{-\eta}(\tau)) \tag{5.2}$$

uniformly in $x \in Supp(X)$. *Moreover, as* $\tau \to 0$,

$$Q'_{-\eta}(\tau) \sim \frac{Q_{-\eta}(e\tau) - Q_{-\eta}(\tau)}{\tau}. \tag{5.3}$$

Finally, there exists $M \in (0, \infty)$ *such that for all* τ *small enough and all* $x \in Supp(X)$,

$$f_{-U|X}(Q_{-U}(\tau|x_{-1})|x) \leq M f_{-\eta}(Q_{-\eta}(\tau)). \tag{5.4}$$

Proof. **1. Equivalence** (5.2)**.**

We first show that, uniformly in $x \in \text{Supp}(X)$,

$$\frac{Q_{-U|X}(\tau|x) - Q_{-\eta}(\tau)H(x_{-1})}{Q_{-\eta}(e\tau) - Q_{-\eta}(\tau)} \to 0. \tag{5.5}$$

Since (2.4) holds uniformly in x, for any sequences $z_n \to \infty$ and $x_n \to x_0$, we have

$$\frac{F_{-\eta}(-z_n/H(x_{-1n}))}{F_{-U|X}(-z_n|x_n)} \to 1.$$

Let $-z_n = Q_{-U|X}(\tau_n|x_n)$ for some arbitrary sequence $\tau_n \to 0$, then we have

$$\frac{F_{-\eta}(Q_{-U|X}(\tau_n|x_n)/H(x_{-1n}))}{\tau_n} \to 1.$$

In addition, let $s_n = \frac{Q_{-U|X}(\tau_n|x_n)-H(x_{-1n})Q_{-\eta}(\tau_n)}{H(x_{-1n})a(\tau_n)}$ and $a(\tau) = Q_{-\eta}(e\tau) - Q_{-\eta}(\tau)$, then

$$\frac{F_{-\eta}(Q_{-U|X}(\tau_n|x_n)/H(x_{-1n}))}{\tau_n} = \frac{F_{-\eta}(Q_{-\eta}(\tau_n) + s_n a(\tau_n))}{\tau_n} \to 1. \tag{5.6}$$

Next, $S_{U|X}$ is in the attraction domain of type I extreme value distribution. Thus, by Proposition 0.10 in Resnick (1987), $S_{U|X}(z|x)$ is Γ-varying, i.e. for fixed x, there exists an auxiliary function f_x such that as $z \to \infty$,

$$\frac{S_{U|X}(z + f_x(z)t|x)}{S_{U|X}(z|x)} \to e^t.$$

In addition, $f_x(z)$ is slowly varying in z by Proposition 0.12 of Resnick (1987). So $z(1 + f_x(z)t/z) \to \infty$ and

$$\frac{S_{U|X}(z + f_x(z)t|x)}{S_\eta(z/H(x_{-1}) + f_x(z)t/H(x_{-1}))} \to 1.$$

We also have $S_{U|X}(z|x)/S_\eta(z/H(x_{-1})) \to 1$. So

$$\frac{S_\eta(z/H(x_{-1}) + f_x(z)t/H(x_{-1}))}{S_\eta(z/H(x_{-1}))} \to e^t.$$

Let $z' = z/H(x_{-1})$ and $f'(z') = f_x(H(x_{-1})z')/H(x_{-1})$, then we have

$$\frac{S_\eta(z' + R'(z')t)}{S_\eta(z')} \to e^t,$$

i.e. S_η is Γ-varying. By Proposition 0.10 in Resnick (1987) again, it means that S_η is in the attraction domain of type I extreme value distribution and thus $Q_{-\eta}$ is Π-varying, i.e.

$$\frac{Q_{-\eta}(t\tau) - Q_{-\eta}(\tau)}{Q_{-\eta}(e\tau) - Q_{-\eta}(\tau)} \to \log(t) \tag{5.7}$$

as $\tau \to 0$, locally uniformly in t. Invert both sides, by Proposition 0.1 in Resnick (1987) and the fact that $F_{-\eta}(Q_{-\eta}(\tau)+sa(\tau))$ is monotone increasing in s, we have

$$\frac{F_{-\eta}(Q_{-\eta}(\tau)+sa(\tau))}{\tau}\to e^s$$

locally uniformly in s. This and (5.6) imply that any convergent subsequence of s_n has limit zero, which means $s_n\to 0$. Last, since $H(x_{-1})$ is bounded away from zero uniformly in $x\in\text{Supp}(X)$, we have

$$\frac{Q_{-U|X}(\tau_n|x_n)-H(x_{-1n})Q_{-\eta}(\tau_n)}{a(\tau_n)}\to 0.$$

Since x_n and τ_n are arbitrary, we have, as $\tau\to 0$,

$$\frac{Q_{-U|X}(\tau|x)-H(x_{-1})Q_{-\eta}(\tau)}{a(\tau)}\to 0$$

uniformly in $x\in\text{Supp}(X)$.

Given (5.5), (5.2) holds uniformly in $x\in\text{Supp}(X)$ because

$$\begin{aligned}&\frac{Q_{-U|X}(e\tau|x)-Q_{-U|X}(\tau|x)}{Q_{-\eta}(e\tau)-Q_{-\eta}(\tau)}\\ =\;&\frac{Q_{-U|X}(e\tau|x)-Q_{-\eta}(e\tau)H(x_{-1})}{Q_{-\eta}(e\tau)-Q_{-\eta}(\tau)}+H(x_{-1})+\frac{H(x_{-1})Q_{-U|X}(\tau)-Q_{-\eta}(\tau|x)}{Q_{-\eta}(e\tau)-Q_{-\eta}(\tau)}\\ \to\;&H(x_{-1}).\end{aligned}$$

2. Equivalence (5.3)**.**

Because $\tau\mapsto -Q_{-\eta}(\tau)$ is Π-varying with auxiliary function $\tau\mapsto Q_{-\eta}(e\tau)-Q_{-\eta}(\tau)$, we have

$$\frac{Q_{-\eta}(b\tau)-Q_{-\eta}(a\tau)}{Q_{-\eta}(e\tau)-Q_{-\eta}(\tau)}\to\log(b)-\log(a). \tag{5.8}$$

By monotonicity of $Q'_{-\eta}$,

$$\frac{Q'_{-\eta}(b\tau)\tau(b-a)}{Q_{-\eta}(e\tau)-Q_{-\eta}(\tau)}\geq\frac{Q_{-\eta}(b\tau)-Q_{-\eta}(a\tau)}{Q_{-\eta}(e\tau)-Q_{-\eta}(\tau)}\geq\frac{Q'_{-\eta}(a\tau)\tau(b-a)}{Q_{-\eta}(e\tau)-Q_{-\eta}(\tau)},$$

for any $b>a>0$. Therefore, using (5.8),

$$\limsup_{\tau\to 0}\frac{Q'_{-\eta}(a\tau)\tau}{Q_{-\eta}(e\tau)-Q_{-\eta}(\tau)}\leq\frac{\log(b)-\log(a)}{b-a}.$$

Letting $b \downarrow a$, we obtain

$$\limsup_{\tau \to 0} \frac{Q'_{-\eta}(a\tau)\tau}{Q_{-\eta}(e\tau) - Q_{-\eta}(\tau)} \leq \frac{1}{a},$$

for any $a > 0$. Similarly, we obtain from the other inequality

$$\liminf_{\tau \to 0} \frac{Q'_{-\eta}(b\tau)\tau}{Q_{-\eta}(e\tau) - Q_{-\eta}(\tau)} \geq \frac{1}{b},$$

for any $b > 0$. By letting $a = b = 1$, we finally obtain (5.3).

3. Inequality (5.4)

From the previous proof, we have, for $b > a > 0$,

$$Q'_{-U|X}(b\tau|x)\tau(b-a) \geq Q_{-U|X}(b\tau|x) - Q_{-U|X}(a\tau|x)$$

and

$$Q'_{-\eta}(a\tau)\tau(b-a) \leq Q_{-\eta}(b\tau) - Q_{-\eta}(a\tau).$$

In addition, by the uniform equivalence in (5.2), there exists M independent of x and some τ_0 small enough, such that when $\tau \leq \tau_0$,

$$Q_{-U|X}(b\tau|x) - Q_{-U|X}(a\tau|x) \geq \frac{1}{M}\Big(Q_{-\eta}(b\tau) - Q_{-\eta}(a\tau)\Big).$$

Therefore, for $\tau \leq \tau_0$ and any $b > a > 0$,

$$f_{-U|X}(Q_{-U|X}(b\tau|x)|x) \leq M f_{-\eta}(Q_{-\eta}(a\tau)).$$

Let $a = 1$ and $b \downarrow 1$, we finally obtain Inequality (5.4). □

Lemma 5.3. *Suppose that Assumptions 2–5 hold and let* $\breve{U} = Y - X_1\beta_1 - X'_{-1}\beta_{-1,r}$. *Then, for all* $x \in Supp(X)$,

$$\lim_{\tau \to 0} \left| \frac{Q_{-\breve{U}|X}(\tau|x) - Q_{-U|X}(\tau/h|x)}{Q_{-U|X}(e\tau|x) - Q_{-U|X}(\tau|x)} \right| = 0, \tag{5.9}$$

$$f_{-\breve{U}|X}\Big(Q_{-\breve{U}|X}(\tau|x)\Big) \sim h f_{-\eta}(Q_{-\eta}(\tau/h))/H(x_{-1}). \tag{5.10}$$

Proof. For the first point, fix $\Delta \in (0, h)$ and remark first that by (5.8),

$$\lim_{\tau \to 0} \frac{Q_{-\eta}(l\tau) - Q_{-\eta}(\tau)}{Q_{-\eta}(e\tau) - Q_{-\eta}(\tau)} = \log(l)$$

for all $l > 0$. Then

$$\begin{aligned}
& \frac{Q_{-\eta}(\tau/(h-\Delta)) - Q_{-\eta}(\tau/(h+\Delta))}{Q_{-\eta}(e\tau) - Q_{-\eta}(\tau)} \\
= \quad & \frac{Q_{-\eta}(\tau/(h-\Delta)) - Q_{-\eta}(\tau/(h+\Delta))}{Q_{-\eta}(e\tau) - Q_{-\eta}(\tau)} \Big/ \frac{Q_{-\eta}(e\tau) - Q_{-\eta}(\tau)}{Q_{-\eta}(e\tau) - Q_{-\eta}(\tau)} \\
\to \quad & \log(\frac{h+\Delta}{h-\Delta}).
\end{aligned}$$

Besides, by definition of the quantiles of $-\breve{U}|X = x$, we have, for all τ small enough,

$$\begin{aligned}
\tau \quad & \leq \quad P(-\breve{U} \leq Q_{-\breve{U}|X}(\tau|x)|X = x) \\
& = \quad P(Y \geq x_1\beta_1 + x'_{-1}\beta_{-1,r} + Q_{\breve{U}|X}(1-\tau|x)|X = x) \\
& = \quad P(Y^* \geq x_1\beta_1 + x'_{-1}\beta_{-1,r} + Q_{\breve{U}|X}(1-\tau|x), D = 1|X = x) \\
& = \quad \int_{x_1\beta_1 + x'_{-1}\beta_{-1,r} + Q_{\breve{U}|X}(1-\tau|x)}^{\infty} P(D = 1|Y^* = y, X = x) dP^{Y^*|X=x}(y).
\end{aligned}$$

For τ small enough, $P(D = 1|Y^* = y, X = x) \in [h - \Delta, h + \Delta]$ for all $y > x_1\beta_1 + x'_{-1}\beta_{-1,r} + Q_{\breve{U}|X}(1 - \tau|x)$. Thus,

$$\tau \leq (h + \Delta) P\left[-U \leq Q_{-\breve{U}|X}(\tau|x)\right].$$

Similarly, using $\tau \geq P(-\breve{U} < -Q_{\breve{U}|X}(\tau|x)|X = x)$,

$$\tau \geq (h - \Delta) P\left[-U \leq Q_{-\breve{U}|X}(\tau|x)\right].$$

Then, by definition of the quantiles of $-U$,

$$Q_{-U|X}(\tau/(h+\Delta)|x) \leq Q_{-\breve{U}|X}(\tau|x) \leq Q_{-U|X}(\tau/(h-\Delta)|x).$$

This implies

$$\begin{aligned}
& \limsup_{\tau} \left| \frac{Q_{-\breve{U}|X}(\tau|x) - Q_{-U|X}(\tau/h|x)}{Q_{-U|X}(e\tau|x) - Q_{-U|X}(\tau|x)} \right| \\
\leq & \limsup_{\tau} \frac{\max\left(Q_{-U|X}(\tau/(h-\Delta)|x) - Q_{-U|X}(\tau/h|x),\, Q_{-U|X}(\tau/h|x) - Q_{-U|X}(\tau/(h+\Delta)|x)\right)}{Q_{-U|X}(e\tau|x) - Q_{-U|X}(\tau|x)} \\
\leq & \limsup_{\tau} \frac{Q_{-U|X}(\tau/(h-\Delta)|x) - Q_{-U|X}(\tau/(h+\Delta)|x)}{Q_{-U|X}(e\tau|x) - Q_{-U|X}(\tau|x)}. \qquad (5.11)
\end{aligned}$$

By (5.2),

$$Q_{-U|X}(\tau/(h-\Delta)|x) - Q_{-U|X}(\tau/(h+\Delta)|x) \sim H(x_{-1})\Big(Q_{-\eta}(\tau/(h-\Delta)) - Q_{-\eta}(\tau/(h+\Delta))\Big)$$

and

$$Q_{-U|X}(e\tau|x) - Q_{-U|X}(\tau|x) \sim H(x_{-1})\Big(Q_{-\eta}(e\tau) - Q_{-\eta}(\tau)\Big).$$

So following (5.11), we have

$$\begin{aligned}
&\limsup_{\tau}\left|\frac{Q_{-\breve{U}|X}(\tau|x) - Q_{-U|X}(\tau/h|x)}{Q_{-U|X}(e\tau|x) - Q_{-U|X}(\tau|x)}\right| \\
\le &\limsup_{\tau} \frac{Q_{-U|X}(\tau/(h-\Delta)|x) - Q_{-U|X}(\tau/(h+\Delta)|x)}{H(x_{-1})\Big(Q_{-\eta}(\tau/(h-\Delta)) - Q_{-\eta}(\tau/(h+\Delta))\Big)} \times \frac{Q_{-\eta}(\tau/(h-\Delta)) - Q_{-\eta}(\tau/(h+\Delta))}{Q_{-\eta}(e\tau) - Q_{-\eta}(\tau)} \\
&\times \frac{H(x_{-1})\Big(Q_{-\eta}(e\tau) - Q_{-\eta}(\tau)\Big)}{Q_{-U|X}(e\tau|x) - Q_{-U|X}(\tau|x)} \\
= &\log(h+\Delta) - \log(h-\Delta).
\end{aligned}$$

By letting Δ tend to 0, the left-hand side tends to zero. The first result follows.

Now let us turn to the last result. We first show that for any fixed x, $Q_{-\breve{U}|X}(\tau|x)$ is $\Pi-$varying. We have

$$\begin{aligned}
&\frac{Q_{-\breve{U}|X}(e\tau|x) - Q_{-\breve{U}|X}(\tau|x)}{Q_{-\eta}(e\tau) - Q_{-\eta}(\tau)} \\
= \quad &\frac{Q_{-\breve{U}|X}(e\tau|x) - Q_{-U|X}(e\tau/h|x)}{Q_{-\eta}(e\tau) - Q_{-\eta}(\tau)} - \frac{Q_{-\breve{U}|X}(\tau|x) - Q_{-U|X}(\tau/h|x)}{Q_{-\eta}(e\tau) - Q_{-\eta}(\tau)} \\
&+ \frac{Q_{-U|X}(e\tau/h|x) - Q_{-U|X}(\tau/h|x)}{Q_{-\eta}(e\tau) - Q_{-\eta}(\tau)}
\end{aligned}$$

By the first result of this lemma, the first and second term converge to zero. By (5.2),

$$\frac{Q_{-U|X}(e\tau/h|x) - Q_{-U|X}(\tau/h|x)}{Q_{-\eta}(e\tau) - Q_{-\eta}(\tau)} \sim H(x_{-1})\frac{Q_{-\eta}(e\tau/h) - Q_{-\eta}(\tau/h)}{Q_{-\eta}(e\tau) - Q_{-\eta}(\tau)} \sim H(x_{-1}).$$

Therefore

$$\frac{Q_{-\breve{U}|X}(e\tau|x) - Q_{-\breve{U}|X}(\tau|x)}{Q_{-\eta}(e\tau) - Q_{-\eta}(\tau)} \sim H(x_{-1}).$$

Then

$$\begin{aligned}
& \frac{Q_{-\breve{U}|X}(l\tau|x) - Q_{-\breve{U}|X}(\tau|x)}{Q_{-\breve{U}|X}(e\tau|x) - Q_{-\breve{U}|X}(\tau|x)} \\
= \quad & \frac{Q_{-\breve{U}|X}(l\tau|x) - Q_{-\breve{U}|X}(\tau|x)}{Q_{-\eta}(l\tau) - Q_{-\eta}(\tau)} \times \frac{Q_{-\eta}(e\tau) - Q_{-\eta}(\tau)}{Q_{-\breve{U}|X}(e\tau|x) - Q_{-\breve{U}|X}(\tau|x)} \times \frac{Q_{-\eta}(l\tau) - Q_{-\eta}(\tau)}{Q_{-\eta}(e\tau) - Q_{-\eta}(\tau)} \\
\to \quad & \ln(l),
\end{aligned}$$

which proves that $Q_{-\breve{U}|X}(.|x)$ is $\Pi-$varying. Now, remark that for y small enough,

$$\begin{aligned}
P(-\breve{U} \leq y|X = x) &= P(-Y + x_1\beta_1 + x'_{-1}\beta_{-1,r} \leq y|X = x) \\
&= P(Y^* \geq -y + x_1\beta_1 + x'_{-1}\beta_{-1,r}, D = 1|X = x) \\
&= P\left(-U \leq y|D = 1, X = x\right) P(D = 1|X = x).
\end{aligned}$$

Therefore, $f_{-\breve{U}|X}(y|X = x) = f_{-U|X,D=1}(y|x)P(D = 1|X = x)$. This equality, combined with Assumption 5(iii) and the fact that X is bounded, ensures that the pdf of $-\breve{U}|X$ is monotone increasing at the lower tail. As a result, $Q'_{-\breve{U}|X}(.|x)$ is decreasing at the lower tail and we have, by the same reasoning as in Lemma 5.2,

$$Q_{-\breve{U}|X}(\tau|x)' \sim \frac{(Q_{-\breve{U}|X}(e\tau|x) - Q_{-\breve{U}|X}(\tau|x))}{\tau}. \tag{5.12}$$

Combining Equations (5.3) and (5.12), we obtain

$$\frac{f_{-\breve{U}|X}\left(Q_{-\breve{U}|X}(\tau|x)\right)}{f_{-\eta}(Q_{-\eta}(\tau/h))} = \frac{Q_{-\eta}(\tau/h)'}{Q_{-\breve{U}|X}(\tau|x)'} \sim \frac{h(Q_{-\eta}(e\tau) - Q_{-\eta}(\tau))}{Q_{-\breve{U}|X}(e\tau|x) - Q_{-\breve{U}|X}(\tau|x)} \sim \frac{h}{H(x_{-1})}.$$

This proves the second result of the lemma. □

Lemma 5.4. *Suppose that Assumptions 1–6 hold and let $\Lambda_n(z,\tau)$ be defined as in* (B.1). *Then*

$$\Lambda_n(z, \tau_n) \xrightarrow{p} \frac{1}{2} z' \mathcal{Q}_H z.$$

The same result holds for its bootstrap counterpart $\Lambda^*_n(z, \tau_n)$.

Proof. By Lemma 9.6 in Chernozhukov (2005), the variance of $\Lambda_n(z,\tau)$ converges to 0. Thus it suffices to prove that $E[\Lambda_n(z,\tau_n)] \to \frac{1}{2}z'\mathcal{Q}_H z$. Let us define, for any $(s,t) \in \mathbb{R}^2$,

$$m(s,t) = \begin{cases} 1 & \text{if } 0 < s \le t, \\ -1 & \text{if } t \le s < 0, \\ 0 & \text{otherwise.} \end{cases}$$

We have

$$\begin{aligned} & E\left[\Lambda_n(z,\tau_n)\right] \\ = \; & \frac{\lambda_n}{\sqrt{\tau_n n}} nE\left[\int_0^{(z_1+X'z_2)/\lambda_n} \mathbb{1}\{-\breve{U} - Q_{-U|X}(\tau_n/h|X) \le s\} - \mathbb{1}\{-\breve{U} - Q_{-U|X}(\tau_n/h|X) \le 0\} ds\right] \\ = \; & \frac{n}{\sqrt{\tau_n n}} E\left[\int_0^{z_1+X'z_2} \mathbb{1}\{-\breve{U} - Q_{-U|X}(\tau_n/h|X) \le s/\lambda_n\} - \mathbb{1}\{-\breve{U} - Q_{-U|X}(\tau_n/h|X) \le 0\} ds\right] \\ = \; & nE\left[\int_0^{z_1+X'z_2} \frac{F_{-\breve{U}|X}(-Q_{-U|X}(\tau_n/h|X) + s/\lambda_n) - F_{-\breve{U}|X}(-Q_{-U|X}(\tau_n/h|X))}{\sqrt{\tau_n n}} ds\right] \\ = \; & E\left[\int_{-\infty}^{+\infty} m(s, z_1 + X'z_2)s \, \frac{nf_{-\breve{U}|X}\left[-Q_{-U|X}(\tau_n/h|X) + V_s\right]}{\lambda_n\sqrt{\tau_n n}} ds\right], \qquad (5.13) \end{aligned}$$

where for each s, V_s is a random variable satisfying $V_s \in [0, s/\lambda_n]$. Let

$$U_n(s) = m(s, z_1 + X'z_2)s \, \frac{nf_{-\breve{U}|X}\left[-Q_{-U|X}(\tau_n/h|X) + V_s\right]}{\lambda_n\sqrt{\tau_n n}}.$$

We first show that

$$U_n(s) \xrightarrow{p} \frac{m(s, z_1 + X'z_2)s}{H(X_{-1})}. \qquad (5.14)$$

By Lemma 5.2 and 5.3,

$$1/\lambda_n = \frac{\sqrt{\tau_n}Q'_{-\eta}(\tau_n/h)}{\sqrt{n}h} \sim \frac{Q_{-\eta}(e\tau_n) - Q_{-\eta}(\tau_n)}{\sqrt{n\tau_n}} \sim \frac{Q_{-U|X}(e\tau_n|X) - Q_{-U|X}(\tau_n|X)}{H(X_{-1})\sqrt{n\tau_n}}.$$

Thus $V_s = o_P(Q_{-U|X}(e\tau_n|X) - Q_{-U|X}(\tau_n|X))$. Moreover, by Lemma 5.3,

$$Q_{-U|X}(\tau_n/h|X) - Q_{-\breve{U}|X}(\tau_n|X) = o_P(Q_{-U|X}(e\tau_n|X) - Q_{-U|X}(\tau_n|X)).$$

Then, following the same argument as Chernozhukov (2005) after his Equation (9.57),

$$
\begin{aligned}
& \frac{f_{-\breve{U}|X}\left[Q_{-U|X}(\tau_n/h|X)+V_s\right]}{f_{-\breve{U}|X}\left(Q_{-\breve{U}|X}(\tau_n|x)\right)} \\
= \quad & \frac{f_{-\breve{U}|X}\left[Q_{-\breve{U}|X}(\tau_n|X)+(Q_{-U|X}(\tau_n/h|X)-Q_{-\breve{U}|X}(\tau_n|X))+V_s\right]}{f_{-\breve{U}|X}\left(Q_{-\breve{U}|X}(\tau_n|x)\right)} \\
\xrightarrow{p} \quad & 1.
\end{aligned} \tag{5.15}
$$

Besides, by Lemma 5.3,

$$
f_{-\breve{U}|X}\left(Q_{-\breve{U}|X}(\tau_n|X)\right) \sim hf_{-\eta}(Q_{-\eta}(\tau_n/h))/H(X_{-1}) = \frac{\lambda_n}{H(X_{-1})}\sqrt{\frac{\tau_n}{n}}. \tag{5.16}
$$

This implies that (5.14) holds.

Next, we prove that for n sufficiently large,

$$
|U_n(s)| \leq U(s), \quad \text{with} \quad E\left(\int_{-\infty}^{\infty} U(s)ds\right) < \infty. \tag{5.17}
$$

We bound $|U_n(s)|$ for $|s| \leq |z_1 + X'z_2|$, since $m(s, z_1 + X'z_2) = 0$ otherwise.

First, because X is bounded and $\lambda_n \to \infty$, $\sup_{x\in \mathrm{Supp}(X)} \frac{|z_1+X'z_2|}{\lambda_n} \to 0$. Thus, for any $|s| \leq |z_1 + X'z_2|$, and n sufficiently large,

$$
\begin{aligned}
& P\Big(Y \in \Big(\gamma(1-\tau_n/h)+X'\beta(1-\tau_n/h), \gamma(1-\tau_n/h)+X'\beta(1-\tau_n/h)+s/\lambda_n\Big], D=0\Big) \\
\leq \quad & P\Big(0 \in \Big(\gamma(1-\tau_n/h)+X'\beta(1-\tau_n/h), \gamma(1-\tau_n/h)+X'\beta(1-\tau_n/h)+s/\lambda_n\Big]\Big) \\
\leq \quad & \mathbb{1}\left\{\sup_{x\in \mathrm{Supp}(X)} P\left(Y^* \leq |s|/\lambda_n | X = x\right) \geq 1-\tau_n/h\right\} \\
= \quad & 0.
\end{aligned}
$$

Hence, by definition of Y and Y^*,

$$
\begin{aligned}
& \Big\{Y \in \Big(\gamma(1-\tau_n/h)+X'\beta(1-\tau_n/h), \gamma(1-\tau_n/h)+X'\beta(1-\tau_n/h)+s/\lambda_n\Big]\Big\} \\
= \quad & \Big\{Y^* \in \Big(\gamma(1-\tau_n/h)+X'\beta(1-\tau_n/h), \gamma(1-\tau_n/h)+X'\beta(1-\tau_n/h)+s/\lambda_n\Big], D=1\Big\} \\
\subset \quad & \Big\{Y^* \in \Big(\gamma(1-\tau_n/h)+X'\beta(1-\tau_n/h), \gamma(1-\tau_n/h)+X'\beta(1-\tau_n/h)+s/\lambda_n\Big]\Big\}.
\end{aligned}
$$

Taking conditional expectations, this implies that for any $|s| \leq |z_1 + X'z_2|$ and n sufficiently large,

$$
\begin{aligned}
& \frac{|s|}{\lambda_n} f_{-\breve{U}|X}\left[Q_{-U|X}(\tau_n/h) + V_s|x\right] \\
\leq \quad & \left| F_{-U|X}\left(Q_{-U|X}(\tau_n/h|x) + \frac{s}{\lambda_n}|x\right) - F_{-U}\left(Q_{-U|X}(\tau_n/h|x)|x\right)\right|.
\end{aligned}
$$

By the mean value theorem,

$$
\begin{aligned}
& \left| F_{-U|X}\left(Q_{-U|X}(\tau_n/h|x) + \frac{s}{\lambda_n}|x\right) - F_{-U}\left(Q_{-U|X}(\tau_n/h|x)|x\right)\right| \\
= \quad & \frac{|s|}{\lambda_n} f_{-U|X}\left(Q_{-U|X}(\tau_n/h|x) + V_s'|x\right), \qquad (5.18)
\end{aligned}
$$

where $V_s' \in [0, s/\lambda_n]$. Because s is bounded for all $|s| \leq |z_1 + x'z_2|$ and all $x \in \text{Supp}(X)$, $|V_s'| \leq K/\lambda_n$. Now, by Lemma 5.2, there exists a constant $c_1 > 0$ independent of x, such that for n sufficiently large,

$$
Q'_{-\eta}(\tau_n/h) \leq \frac{c_1 h}{\tau_n}\Bigg(Q_{-\eta}(e\tau_n/h) - Q_{-\eta}(\tau_n/h)\Bigg).
$$

By (5.2), for n sufficiently large, there exists a constant M' independent of n and x such that

$$
Q_{-U|X}(e\tau_n/h|x) - Q_{-U|X}(\tau_n/h|x) \geq M'\Bigg(Q_{-\eta}(e\tau_n/h) - Q_{-\eta}(\tau/h)\Bigg).
$$

Combining the above two inequalities, we have, for any $x \in \text{Supp}(X)$,

$$
\begin{aligned}
\lambda_n\left[Q_{-U|X}(e\tau_n/h|x) - Q_{-U|X}(\tau_n/h|x)\right] \quad &= \quad \frac{\sqrt{n}hQ_{-U|X}(e\tau_n/h|x) - Q_{-U|X}(\tau_n/h|x)}{\sqrt{\tau_n}Q'_{-\eta}(\tau_n/h)} \\
&\geq \quad \sqrt{n\tau_n}\frac{Q_{-U|X}(e\tau_n/h|x) - Q_{-U|X}(\tau_n/h|x)}{c_1\Bigg(Q_{-\eta}(e\tau_n/h) - Q_{-\eta}(\tau_n/h)\Bigg)} \\
&\geq \quad \frac{M'\sqrt{n\tau_n}}{c_1} \\
&\to \quad \infty.
\end{aligned}
$$

Hence, for n sufficiently large,

$$
Q_{-U|X}(e\tau_n/h|X) \geq Q_{-U|X}(\tau_n/h|X) + \frac{K}{\lambda_n} \geq Q_{-U|X}(\tau_n/h|X) + V_s'.
$$

Plugging this inequality in (5.18) and using the monotonicity of $f_{-U|X}$, we obtain

$$f_{-\breve{U}|X}\left[Q_{-U|X}(\tau_n/h|X)+V_s|X\right]\leq f_{-U|X}\left(Q_{-U|X}(e\tau_n/h|X)|X\right).$$

Then by (5.4) in Lemma 5.2, we get

$$|U_n(s)|\leq |s|M\mathbb{1}\{|s|\leq |z_1+X'z_2|\}\frac{nf_{-\eta}\left(Q_{-\eta}(e\tau_n/h)\right)}{\lambda_n\sqrt{\tau_n n}}.$$

Finally, by Lemma 5.2 and (5.8),

$$\begin{aligned}\frac{nf_{-\eta}\left(Q_{-\eta}(e\tau_n/h)\right)}{\lambda_n\sqrt{\tau_n n}} &= \frac{f_{-\eta}\left(Q_{-\eta}(e\tau_n/h)\right)}{hf_{-\eta}\left(Q_{-\eta}(\tau_n/h)\right)}\\ &\sim \frac{Q_{-\eta}(e\tau_n/h)-Q_{-\eta}(\tau_n/h)}{h[Q_{-\eta}(e^2\tau_n/h)-Q_{-\eta}(e\tau_n/h)]}\\ &\to \frac{1}{h}.\end{aligned}$$

Since the left-hand side admits a limit, it is bounded by a constant $K>0$. Hence, we finally have, for n sufficiently large, $|U_n(s)|\leq U(s)$ with

$$U(s)=KM|s|\mathbb{1}\{|s|\leq |z_1+X'z_2|\}$$

which satisfies (5.17). Then, by Fubini's theorem, $E\left(\int_{-\infty}^{+\infty}U_n(s)ds\right)=\int_{-\infty}^{+\infty}E\left(U_n(s)\right)ds$. Second, since $E(U(s))<+\infty$, the sequence $(U_n(s))_n$ is asymptotically uniformly integrable. Together with (5.14), this implies, by Theorem 2.20 in van der Vaart (2000), that $E(U_n(s))\to E[m(s,z_1+X'z_2)s/H(X_{-1})]$. As a result,

$$E\left[\Lambda_n(z,\tau_n)\right]\to E\left[\frac{1}{H(X_{-1})}\int_0^{z_1+X'z_2}sds\right]=\frac{1}{2}z'\mathcal{Q}_Hz.$$

Finally, we turn to the bootstrap counterpart $\Lambda_n^*(z,\tau)$ of $\Lambda_n(z,\tau)$, which satisfies, for all $z=(z_1,z_2')'\in\mathbb{R}\times\mathbb{R}^d$,

$$\begin{aligned}\Lambda_n^*(z,\tau)=\frac{\lambda_n}{\sqrt{\tau n}}\sum_{i=1}^n w_{n,i}\int_0^{(z_1+X_i'z_2)/\lambda_n}&\Big[\mathbb{1}\{-Y_i+\gamma(1-\tau/h)+X_i'\beta(1-\tau/h)\leq s\}\\ &-\mathbb{1}\{-Y_i+\gamma(1-\tau/h)+X_i'\beta(1-\tau/h)\leq 0\}\Big]ds.\end{aligned}$$

First,

$$
\begin{aligned}
E\left[\Lambda_n^*(z,\tau)\right] &= E\left[E(\Lambda_n^*(z,\tau)|\{w_{n,i}\}_{i=1}^n)\right] \\
&= E\Bigg\{\sum_{i=1}^n w_{n,i} E\Bigg[\frac{\lambda_n}{\sqrt{\tau n}}\int_0^{(z_1+X_i'z_2)/\lambda_n}\Bigg[\mathbb{1}\{\gamma(1-\tau/h)+X_i'\beta(1-\tau/h)\le s+Y_i\} \\
&\quad -\mathbb{1}\{\gamma(1-\tau/h)+X_i'\beta(1-\tau/h)\le Y_i\}\Bigg]ds\Bigg]\Bigg\} \\
&= \frac{1}{n}E\left\{\sum_{i=1}^n w_{n,i} E\left[\Lambda_n(z,\tau)\right]\right\} \\
&= E\left[\Lambda_n(z,\tau)\right] \\
&\to \frac{1}{2} z Q_H' z.
\end{aligned}
$$

Next, we show that $V\left(\Lambda_n^*(z,\tau)\right)\to 0$. To see this, we note that by variance decomposition,

$$
\begin{aligned}
V\left(\Lambda_n^*(z,\tau)\right) &= E\left[V\left(\Lambda_n^*(z,\tau)|\{w_{n,i}\}_{i=1}^n\right)\right] + V\left[E\left(\Lambda_n^*(z,\tau)|\{w_{n,i}\}_{i=1}^n\right)\right] \\
&= E\Bigg[\sum_{i=1}^n w_{n,i}^2 \frac{\lambda_n^2}{\tau n} V\Bigg(\int_0^{(z_1+X_i'z_2)/\lambda_n}\Bigg[\mathbb{1}\{\gamma(1-\tau/h)+X_i'\beta(1-\tau/h)\le s+Y_i\} \\
&\quad -\mathbb{1}\{\gamma(1-\tau/h)+X_i'\beta(1-\tau/h)\le Y_i\}\Bigg]ds\Bigg)\Bigg] + V\left[\frac{1}{n}\sum_{i=1}^n w_{n,i} E\left(\Lambda_n(z,\tau)\right)\right] \\
&= (1-\frac{1}{n})V\left(\Lambda_n(z,\tau)\right) + 0 \\
&\to 0.
\end{aligned}
$$

□

References

Chernozhukov, V. 2005. Extremal Quantile Regression. *The Annals of Statistics*, **33**(2), pp. 806–839.

D'Haultfoeuille, X., & Maurel, A. 2013. Another Look at the Identification at Infinity of Sample Selection Models. *Econometric Theory*, **29**(1), 213–224.

D'Haultfoeuille, X., Maurel, A., & Zhang, Y. 2014. Extremal Quantile Regressions for Selection Models and the Black-White Wage Gap. *NBER Working Paper No. 20257*.

D'Haultfoeuille, X., Maurel, A., & Zhang, Y. 2017. Extremal Quantile Regressions for Selection Models and the Black-White Wage Gap. Working Paper.

Heckman, J.J., & Vytlacil, E.J. 2007. Econometric evaluation of social programs, Part II. *In:* Heckman, J.J., & Leamer, E.E. (eds), *Handbook of Econometrics*, vol. 6B. Elsevier.

Pollard, D. 1991. Asymptotics for Least Absolute Deviation Regression Estimators. *Econometric Theory*, **7**(2), pp. 186–199.

Resnick, S. 1987. *Extreme values, regular variation, and point processes.* New York: Springer-Verlag.

van der Vaart, A. W. 2000. *Asymptotic Statistics.* Cambridge Series in Statistical and Probabilistic Mathematics.