Change-in-Change Asymptotics

(preliminary and incomplete)

Xavier D'Haultfœuille* Jérémy L'Hour† Martin Mugnier‡

January 4, 2021

Abstract

XXX

*CREST-ENSAE, xavier.dhaultfoeuille@ensae.fr.
†CREST-ENSAE, jeremy.l.hour@ensae.fr
‡CREST-ENSAE, martin.mugnier@ensae.fr.

1 Asymptotic Results (Observed Rank)

We consider $\theta_0 = E[F_Y^{-1}(U)]$ and the estimator

$$\widehat{\theta}_n = \frac{1}{n}\sum_{j=1}^{n} \widehat{F}_Y^{-1}(U_j), \tag{1}$$

where $\widehat{F}_Y^{-1}$ is the empirical quantile function obtained from $(Y_i)_{i=1,\dots,n}$.

Assumption 1 (Sampling) *Assume: (i) $(Y_i)_{i=1,\dots,n}$ are independent draws from the distribution F_Y and $(U_j)_{j=1,\dots,n}$ are independent draws from the distribution F_U. (ii) U_j is independent of Y_i for any i and j. (iii) $Supp(U_i) \subseteq [0,1]$. (iv) F_Y and F_U are absolutely continuous with respect to the Lebesgue measure.*

Notice that (1) is a L-statistics (Shorack and Wellner, 1986, Chapter 19), *i.e.* $\widehat{\theta}_n = n^{-1}\sum_{i=1}^{n} c_{ni} Y_{(i)}$, where $Y_{(1)} < \dots < Y_{(n)}$ is the order statistics and $c_{ni} = \#\{U_j : U_j \in ((i-1)/n, i/n]\}$. However, contrary to the textbook case, the weights c_{ni} are random variables,

$$(c_{n1}, \dots, c_{nn}) \sim \mathcal{M}(n, F_U(1/n) - F_U(0/n), \dots, F_U(n/n) - F_U((n-1)/n)).$$

In order to study its asymptotic behavior, let us decompose it into two parts, each only depending at the first order on the random sample $(Y_i)_{i=1,\dots,n}$ or $(U_i)_{i=1,\dots,n}$ but not both. Notice that θ_0 can be written as

$$\theta_0 = \int_0^1 F_Y^{-1}\, dF_U.$$

Let $\xi_i := F_Y(Y_i) \sim \mathcal{U}[0,1]$, $\mathbb{G}_n$ denote the empirical cumulative distribution function obtained from $(\xi_i)_{i=1,\dots,n}$. Also, for any function $F : [0,1] \longrightarrow [0,1]$, we let F^{-1} denote its left-continuous generalized inverse:

$$F^{-1}(y) := \inf\{x \in [0,1] \,:\, y \le F(x)\} \qquad \forall y \in [0,1].$$

Thus, $\mathbb{G}_n^{-1}(\tau)$ is the usual empirical quantile of order τ.

The estimator (1) can be expressed as[1]

$$\widehat{\theta}_n = \int_0^1 F_Y^{-1} \circ \mathbb{G}_n^{-1} \, d\widehat{F}_U,$$

where $\widehat{F}_U$ is the empirical cumulative distribution function obtained from $(U_i)_{i=1,\dots,n}$. We show that

$$\sqrt{n}(\widehat{\theta}_n - \theta_0) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \varepsilon_i + \frac{1}{\sqrt{n}} \sum_{i=1}^n \zeta_i + o_P(1),$$

where $\varepsilon_i = -\int_0^1 [\mathbb{1}\{U_i \leq t\} - F_U(t)] \, dF_Y^{-1}(t)$ and $\zeta_i = -\int_0^1 [\mathbb{1}\{F_Y(Y_i) \leq t\} - t] f_U(t) \, dF_Y^{-1}(t)$ are independent, square-integrable, random variables, allowing to apply a standard CLT.

Assumption 2 (Regularity Conditions on Densities)

(i) There exist $b_1, b_2 > 0$ and $C_U > 0$ such that for all $t \in (0,1)$:

$$f_U(t) \leq C_U t^{-b_1} (1-t)^{-b_2}.$$

(ii) There exist $d_1, d_2 > 0$ and $C_Y > 0$ such that for all $t \in (0,1)$:

$$|F_Y^{-1}(t)| \leq C_Y t^{-d_1} (1-t)^{-d_2}.$$

(iii) $b_1 + d_1 < 1/2$ and $b_2 + d_2 < 1/2$.

Point 2 of Assumption 2 holds under the following moment condition on Y.

Lemma 1 (Lower-Level Conditions on Y) *Assume $E[|Y|^p] < \infty$ for $p > 1$, then Assumption 2(ii) is verified with $d_1 = d_2 = 1/p$.*

Theorem 1 (Asymptotic Normality) *Under Assumptions 1 and 2, as $n \to \infty$,*

$$\sqrt{n}(\widehat{\theta}_n - \theta_0) \xrightarrow{d} \mathcal{N}(0, \sigma^2),$$

with

$$\sigma^2 = \int_0^1 \int_0^1 [F_U(s \wedge t) - F_U(s) F_U(t) + [s \wedge t - st] f_U(s) f_U(t)] \; dF_Y^{-1}(s) \, dF_Y^{-1}(t).$$

All proofs are gathered in the appendix.

[1] Letting $\lceil a \rceil$ be the least integer greater than or equal to a, notice that $\mathbb{G}_n^{-1}(x) = \xi_{(\lceil nx \rceil)} = F_Y(Y_{(\lceil nx \rceil)})$ and F_Y absolutely continuous with respect to the Lebesgue measure imply $F_Y^{-1} \circ \mathbb{G}_n^{-1}(x) = F_Y^{-1} \circ F_Y(Y_{(\lceil nx \rceil)}) = Y_{(\lceil nx \rceil)} = \widehat{F}_Y^{-1}(x)$.

2 Asymptotic Results (Estimated Rank)

In many applications, the random variable U_i has a known form $U_i = F_Z(X_i)$ for some observed random variable X_i and a unknown cumulative distribution function F_Z. As a consequence, we do not directly observe the random variables $(U_i)_{i=1,\dots,n}$, instead we are left with estimated quantities $(\widehat{U}_i)_{i=1,\dots,n}$. In these cases, $\widehat{U}_i$ is the image of some observed random variable X_i through an empirical cumulative distribution function that comes from another independent sample $(Z_i)_{i=1,\dots,n}$, *i.e.* $\widehat{U}_i = \widehat{F}_Z(X_i)$.

Assumption 3 (Pooled Independent Samples) *(i)* $(Y_i)_{i=1,\dots,n}$ *(resp.* $(Z_i)_{i=1,\dots,n}$ *and* $(X_i)_{i=1,\dots,n}$*) are independent draws from the distribution* F_Y *(resp.* F_Z *and* F_X*). (ii)* (Y_i, Z_j, X_k) *are mutually independent for any value of* i,j *and* k*. (iii)* F_Y, F_Z *and* F_X *are absolutely continuous with respect to the Lebesgue measure.*

Notice that U is distributed with cdf $F_U = F_X \circ F_Z^{-1}$ and density

$$f_U(t) = \frac{f_X(F_Z^{-1}(t))}{f_Z(F_Z^{-1}(t))} \mathbb{1}\{t \in [0,1]\}.$$

Notice that $\widehat{U}_i = \mathbb{H}_n(U_i)$, where $\mathbb{H}_n$ is the empirical cdf obtained from $(F_Z(Z_j))_{j=1,\dots,n}$ with $F_Z(Z_j) \sim \mathcal{U}[0,1]$. We consider the estimator:

$$\check{\theta}_n := \frac{1}{n}\sum_{j=1}^{n} \widehat{F}_Y^{-1}(\widehat{U}_j) = \frac{1}{n}\sum_{j=1}^{n} \widehat{F}_Y^{-1}(\mathbb{H}_n(U_j)).$$

Note: we can also use a smooothed version of $\widehat{F}_Z$. Following Shorack and Wellner (1986), we let $\widehat{F}_Z(Z_{(i)}) = i/(n+1)$ for $i = 1...n$, $\widehat{F}_Z(\cdot)$ linear between $Z_{(i)}$ and $Z_{(i+1)}$ for $i < n$. For $z < Z_{(1)}$ and $z > Z_{(n)}$, we extrapolate linearly until reaching 0 and 1 respectively. One can show that this extrapolation is equivalent to defining $Z_{(0)} = 2Z_{(1)} - Z_{(2)}$ and $Z_{(n+1)} = 2Z_{(n)} - Z_{(n-1)}$ instead of 0 and 1 as in Shorack and Wellner (1986). With this estimator, $\mathbb{H}_n(\cdot)$ is defined as

$$\mathbb{H}_n(u) = \frac{1}{n+1}\left(i + \frac{F_Z^{-1}(u) - Z_{(i)}}{Z_{(i+1)} - Z_{(i)}}\right) \quad \text{if } Z_{(i)} \leq F_Z^{-1}(u) \leq Z_{(i+1)}$$

for $i = 0, ..., n$. Finally, $\mathbb{H}_n$ is constant and equal to 0 on $[0, F_Z(Z_{(0)})]$ and constant, equal to 1 on $[F_Z(Z_{(n+1)}), 1]$. (these two sets may or may not be empty). Similarly as before, $\check{\theta}_n$ can be expressed as:

$$\check{\theta}_n = \int_0^1 F_Y^{-1} \circ \mathbb{G}_n^{-1} \circ \mathbb{H}_n \, d\widehat{F}_U.$$

We show that the estimator can be decomposed into three independent parts:

$$\sqrt{n}(\check{\theta}_n - \theta_0) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \varepsilon_i + \frac{1}{\sqrt{n}} \sum_{i=1}^n \zeta_i + \frac{1}{\sqrt{n}} \sum_{i=1}^n \varphi_i + o_P(1),$$

where $\varepsilon_i = -\int_0^1 [\mathbb{1}\{F_Z(X_i) \leq t\} - F_X(F_Z^{-1}(t))] \, dF_Y^{-1}(t)$, $\zeta_i = -\int_0^1 [\mathbb{1}\{F_Y(Y_i) \leq t\} - t] f_U(t) \, dF_Y^{-1}(t)$ and $\varphi_i := \int_0^1 [\mathbb{1}\{F_Z(Z_i) \leq t\} - t] \, f_U(t) \, dF_Y^{-1}(t)$ are independent, square-integrable, random variables, allowing to apply a standard CLT.

Theorem 2 (Asymptotic Normality) *Under Assumptions 2-3, as $n \to \infty$,*

$$\sqrt{n}(\check{\theta}_n - \theta_0) \xrightarrow{d} \mathcal{N}(0, \sigma^2),$$

with

$$\sigma^2 = \int_0^1 \int_0^1 \left[F_X(F_Z^{-1}(s \wedge t)) - F_X(F_Z^{-1}(s)) F_X(F_Z^{-1}(t)) \right] \, dF_Y^{-1}(s) \, dF_Y^{-1}(t)$$
$$+ 2 \int_0^1 \int_0^1 [s \wedge t - st] \frac{f_X(F_Z^{-1}(s))}{f_Z(F_Z^{-1}(s))} \frac{f_X(F_Z^{-1}(t))}{f_Z(F_Z^{-1}(t))} \, dF_Y^{-1}(s) \, dF_Y^{-1}(t).$$

The proof is long and technical. Describe the main steps and ingredients:

1. Decompose into three terms. Two are the same as in Theorem 1, the third is new. We decompose it further into several terms: remainder terms plus a L-statistic.

2. For some remainder term, similar technique as in Theorem 1 but a bit more complex. Use in particular the fact that (i) order statistic of uniforms and uniform spacings are distributed as beta; (ii) mean absolute deviation of beta distributions.

3. For another remainder term, use convergence of the supremum of the weighted empirical quantile process (see in particular Csorgo et al., 1986, Corollary 4.3.1).

4. For the $L-$statistic, results in Shorack and Wellner (1986) do not apply here. Instead, we use the necessary and sufficient condition for its asymptotic normality in Hecker (1976).

3 Application to Change-in-Change

Faire une revue de litérature sur le Change-in-Change et sur d'autres possibles estimateurs rentrant dans notre cadre.

We study the Change-in-Change estimator of Athey and Imbens (2006). Let $Y_{gt,i}$ the outcome at time t for individual i in group g. The Change-in-Change estimand of the Average Treatment Effect (ATE) is

$$\tau^{CIC} = E[Y_{11}] - E[F_{Y_{01}}^{-1}(F_{Y_{00}}(Y_{10}))].$$

The idea is to estimate the counterfactual by averaging the quantile the treated population would have had, had they been in the untreated group at the initial date and kept the same rank in the second period. In our more simplistic framework, we have $U = F_{Y,00}(Y_{10})$, with $Y_{10} \sim F_{Y,10}$. Assume that all the cdf are absolutely continuous with respect to the Lebesgue measure, then $F_U = F_{Y,10} \circ F_{Y,00}^{-1}$ and its density is:

$$f_U(t) = \frac{f_{Y,10}(F_{Y,00}^{-1}(t))}{f_{Y,00}(F_{Y,00}^{-1}(t))} \mathbb{1}\{t \in [0,1]\}.$$

Clearly, if the outcome distribution is the same for the treated and the untreated at the initial date ($f_{Y,10} = f_{Y,00}$) then U is uniformly distributed. Athey and Imbens (2006) require the density of Y_{gt} for each g and t to be bounded from below and bounded from above on a compact support (see Assumption 5 therein). This assumption yields that f_U will also be bounded. In general, f_U will be bounded if and only if the ratio $f_{Y,10}/f_{Y,00}$ is bounded, which may not be the case for many usual distributions typically encountered with economic data. Our method of proof does not require any constant bound on f_U, thus extending the cases where the Change-in-Change is a relevant tool.

Examples: In the following examples we study the tail behavior of f_U with respect to the underlying distribution of the treated and untreated outcomes at the initial date. We show that under many standard distributions, Assumption 2 (i) is verified.

1. Exponential Distribution. Assume that $Y_{g0} \sim \mathcal{E}(\lambda_g)$, in that case

$$f_U(t) = \frac{\lambda_1}{\lambda_0}(1-t)^{\lambda_1/\lambda_0 - 1}\mathbb{1}\{t \in [0,1]\},$$

and $U \sim \text{Beta}(1, \lambda_1/\lambda_0)$.

2. Pareto Distribution. Assume that Y_{g0} has cdf $1 - (\beta_g/x)^{\alpha_g}$, in that case
$$f_U(t) = \frac{\alpha_1}{\alpha_0}\left(\frac{\beta_1}{\beta_0}\right)^{\alpha_1}(1-t)^{\alpha_1/\alpha_0 - 1}\mathbb{1}\{1 - (\beta_0/\beta_1)^{\alpha_0} < t < 1\},$$
which is a "truncated" Beta distribution.

3. Normal Distribution. Assume that $Y_{g0} \sim \mathcal{N}(\mu_g, \sigma_g^2)$, in that case
$$f_U(t) = \frac{\sigma_0}{\sigma_1}\exp\left[-\frac{1}{2\sigma_1^2}\Big((\sigma_0+\sigma_1)\Phi^{-1}(t) + \mu_0 - \mu_1\Big)\Big((\sigma_0-\sigma_1)\Phi^{-1}(t) + \mu_0 - \mu_1\Big)\right]\mathbb{1}\{t \in [0,1]\}.$$
Consider the special case: $\mu_1 = \mu_0$ and $\sigma_1 > \sigma_0$. For $t \in (1/2, 1)$, using the inequality $\Phi^{-1}(t) \leq \sqrt{-2\ln(2(1-t))}$ yields $f_U(t) \leq (\sigma_0/\sigma_1)(2(1-t))^{\sigma_0^2/\sigma_1^2 - 1}$. Symmetrically, for $t \in (0, 1/2)$, $\Phi^{-1}(t) \geq -\sqrt{-2\ln(2t)}$ yields $f_U(t) \leq (\sigma_0/\sigma_1)(2t)^{\sigma_0^2/\sigma_1^2 - 1}$.

4. Logistic Distribution. Assume that Y_{g0} has cdf $1/(1 + \exp(-(t-\mu_g)/\beta_g))$, in that case
$$f_U(t) = \frac{\beta_0}{\beta_1}\frac{(1/t - 1)^{\beta_0/\beta_1 - 1}\, e^{(\mu_1-\mu_0)/\beta_1}}{t^2\left(1 + (1/t-1)^{\beta_0/\beta_1}\, e^{(\mu_1-\mu_0)/\beta_1}\right)^2}\mathbb{1}\{t \in [0,1]\}.$$

5. Gumbel Distribution. Assume that Y_{g0} has cdf $e^{-(t-\mu_g)/\beta_g}\exp\left(-e^{-(t-\mu_g)/\beta_g}\right)/\beta_g$, in that case
$$f_U(t) = \frac{\beta_0}{\beta_1}e^{(\mu_1-\mu_0)/\beta_1}\ln(t)^{\beta_0/\beta_1 - 1}\exp\left(-\ln(t)^{\beta_0/\beta_1}e^{(\mu_1-\mu_0)/\beta_1}\right)\mathbb{1}\{t \in [0,1]\}.$$
If $\beta_1 = \beta_0 = 1$ and $\mu_0 > \mu_1$, $U \sim \text{Beta}(1 - e^{\mu_1 - \mu_0}, 1)$. □

We never observe U_i directly, instead F_{00} is replaced by its empirical counterpart $\widehat{F}_{00}$ and we have $\widehat{U}_i = \widehat{F}_{00}(Y_{i,10})$. Notice that:
$$\begin{aligned}\widehat{U}_i &= \frac{1}{n}\sum_{j=1}^{n}\mathbb{1}\{Y_{00,j} \leq Y_{10,i}\}\\ &= \frac{1}{n}\sum_{j=1}^{n}\mathbb{1}\{F_{00}(Y_{00,j}) \leq U_i\}\\ &= \mathbb{H}_n(U_i),\end{aligned}$$
where $\mathbb{H}_n$ is the empirical cdf of $F_{00}(Y_{00,j}) \sim \mathcal{U}[0,1]$.

4 Theorem 2 generalizes Theorem 5.1 in Athey and Imbens (2006).

By several changes of variables, one may easily verify that

$$\zeta_i = -\int_{F_Z^{-1}(0)}^{F_Z^{-1}(1)} \left[\mathbb{1}\{F_Y(Y_i) \le F_Z(x)\} - F_Z(x)\right] \frac{1}{f_Y\left(F_Y^{-1}(F_Z(x))\right)} \times f_X(x)\, dx, \tag{2}$$

$$\varphi_i = \int_{F_Z^{-1}(0)}^{F_Z^{-1}(1)} \left[\mathbb{1}\{Z_i \le x\} - F_Z(x)\right] \frac{1}{f_Y\left(F_Y^{-1}(F_Z(x))\right)} \times f_X(x)\, dx. \tag{3}$$

Athey and Imbens (2006) impose $Supp(Y_{10}) \subset Supp(Y_{00})$, which is equivalent to $Supp(X) \subset Supp(Z)$ in our notation. Under this assumption, we have

$$E(\zeta_i^2) = V^q,$$
$$E(\varphi_i^2) = V^p,$$

where V^q and V^p are defined in Theorem 5.1 in Athey and Imbens (2006). Now, it remains to analyze the last variance term V^r. We have:

$$\begin{aligned} V^r &:= V\left(F_{Y_{01}}^{-1}\left(F_{Y_{00}}(Y_{10})\right)\right) \\ &= V\left(F_Y^{-1}\left(F_Z(X)\right)\right) \\ &= \int_{Supp(X)} \left(F_Y^{-1}\left(F_Z\left(x\right)\right)\right)^2 f_X(x)\, dx - \left(\underbrace{\int_{Supp(X)} F_Y^{-1}\left(F_Z\left(x\right)\right) f_X(x)\, dx}_{=:\theta_0}\right)^2. \end{aligned}$$

Also, by an integration by part

$$\begin{aligned} E(\varepsilon_i^2) &:= \int_0^1\int_0^1 \Big[F_X(F_Z^{-1}(s\wedge t)) - F_X(F_Z^{-1}(s))F_X(F_Z^{-1}(t))\Big]\, dF_Y^{-1}(s)\, dF_Y^{-1}(t) \\ &= \int_0^1 \Big\{ \Big[F_Y^{-1}(s)\Big(F_X(F_Z^{-1}(s\wedge t)) - F_X(F_Z^{-1}(s))F_X(F_Z^{-1}(t))\Big)\Big]_{s=0}^{s=1} \\ &\quad - \int_0^1 F_Y^{-1}(s)\Big[\frac{f_X(F_Z^{-1}(s\wedge t))}{f_Z(F_Z^{-1}(s\wedge t))}\mathbb{1}\{s\le t\} - \frac{f_X(F_Z^{-1}(s))F_X(F_Z^{-1}(t))}{f_Z(F_Z^{-1}(s))}\Big]\, ds\Big\}\, dF_Y^{-1}(t) \\ &= -\int_0^1 \left(\int_0^1 F_Y^{-1}(s)\frac{f_X(F_Z^{-1}(s\wedge t))}{f_Z(F_Z^{-1}(s\wedge t))}\mathbb{1}\{s\le t\}\, ds\right) dF_Y^{-1}(t) \\ &\quad + \int_0^1 \left(\int_0^1 F_Y^{-1}(s)\frac{f_X(F_Z^{-1}(s))F_X(F_Z^{-1}(t))}{f_Z(F_Z^{-1}(s))}\, ds\right) dF_Y^{-1}(t). \end{aligned} \tag{4}$$

The third equality follows because $Supp(X) \subset Supp(Y)$ implies

$$[F_Y^{-1}(s)\Big(F_X(F_Z^{-1}(s \wedge t)) - F_X(F_Z^{-1}(s))F_X(F_Z^{-1}(t))\Big)\Big]_{s=0}^{s=1} = 0.$$

Focus on the second term in (4). By the change of variable $x = F_Z^{-1}(s)$, we obtain

$$\begin{aligned}
&\int_0^1 \left(\int_0^1 F_Y^{-1}(s)\frac{f_X(F_Z^{-1}(s))F_X(F_Z^{-1}(t))}{f_Z(F_Z^{-1}(s))}\,ds\right)\,dF_Y^{-1}(t)\\
&\quad = \left(\int_0^1 F_Y^{-1}(s)\frac{f_X(F_Z^{-1}(s))}{f_Z(F_Z^{-1}(s))}\,ds\right)\left(\int_0^1 F_X(F_Z^{-1}(t))\,dF_Y^{-1}(t)\right)\\
&= \left(\int_{F_Z^{-1}(0)}^{F_Z^{-1}(1)} F_Y^{-1}(F_Z(x))f_X(x)\,dx\right)\left(\int_0^1 F_X(F_Z^{-1}(t))\,dF_Y^{-1}(t)\right)\\
&= \left(\int_{F_Z^{-1}(0)}^{F_Z^{-1}(1)} F_Y^{-1}(F_Z(x))f_X(x)\,dx\right)\left([F_Y^{-1}(t)F_X(F_Z^{-1}(t))]_{t=0}^{t=1} - \int_{F_Z^{-1}(0)}^{F_Z^{-1}(1)} F_Y^{-1}(F_Z(x))f_X(x)\,dx\right)\\
&= -\left(\int_{F_Z^{-1}(0)}^{F_Z^{-1}(1)} F_Y^{-1}(F_Z(x))f_X(x)\,dx\right)^2 + \left(\int_{F_Z^{-1}(0)}^{F_Z^{-1}(1)} F_Y^{-1}(F_Z(x))f_X(x)\,dx\right)F_Y^{-1}(1)\\
&= -\theta_0^2 + \theta_0 F_Y^{-1}(1),
\end{aligned} \tag{5}$$

where we obtain the third equality by an integration by part followed by the same change of variable than before, and the last two equalities holds because $Supp(X) \subset Supp(Z)$. Now, focus on the first term in (4). By the same change of variable again, we have

$$\begin{aligned}
&-\int_0^1 \left(\int_0^1 F_Y^{-1}(s)\frac{f_X(F_Z^{-1}(s \wedge t))}{f_Z(F_Z^{-1}(s \wedge t))}\mathbb{1}\{s \le t\}\,ds\right)\,dF_Y^{-1}(t)\\
&\quad = -\int_0^1 \left(\int_{F_Z^{-1}(0)}^{F_Z^{-1}(t)} F_Y^{-1}(F_Z(x))f_X(x)\,dx\right)\,dF_Y^{-1}(t).
\end{aligned} \tag{6}$$

By Leibniz's derivation rule for integrals, we have

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{dt}}\left(\int_{F_Z^{-1}(0)}^{F_Z^{-1}(t)} F_Y^{-1}(F_Z(x))f_X(x)\,dx\right) &= \frac{1}{f_Z(F_Z^{-1}(t))}F_Y^{-1}(F_Z(F_Z^{-1}(t)))f_X(F_Z^{-1}(t))\\
&= \frac{f_X(F_Z^{-1}(t))}{f_Z(F_Z^{-1}(t))}F_Y^{-1}(t).
\end{aligned}$$

Hence, an integration by part of (6) yields

$$\begin{aligned}
&-\int_0^1 \left(\int_0^1 F_Y^{-1}(s)\frac{f_X(F_Z^{-1}(s \wedge t))}{f_Z(F_Z^{-1}(s \wedge t))}\mathbb{1}\{s \le t\}\,ds\right)\,dF_Y^{-1}(t)\\
&= -\theta_0 F_Y^{-1}(1) + \int_0^1 F_Y^{-1}(t)^2\frac{f_X(F_Z^{-1}(t))}{f_Z(F_Z^{-1}(t))}\,dt\\
&= -\theta_0 F_Y^{-1}(1) + \int_{F_Z^{-1}(0)}^{F_Z^{-1}(1)} \left(F_Y^{-1}(F_Z(x))\right)^2 f_X(x)\,dx,
\end{aligned} \tag{7}$$

where we used the change of variable $x = F_Z^{-1}(t)$ to obtain the last equality. Now, by combining (5) and (7) and noting that

$$\int_{F_Z^{-1}(0)}^{F_Z^{-1}(1)} \left(F_Y^{-1}(F_Z(x))\right)^2 f_X(x)\, dx = \int_{Supp(X)} \left(F_Y^{-1}(F_Z(x))\right)^2 f_X(x)\, dx,$$

we obtain

$$\begin{aligned}
E(\varepsilon_i^2) &= -F_Y^{-1}(1)\theta_0 + \int_{Supp(X)} \left(F_Y^{-1}(F_Z(x))\right)^2 f_X(x)\, dx - \theta_0^2 + \theta_0 F_Y^{-1}(1) \\
&= \int_{Supp(X)} \left(F_Y^{-1}(F_Z(x))\right)^2 f_X(x)\, dx - \left(\int_{Supp(X)} F_Y^{-1}(F_Z(x)) f_X(x)\, dx\right)^2 \\
&= V^r.
\end{aligned}$$

Notice that we can also write ζ_i, φ_i and ε_i as:

$$\begin{aligned}
\zeta_i &= -E_U\left[(\mathbb{1}\{F_Y(Y_i) \le U\} - U)\frac{1}{f_Y\left(F_Y^{-1}(U)\right)}\right] \\
\varphi_i &= E_U\left[(\mathbb{1}\{F_Z(Z_i) \le U\} - U)\frac{1}{f_Y\left(F_Y^{-1}(U)\right)}\right], \\
{}''\varepsilon_i'' &= -(F_Y^{-1}(U_i) - E_U\left[F_Y^{-1}(U)\right]) = \theta_0 - F_Y^{-1}(U_i),
\end{aligned}$$

where the expectation is taken over the distribution of U only.

4.1 Variance estimation

Notice that the asymptotic variance in Theorem 2 can be rewritten as $E\left[\zeta_i^2 + \varphi_i^2 + \epsilon_i^2\right]$. So, we need to estimate the following three terms:

$$\begin{aligned}
\zeta_i &= -E_U\left[(\mathbb{1}\{F_Y(Y_i) \le U\} - U)\frac{1}{f_Y\left(F_Y^{-1}(U)\right)}\right] \\
\varphi_i &= E_U\left[(\mathbb{1}\{F_Z(Z_i) \le U\} - U)\frac{1}{f_Y\left(F_Y^{-1}(U)\right)}\right], \\
\epsilon_i &= -(F_Y^{-1}(U_i) - E_U\left[F_Y^{-1}(U)\right]) = \theta_0 - F_Y^{-1}(U_i),
\end{aligned}$$

The three terms above are straightforward to estimate except for their dependence on the density of Y. Non-parametric estimation of the density of Y requires the use of a kernel and a choice of bandwith. To overcome this difficulty, we instead draw inspiration from Lewbel and Schennach (2007) and notice that the function $x \to 1/f_Y(F_Y^{-1}(x))$ is the derivative of the quantile function of Y, $F_Y^{-1}(x)$. Let us define the empirical quantile function of Y, for $x \in (0,1)$:

$$\widehat{F}_Y^{-1}(x) := \inf \left\{ y \text{ s.t. } \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}\{Y_i \leq y\} \geq x \right\}.$$

Let us also use the estimated ranks as defined in Section 2: $\widehat{U}_i = \widehat{F}_Z(X_i)$, and denote $\widehat{U}_{(i)}$ the i-th value in the ordered sample $\widehat{U}_{(1)} \leq \widehat{U}_{(2)} \leq \ldots \leq \widehat{U}_{(n)}$. We can approximate the value of the function $x \to 1/f_Y(F_Y^{-1}(x))$ evaluated at $\left(\widehat{U}_{(i+1)} + \widehat{U}_{(i)}\right)/2$ by:

$$\frac{\widehat{F}_Y^{-1}(\widehat{U}_{(i+1)}) - \widehat{F}_Y^{-1}(\widehat{U}_{(i)})}{\widehat{U}_{(i+1)} - \widehat{U}_{(i)}}.$$

With that in mind, we propose the following estimators for the quantities above:

$$\begin{aligned}
\widehat{\zeta}_i &= -\frac{1}{n} \sum_{j=1}^{n-1} \left(\mathbb{1}\left\{ \frac{i+1}{n} \leq \frac{\widehat{U}_{(j+1)} + \widehat{U}_{(j)}}{2} \right\} - \frac{\widehat{U}_{(j+1)} + \widehat{U}_{(j)}}{2} \right) \frac{\widehat{F}_Y^{-1}(\widehat{U}_{(j+1)}) - \widehat{F}_Y^{-1}(\widehat{U}_{(j)})}{\widehat{U}_{(j+1)} - \widehat{U}_{(j)}}, \\
\widehat{\varphi}_i &= -\widehat{\zeta}_i, \\
\widehat{\epsilon}_i &= \check{\theta} - F_Y^{-1}(\widehat{U}_i).
\end{aligned}$$

Finally, the estimator for σ^2 is given by:

$$\widehat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} \widehat{\zeta}_i^2 + \widehat{\varphi}_i^2 + \widehat{\epsilon}_i^2.$$

(On a $\widehat{\varphi}_i = -\widehat{\zeta}_i$ dans le cas où l'on a autant de Z_i que de Y_i, ce qui n'est pas nécess-sairement vrai dans le cas général. Notons bien que le terme $\frac{i+1}{n}$ sert d'estimateur à, par exemple, $\widehat{F}_Z(Z_{(i)})$).

References

Athey, S. and Imbens, G. W. (2006). Identification and inference in nonlinear difference-in-differences models. *Econometrica*, 74(2):431–497.

Berend, D. and Kontorovich, A. (2013). A sharp estimate of the binomial mean absolute deviation with applications. *Statistics & Probability Letters*, 83(4):1254–1259.

Csorgo, M., Csorgo, S., Horváth, L., and Mason, D. M. (1986). Weighted empirical and quantile processes. *The Annals of Probability*, pages 31–85.

Falkner, N. and Teschl, G. (2012). On the substitution rule for lebesgue–stieltjes integrals. *Expositiones Mathematicae*, 30(4):412 – 418.

Hecker, H. (1976). A characterization of the asymptotic normality of linear combinations of order statistics from the uniform distribution. *Ann. Statist.*, 4(6):1244–1246.

Lewbel, A. and Schennach, S. M. (2007). A simple ordered data estimator for inverse density weighted expectations. *Journal of Econometrics*, 136(1):189–211.

Shorack, G. and Wellner, J. (1986). *Empirical Processes with Applications to Statistics*. John Wiley and Sons.

Van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes*. Springer Series in Statistics. Springer, New York, NY.

A Proofs of the main results

Below, we use "$\lesssim$" to indicate an inequality up to universal constant. In most cases below, this means a constant independent of x and n.

A.1 Proof of Lemma 1

Observe that $E[|Y|] < \infty$ implies $tS_Y(t) \to 0$ and $tF_Y(-t) \to 0$ as $t \to \infty$. Thus $E[|Y|^p] < \infty$ implies $t^p S_Y(t) \to 0$ and $t^p F_Y(-t) \to 0$ as $t \to \infty$. The convergence to 0 of $t^p S_Y(t)$ implies that there exists $C > 0$ and t_1 such that for all $t \geq t_1$,

$$|t|^p(1 - F_Y(t)) \leq C.$$

This implies that for all $u \geq F_Y(t_1)$, $|F_Y^{-1}(u)|^p(1-u) \leq C$ or, equivalently,

$$|F_Y^{-1}(u)| \leq C(1-u)^{-1/p}.$$

Hence, there exists $C_1 > 0$ such that for $u \geq F_Y(t_1)$,

$$|F_Y^{-1}(u)| \leq C_1[u(1-u)]^{-1/p}.$$

Using $t^p F_Y(-t) \to 0$ and a similar reasoning, there exists C_2 and $t_2 \leq t_1$ such that for all $u \leq t_2$, $|F_Y^{-1}(u)| \leq C_2[u(1-u)]^{-1/p}$. The result follows since $|F_Y^{-1}(u)[u(1-u)]^{1/p}|$ is bounded on $[t_2, t_1]$. $\square$

B Proof of Theorem 1

Consider the following decomposition:

$$\widehat{\theta}_n - \theta_0 = \underbrace{\int_0^1 F_Y^{-1} \circ \mathbb{G}_n^{-1} \, dF_U - \int_0^1 F_Y^{-1} \, dF_U}_{:=T_{1n}} + \underbrace{\int_0^1 F_Y^{-1} \circ \mathbb{G}_n^{-1} \, d\widehat{F}_U - \int_0^1 F_Y^{-1} \circ \mathbb{G}_n^{-1} \, dF_U}_{:=T_{2n}}.$$

The proof proceeds in three steps. In the first step, we prove that T_{1n} is linear up to a negligible remainder term. In the second step, we prove the same result for T_{2n}. The last step concludes.

First step: linearization of T_{1n}. By Lemma 2 followed by an integration by part,

$$\begin{aligned}\sqrt{n}T_{1n} &= \sqrt{n}\left[\int_0^1 F_Y^{-1}\, dF_U \circ \mathbb{G}_n - \int_0^1 F_Y^{-1}\, dF_U\right] \\ &= \sqrt{n}\left\{\int_{\xi_{(1)}}^{\xi_{(n)}} F_Y^{-1}\, d\left[F_U \circ \mathbb{G}_n - F_U\right] - \int_0^{\xi_{(1)}} F_Y^{-1} dF_U - \int_{\xi_{(n)}}^1 F_Y^{-1} dF_U\right\} \\ &= -\sqrt{n}\left\{\int_{\xi_{(1)}}^{\xi_{(n)}} [F_U \circ \mathbb{G}_n - F_U]\, dF_Y^{-1} - \int_0^{\xi_{(1)}} F_Y^{-1} dF_U - \int_{\xi_{(n)}}^1 F_Y^{-1} dF_U\right\},\end{aligned}$$

where the last equality relies on Assumption 2, $d_1 + b_1 < 1$ and $d_2 + b_2 < 1$. Next, using Assumption 2 again,

$$\begin{aligned}\left|\int_0^{\xi_{(1)}} F_Y^{-1} dF_U\right| &\lesssim \mathbb{1}\left\{\xi_{(1)} \geq 1/2\right\}\left|\int_0^1 F_Y^{-1} dF_U\right| + \mathbb{1}\left\{\xi_{(1)} < 1/2\right\}\int_0^{\xi_{(1)}} t^{-b_1-d_1} dt \\ &\lesssim \mathbb{1}\left\{\xi_{(1)} \geq 1/2\right\} + \xi_{(1)}^{1-b_1-d_1}.\end{aligned}$$

Thus, because $\xi_{(1)} = O_p(1/n)$ and $b_1 + d_1 < 1/2$, $\sqrt{n}\int_0^{\xi_{(1)}} F_Y^{-1} dF_U = o_p(1)$. Similarly, $\sqrt{n}\int_{\xi_{(n)}}^1 F_Y^{-1} dF_U = o_p(1)$. Hence,

$$\sqrt{n}T_{1n} = -\sqrt{n}\int_{\xi_{(1)}}^{\xi_{(n)}} [\mathbb{G}_n - I] d\Lambda + R_n + o_p(1),$$

where Λ is the measure defined by $d\Lambda/dF_Y^{-1} = f_U$ and

$$R_n := \sqrt{n}\left(\int_{\xi_{(1)}}^{\xi_{(n)}} [\mathbb{G}_n - \mathrm{I}] f_U\, dF_Y^{-1} - \int_{\xi_{(1)}}^{\xi_{(n)}} [F_U \circ \mathbb{G}_n - F_U]\, dF_Y^{-1}\right).$$

We show below that $R_n = o_p(1)$, which further proves that

$$\sqrt{n}T_{1n} = \frac{1}{\sqrt{n}}\sum_{i=1}^n \eta_i + o_p(1), \tag{8}$$

with $\eta_i := -\int_0^1 [\mathbb{1}\{F_Y(Y_i) \leq t\} - t] d\Lambda(t)$. MODIFY THIS...

By the mean value theorem, there exists $T_n(t) \in (\mathbb{G}_n(t), t)$ such that

$$R_n = \sqrt{n}\int_{\xi_{(1)}}^{\xi_{(n)}} \underbrace{[f_U - f_U \circ T_n]}_{:=A_n} [\mathbb{G}_n - \mathrm{I}]\; dF_Y^{-1}.$$

By Assumption 2, there exists $\delta > 0$ such that $b_j + d_j < 1/2 - \delta$. Further, let $\delta_j > 0$ be such that

$$b_j + d_j < 1/2 - \delta - \delta_j. \tag{9}$$

Then let $q(t) = t^{1/2-\delta_1}(1-t)^{1/2-\delta_2}$. From what precedes, we have

$$|R_n| \quad \leq \sup_{t\in(0,1)} \left| \frac{\sqrt{n}(\mathbb{G}_n(t) - t)}{q(t)} \right| \int_{\xi_{(1)}}^{\xi_{(n)}} |A_n(t)| q(t)\, dF_Y^{-1}(t). \tag{10}$$

We now show that the second term tends to 0 almost surely. First, by convergence of $\mathbb{G}_n(t)$ to t, we have, for all $t \in (0,1)$, $T_n(t) \xrightarrow{\text{a.s.}} t$. Then, by continuity of f_U, $A_n(t) \xrightarrow{\text{a.s.}} 0$ for all $t \in (0,1)$. Fix $\varepsilon > 0$. By Theorem 10.6.1 in Shorack and Wellner (1986), we have, for all $t \geq \xi_{(1)}$ and all n large enough,

$$\mathbb{G}_n(t) \leq (1+\varepsilon)t^{1-\delta/2} \leq (1+\varepsilon)t^{1-\delta}.$$

Now, let $B(t) := C_U t^{-b_1}(1-t)^{-b_2}$. Then, by Assumption 2 and because B is a convex function, we obtain, for all $t \in [\xi_{(1)}, \xi_{(n)}]$,

$$\begin{aligned} |A_n(t)| &\leq [B(\mathbb{G}_n(t)) \vee B(t)] + B(t) \\ &\lesssim t^{-b_1-\delta}(1-t)^{-b_2-\delta}, \quad \text{a.s.} \end{aligned}$$

Therefore,

$$|A_n(t)| \mathbb{1}\left\{t \in [\xi_{(1)}, \xi_{(n)}]\right\} q(t) \lesssim t^{1/2-b_1-\delta-\delta_1}(1-t)^{1/2-b_2-\delta-\delta_2}.$$

Moreover, by (9) and Lemma 3,

$$\int_0^1 t^{1/2-b_1-\delta-\delta_1}(1-t)^{1/2-b_2-\delta-\delta_2}\, dF_Y^{-1}(t) < \infty.$$

Then, by the dominated convergence theorem,

$$\int_{\xi_{(1)}}^{\xi_{(n)}} |A_n(t)| q(t)\, dF_Y^{-1}(t) \xrightarrow{\text{a.s.}} 0. \tag{11}$$

Next, by Equation (2) in Chapter 2, Section 7 (page 141) in Shorack and Wellner (1986), we have

$$\sup_{t\in(0,1)} \left| \frac{\sqrt{n}(\mathbb{G}_n(t) - t)}{q(t)} \right| = o_p(1).$$

This, together with (10) and (11), implies that $R_n = o_p(1)$.

Second step: linearization of T_{2n}. By Lemma 2 followed by an integration by part,

$$\sqrt{n}T_{2n} = \sqrt{n}\int_0^1 F_Y^{-1}\, d\left[\widehat{F}_U \circ \mathbb{G}_n - F_U \circ \mathbb{G}_n\right]$$

$$= \left[F_Y^{-1}(t)\left(\widehat{F}_U(\mathbb{G}_n(t)) - F_U(\mathbb{G}_n(t))\right)\right]_0^1 - \sqrt{n}\int_0^1 \left[\widehat{F}_U \circ \mathbb{G}_n - F_U \circ \mathbb{G}_n\right] dF_Y^{-1} \quad (12)$$

$$= -\sqrt{n}\int_0^1 \left[\widehat{F}_U \circ \mathbb{G}_n - F_U \circ \mathbb{G}_n\right] dF_Y^{-1}, \quad (13)$$

since for $t \in (0, \xi_{(1)})$, $\mathbb{G}_n(t) = 0$ and $\widehat{F}_U(0) = F_U(0) = 0$ because $(U_i)_{i=1,\dots,n}$ is an iid sample of random variables absolutely continuous with respect to the Lebesgue measure on $[0,1]$. Symmetrically, for $t \in (\xi_{(n)}, 1)$, $\mathbb{G}_n(t) = 1$ and $\widehat{F}_U(1) = F_U(1) = 1$. We now prove that

$$-\sqrt{n}\int_0^1 \left[\widehat{F}_U \circ \mathbb{G}_n - F_U \circ \mathbb{G}_n\right] dF_Y^{-1} = -\sqrt{n}\int_0^1 \left[\widehat{F}_U - F_U\right] dF_Y^{-1} + o_p(1). \quad (14)$$

Let $\mathbb{V}_n = \sqrt{n}(\widehat{F}_U \circ F_U^{-1} - I)$ denote the empirical process associated with the uniform variables $(F_U(U_i))_{i=1,\dots,n}$ and define

$$R_n = \int_0^1 \left(\mathbb{V}_n \circ F_U \circ \mathbb{G}_n - \mathbb{V}_n \circ F_U\right) dF_Y^{-1}.$$

Equation (14) is equivalent to $R_n = o_p(1)$. We actually prove the stronger result that $E[|R_n|] \to 0$. For that purpose, let $I_n(x) = (x, \mathbb{G}_n(x)]$ if $\mathbb{G}_n(x) > x$, $I_n(x) = [\mathbb{G}_n(x), x)$ if $\mathbb{G}_n(x) < x$ and $\emptyset$ otherwise. Finally, let $S_n(x) = \mathrm{sgn}(\mathbb{G}_n(x) - x)$. Observe first that

$$\mathbb{V}_n \circ F_U \circ \mathbb{G}_n(x) - \mathbb{V}_n \circ F_U(x) = S_n(x) Z_n(x), \quad (15)$$

with

$$Z_n(x) = \frac{1}{\sqrt{n}}\sum_{i=1}^n \left[\mathbb{1}\left\{U_i \in I_n(x)\right\} - P_U(I_n(x))\right],$$

where $P_U([a,b]) = P_U((a,b]) = P_U([a,b)) = P_U((a,b)) = F_U(a) - F_U(b)$ for all $(a,b) \in [0,1], a \leq b$. Then,

$$\begin{aligned}
E\left[|R_n| \,|(\xi_i)_i\right] &\leq E\left[\int_0^1 \left|\mathbb{V}_n \circ F_U \circ \mathbb{G}_n - \mathbb{V}_n \circ F_U\right| dF_Y^{-1}\middle|(\xi_i)_i\right] \\
&= \int_0^1 E\left[\left|\mathbb{V}_n \circ F_U \circ \mathbb{G}_n - \mathbb{V}_n \circ F_U\right| \,|(\xi_i)_i\right] dF_Y^{-1} \\
&\leq \int_0^1 E\left[Z_n(x)^2|(\xi_i)_i\right]^{1/2} dF_Y^{-1}(x) \\
&= \int_0^1 V\left[\mathbb{1}\left\{U_1 \in I_n(x)\right\}|(\xi_i)_i\right]^{1/2} dF_Y^{-1}(x) \\
&\leq \int_0^1 \left|P_U(I_n(x))\right|^{1/2} dF_Y^{-1}(x). \quad (16)
\end{aligned}$$

The first equality follows by Fubini-Tonelli's theorem, the second inequality uses (15) and the Cauchy-Schwarz inequality and the second equality holds since conditional on the $(\xi_i)_i$, the variables $\mathbb{1}\{U_i \in I_n(x)\} - P_U(I_n(x))$ are i.i.d. with mean zero. As a result,

$$\begin{aligned} E\left[|R_n|\right] &\leq \int_0^1 E\left[|P_U(I_n(x)|^{1/2}\right] dF_Y^{-1}(x) \\ &\leq \int_0^1 E\left[|P_U(I_n(x)|\right]^{1/2} dF_Y^{-1}(x), \end{aligned} \tag{17}$$

where the first ineqality follows by (16) and Fubini-Tonelli's theorem, whereas the second is due to Jensen's inequality. Now, by the law of large numbers and the continuous mapping theorem, $|F_U(\mathbb{G}_n(x)) - F_U(x)| \xrightarrow{\mathbb{P}} 0$ for all $x \in [0,1]$. Moreover, $|F_U(\mathbb{G}_n(x)) - F_U(x)| \leq 1$. Hence, for all $x \in [0,1]$,

$$E\left[|F_U(x) - F_U(\mathbb{G}_n(x))|\right] \to 0.$$

We now apply the dominated convergence theorem to prove that $E[|R_n|] \to 0$. Because $x \mapsto E[|P_U(I_n(x))|]^{1/2}$ is bounded by 1 for all n, it is actually enough to bound this function for x close to 0 and close to 1. Also, by symmetry, we can focus without loss of generality on the neighborhood of 0. We prove that

$$E\left[|P_U(I_n(x))|\right] \lesssim x^{1-b_1}. \tag{18}$$

Then the result follows by Lemma 3 combined with Assumption 2. To prove (18), we apply Lemma 5 with $Q_n(x) := \mathbb{G}_n(x)$ and $\delta < \exp(-1)/2$. If $x \geq 1/n$, Cauchy-Schwarz inequality yields

$$E\left[|\mathbb{G}_n(x) - x|\right] \leq \left[\frac{x(1-x)}{n}\right]^{1/2} \leq 2x, \tag{19}$$

since $n^{1/2} \geq x^{-1/2}$. If $x < 1/n$, (19) holds as well by Theorem 1 in Berend and Kontorovich (2013). Hence, (19) holds for all $x \in (0, \bar{\delta}/2)$. Next, let $n_0 \in \mathbb{N}$, $n_0 \geq 4/(1-\bar{\delta})^2$. By Kiefer's inequality (see, e.g. Van der Vaart and Wellner, 1996, Corollary A.6.3), we have, for all $x \in [0, \delta]$ and all $n \geq n_0$,

$$\Pr(\mathbb{G}_n(x) > 1/2) \leq (ex)^{n(1-\bar{\delta})^2/4} \lesssim x. \tag{20}$$

Thus, we can apply Lemma 5, which yields (18).

Hence, (14) holds. Combined with (13), this implies that

$$\sqrt{n}T_{2n} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\varepsilon_i + o_p(1), \tag{21}$$

with $\varepsilon_i = -\int_0^1 [\mathbb{1}\{U_i \leq t\} - F_U(t)]\, dF_Y^{-1}(t)$.

Third step: conclusion. By definition of η_i and ε_i, we have$E[\eta_i] = E[\varepsilon_i] = 0$ and

$$\begin{aligned} E[\eta_i^2] &= \int_0^1 \int_0^1 (s \wedge t - st) f_U(s) f_U(t)\, dF_Y^{-1}(s)\, dF_Y^{-1}(t), \\ E[\varepsilon_i^2] &= \int_0^1 \int_0^1 (F_U(s \wedge t) - F_U(s)F_U(t))\, dF_Y^{-1}(s)\, dF_Y^{-1}(t). \end{aligned}$$

Moreover, under Assumption 1, η_i and ε_i are independent. The result follows by the central limit theorem. $\square$

B.1 Proof of Theorem 2

We first decompose the difference $\check{\theta}_n - \theta_0$ into three parts that we study independently:

$$\begin{aligned} \check{\theta}_n - \theta_0 = &\underbrace{\int_0^1 F_Y^{-1} \circ \mathbb{G}_n^{-1}\, dF_U - \int_0^1 F_Y^{-1}\, dF_U}_{=T_{1n}} + \underbrace{\int_0^1 F_Y^{-1} \circ \mathbb{G}_n^{-1}\, d\widehat{F}_U - \int_0^1 F_Y^{-1} \circ \mathbb{G}_n^{-1}\, dF_U}_{=T_{2n}} \\ &+ \underbrace{\int_0^1 F_Y^{-1} \circ \mathbb{G}_n^{-1} \circ \mathbb{H}_n\, d\widehat{F}_U - \int_0^1 F_Y^{-1} \circ \mathbb{G}_n^{-1}\, d\widehat{F}_U}_{:=T_{3n}}. \end{aligned}$$

This decomposition is convenient as T_{1n} and T_{2n} have already been analyzed in the proof of Theorem 1. We then prove the result in eight steps. We first show that

$$\sqrt{n}T_{3n} = -\sqrt{n}\int_0^1 \left[\widehat{F}_U \circ \mathbb{H}_n^{-1} \circ \overline{\mathbb{G}}_n - \widehat{F}_U \circ \mathbb{G}_n\right] dF_Y^{-1}. \tag{22}$$

where $\overline{\mathbb{G}}_n$ is defined below. Second, we show that

$$\sqrt{n}T_{3n} = -\sqrt{n}\underbrace{\int_0^1 \left[F_U \circ \mathbb{H}_n^{-1} \circ \mathbb{G}_n - F_U \circ \mathbb{G}_n\right] dF_Y^{-1}}_{:=J_{1n}} + o_p(1). \tag{23}$$

Let us then write $-\sqrt{n}J_{1n} = \sqrt{n}J_{2n} + R_{1n} + R_{2n} + R_{3n} + R_{4n}$, with:

$$J_{2n} := -\int_{1/n}^{1-1/n} \left[\mathbb{H}_n^{-1}(x) - E[\mathbb{H}_n^{-1}(x)]\right] f_U(x)\, dF_Y^{-1}(x), \tag{24}$$

$$R_{1n} := -\sqrt{n}\left(J_{1n} - \int_{\xi_{(1)}}^{\xi_{(n)}} \left[\mathbb{H}_n^{-1} \circ \mathbb{G}_n - \mathbb{G}_n\right] f_U\, dF_Y^{-1}\right), \tag{25}$$

$$R_{2n} := -\sqrt{n}\left(\int_{\xi_{(1)}}^{\xi_{(n)}} \left[\mathbb{H}_n^{-1} \circ \mathbb{G}_n - \mathbb{G}_n\right] f_U\, dF_Y^{-1} - \int_{\xi_{(1)}}^{\xi_{(n)}} \left[\mathbb{H}_n^{-1} - \mathrm{I}\right] f_U\, dF_Y^{-1}\right), \tag{26}$$

$$R_{3n} := \int_{\xi_{(1)}}^{\xi_{(n)}} \left[x - E[\mathbb{H}_n^{-1}(x)]\right] f_U(x)\, dF_Y^{-1}(x), \tag{27}$$

$$\begin{aligned} R_{4n} :=& \int_{1/n}^{1-1/n} \left[\mathbb{H}_n^{-1}(x) - E[\mathbb{H}_n^{-1}(x)]\right] f_U(x)\, dF_Y^{-1}(x) \\ &- \int_{\xi_{(1)}}^{\xi_{(n)}} \left[\mathbb{H}_n^{-1}(x) - E[\mathbb{H}_n^{-1}(x)]\right] f_U(x)\, dF_Y^{-1}(x), . \end{aligned} \tag{28}$$

In the third to sixth steps, we prove that each of the four terms $R_{1n} - R_{4n}$ tends to 0 in probability. In the seventh step, we show that J_{2n} tend to a normal distribution. The eighth step concludes.

First step: Equation (22) **holds.** Let $\zeta_j = F_Z(Z_j)$, $X_n^0 := [0, \zeta_{(1)}]$ and $X_n^1 := [\zeta_{(n)}, 1]$. For all $t \in [0,1]$, let us also define

$$\begin{aligned} \overline{\mathbb{G}}_n(t) &= \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}\{\xi_i \le t\} + \frac{1}{n}\sum_{i=1}^{n-1} \mathbb{1}\{\xi_i < t < \xi_{i+1}\} \\ &= \mathbb{G}_n(t) + \frac{1}{n}\sum_{i=1}^{n-1} \mathbb{1}\{\xi_i < t < \xi_{i+1}\}. \end{aligned}$$

Then, remark that $\mathbb{G}_n^{-1} \circ \mathbb{H}_n$ is the generalized inverse of $\mathbb{H}_n^{-1} \circ \overline{\mathbb{G}}_n$. Then, by splitting the first integral in $\sqrt{n}T_{3n}$ and applying Lemma 2, we obtain

$$\begin{aligned} \sqrt{n}T_{3n} &= \sqrt{n}\left(\int_{(X_n^0 \cup X_n^1)^c} F_Y^{-1} \circ \mathbb{G}_n^{-1} \circ \mathbb{H}_n\, d\widehat{F}_U - \int_0^1 F_Y^{-1} \circ \mathbb{G}_n^{-1}\, d\widehat{F}_U \right.\\ &\quad \left. + \int_{X_n^0 \cup X_n^1} F_Y^{-1} \circ \mathbb{G}_n^{-1} \circ \mathbb{H}_n\, d\widehat{F}_U\right) \\ &= \sqrt{n}\left(\int_{\xi_{(1)}}^{\xi_{(n)}} F_Y^{-1}\, d\left[\widehat{F}_U \circ \mathbb{H}_n^{-1} \circ \overline{\mathbb{G}}_n\right] - \int_0^1 F_Y^{-1}\, d\left[\widehat{F}_U \circ \mathbb{G}_n\right] \right.\\ &\quad \left. + \int_{X_n^0 \cup X_n^1} F_Y^{-1} \circ \mathbb{G}_n^{-1} \circ \mathbb{H}_n\, d\widehat{F}_U\right) \\ &= \sqrt{n}\int_{\xi_{(1)}}^{\xi_{(n)}} F_Y^{-1}\, d\left[\widehat{F}_U \circ \mathbb{H}_n^{-1} \circ \overline{\mathbb{G}}_n - \widehat{F}_U \circ \mathbb{G}_n\right] + \int_{X_n^0 \cup X_n^1} F_Y^{-1} \circ \mathbb{G}_n^{-1} \circ \mathbb{H}_n\, d\widehat{F}_U, \end{aligned} \tag{29}$$

where we used the fact that $\widehat{F}_U \circ \mathbb{G}_n$ is constant on the two segments $[0,\xi_{(1)}]$ and $[\xi_{(n)},1]$ to obtain the third equality. Remark that

$$\begin{aligned}
&\sqrt{n}\left[F_Y^{-1}(t)\left(\widehat{F}_U \circ \mathbb{H}_n^{-1} \circ \overline{\mathbb{G}}_n(t) - \widehat{F}_U \circ \mathbb{G}_n(t)\right)\right]_{t=\xi_{(1)}}^{t=\xi_{(n)}} \\
&\quad = \sqrt{n}\Big[\mathbb{1}\left\{\zeta_{(n)} < U_{(n)}\right\} F_Y^{-1}(\xi_{(n)})\left(\widehat{F}_U(\zeta_{(n)}) - 1\right) - \mathbb{1}\left\{\zeta_{(1)} \geq U_{(1)}\right\} F_Y^{-1}(\xi_{(1)})\widehat{F}_U(\zeta_{(1)})\Big].
\end{aligned}$$

Also, since $\mathbb{H}_n$ is constant on the two segments X_n^0 and X_n^1, we have

$$\begin{aligned}
&\sqrt{n}\int_{X_n^0 \cup X_n^1} F_Y^{-1} \circ \mathbb{G}_n^{-1} \circ \mathbb{H}_n \, d\widehat{F}_U \\
&\quad = \sqrt{n}\left[\widehat{F}_U(1) - \widehat{F}_U(\zeta_{(n)})\right] F_Y^{-1}(\xi_{(n)}) + \sqrt{n}\left[\widehat{F}_U(\zeta_{(1)}) - \widehat{F}_U(0)\right] F_Y^{-1}(\xi_{(1)}) \\
&\quad = \sqrt{n}\Big[\mathbb{1}\left\{\zeta_{(1)} \geq U_{(1)}\right\} F_Y^{-1}(\xi_{(1)})\widehat{F}_U(\zeta_{(1)}) - \mathbb{1}\left\{\zeta_{(n)} < U_{(n)}\right\} F_Y^{-1}(\xi_{(n)})\left(\widehat{F}_U(\zeta_{(n)}) - 1\right)\Big].
\end{aligned}$$

Thus, an integration by part of the first term in (29) yields (22).

Second step: Equation (23) **holds.** From (22), we have

$$\begin{aligned}
\sqrt{n}T_{3n} = & - \underbrace{\sqrt{n}\int_0^1 \left[F_U \circ \mathbb{H}_n^{-1} \circ \mathbb{G}_n - F_U \circ \mathbb{G}_n\right] dF_Y^{-1}}_{=:J_{1n}} \\
& - \sqrt{n}\int_0^1 \left[\widehat{F}_U \circ \mathbb{H}_n^{-1} \circ \overline{\mathbb{G}}_n - F_U \circ \mathbb{H}_n^{-1} \circ \mathbb{G}_n\right] dF_Y^{-1} \\
& - \sqrt{n}\int_0^1 \left[F_U \circ \mathbb{G}_n - \widehat{F}_U \circ \mathbb{G}_n\right] dF_Y^{-1}.
\end{aligned}$$

We show below that

$$\sqrt{n}\int_0^1 \left[\widehat{F}_U \circ \mathbb{H}_n^{-1} \circ \overline{\mathbb{G}}_n - F_U \circ \mathbb{H}_n^{-1} \circ \mathbb{G}_n\right] dF_Y^{-1} = \sqrt{n}\int_0^1 \left[\widehat{F}_U - F_U\right] dF_Y^{-1} + o_p(1). \quad (30)$$

Once combined with (14), this proves (23). To prove (30), we follow closely the proof of (14). Let $\mathbb{V}_n = \sqrt{n}(\widehat{F}_U \circ F_U^{-1} - I)$,

$$R_n = \int_0^1 \left(\mathbb{V}_n \circ F_U \circ \mathbb{H}_n^{-1} \circ \overline{\mathbb{G}}_n - \mathbb{V}_n \circ F_U\right) dF_Y^{-1},$$

and $\bar{I}_n(x) = (x, \mathbb{H}_n^{-1} \circ \overline{\mathbb{G}}_n(x)]$ if $\mathbb{H}_n^{-1} \circ \overline{\mathbb{G}}_n(x) > x$, $\bar{I}_n(x) = [\mathbb{H}_n^{-1} \circ \overline{\mathbb{G}}_n(x), x)$ if $\mathbb{H}_n^{-1} \circ \overline{\mathbb{G}}_n(x) < x$ and $\emptyset$ otherwise. We prove that $E[|R_n|] \to 0$. Reasoning as to obtain (17) (but conditioning first on $(\xi_i,\zeta_i)_i$ instead of just on $(\xi_i)_i$), we get

$$E\left[|R_n|\right] \leq \int_0^1 E\left[\left|F_U(x) - F_U(\mathbb{H}_n^{-1} \circ \overline{\mathbb{G}}_n(x))\right|\right]^{1/2} dF_Y^{-1}(x).$$

Because $\overline{\mathbb{G}}_n(x) \xrightarrow{\mathbb{P}} x$, by uniform convergence of $\mathbb{H}_n^{-1}$ towards I and the continuous mapping theorem, $|F_U(\mathbb{H}_n^{-1} \circ \overline{\mathbb{G}}_n(x)) - F_U(x)| \xrightarrow{\mathbb{P}} 0$ for all $x \in [0,1]$. Moreover, $|F_U(\mathbb{H}_n^{-1} \circ \overline{\mathbb{G}}_n(x)) - F_U(x)| \leq 1$. Hence, for all $x \in [0,1]$,

$$E\left[|F_U(x) - F_U(\mathbb{H}_n^{-1} \circ \overline{\mathbb{G}}_n(x))|\right] \to 0.$$

Next, we show $E[|R_n|] \to 0$ by proving

$$E\left[|F_U(x) - F_U(\mathbb{H}_n^{-1} \circ \overline{\mathbb{G}}_n(x))|\right] \lesssim x^{1-b_1}. \tag{31}$$

and applying the dominated convergence theorem. As in Theorem 1, we apply Lemma 5 with $Q_n(x) := \mathbb{H}_n^{-1} \circ \overline{\mathbb{G}}_n(x)$. The two conditions of this lemma are checked in Lemma 4. Hence, (31), and thus (23), hold.

Third step: $R_{1n} = o_p(1)$. Recall that R_{1n} is defined in (25). By the mean value theorem, there exists $T_n(t) \in (\mathbb{G}_n(t), \mathbb{H}_n^{-1} \circ \mathbb{G}_n(t))$ such that

$$R_{1n} = \sqrt{n} \int_{\xi_{(1)}}^{\xi_{(n)}} \underbrace{[f_U - f_U \circ T_n]}_{:=A_n} \left[\mathbb{H}_n^{-1} \circ \mathbb{G}_n - \mathbb{G}_n\right] dF_Y^{-1}.$$

By Assumption 2, there exists $\delta > 0$ such that $b_j + d_j < 1/2 - \delta$. Further, let $\delta_j > 0$ be such that

$$b_j + d_j < 1/2 - \delta - \delta_j. \tag{32}$$

Then let $q(t) = t^{1/2-\delta_1}(1-t)^{1/2-\delta_2}$. From what precedes, we have

$$|R_{1n}| \quad \leq \sup_{t \in (1/n, 1-1/n)} \left| \frac{\sqrt{n}(\mathbb{H}_n^{-1}(t) - t)}{q(t)} \right| \int_{\xi_{(1)}}^{\xi_{(n)}} |A_n(t)| q(t) \, dF_Y^{-1}(t). \tag{33}$$

We now show that the second term tends to 0 almost surely. First, by uniform convergence of $\mathbb{H}_n^{-1}$ towards I and convergence of $\mathbb{G}_n(t)$ to t, we have, for all $t \in (0,1)$, $T_n(t) \xrightarrow{\text{a.s.}} t$. Then, by continuity of f_U, $A_n(t) \xrightarrow{\text{a.s.}} 0$ for all $t \in (0,1)$. Fix $\varepsilon > 0$. By Theorem 10.6.1 in Shorack and Wellner (1986), we have, for all $t \geq \xi_{(1)}$ and all n large enough,

$$\mathbb{G}_n(t) \leq (1+\varepsilon) t^{1-\delta/2} \leq (1+\varepsilon) t^{1-\delta}.$$

Moreover, by the same theorem, we have, for all $u \geq 1/n$,

$$\mathbb{H}_n^{-1}(u) \leq (1+\varepsilon)u^{(1-\delta/2)}.$$

Then, since $\mathbb{G}_n(t) \geq 1/n$ for all $t \geq \xi_{(1)}$,

$$\mathbb{H}_n^{-1} \circ \mathbb{G}_n(t) \leq (1+\varepsilon)^2 t^{1-\delta}.$$

Now, let $B(t) := C_U t^{-b_1}(1-t)^{-b_2}$. Then, by Assumption 2 and because B is a convex function, we obtain, for all $t \in [\xi_{(1)}, \xi_{(n)}]$,

$$\begin{aligned}|A_n(t)| &\leq \left[B\left(\mathbb{H}_n^{-1} \circ \mathbb{G}_n(t)\right) \vee B\left(\mathbb{G}_n(t)\right)\right] + B(t) \\ &\lesssim t^{-b_1-\delta}(1-t)^{-b_2-\delta}, \quad \text{a.s.}\end{aligned}$$

Therefore,

$$|A_n(t)| \mathbb{1}\left\{t \in [\xi_{(1)}, \xi_{(n)}]\right\} q(t) \lesssim t^{1/2-b_1-\delta-\delta_1}(1-t)^{1/2-b_2-\delta-\delta_2}.$$

Moreover, by (32) and Lemma 3,

$$\int_0^1 t^{1/2-b_1-\delta-\delta_1}(1-t)^{1/2-b_2-\delta-\delta_2}\, dF_Y^{-1}(t) < \infty.$$

Then, by the dominated convergence theorem,

$$\int_{\xi_{(1)}}^{\xi_{(n)}} |A_n(t)| q(t)\, dF_Y^{-1}(t) \xrightarrow{\text{a.s.}} 0. \tag{34}$$

Next, by Corollary 4.3.1 and Theorem 3.4 in Csorgo et al. (1986),

$$\sup_{t \in (1/n, 1-1/n)} \left| \frac{\sqrt{n}(\mathbb{H}_n^{-1}(t) - t)}{q(t)} \right| = O_p(1).$$

This, together with (33) and (34), implies that $R_{1n} = o_p(1)$.

Fourth step: $R_{2n} = o_p(1)$. Recall that R_{2n} is defined in (26). We actually prove the result in L^1. Let $\mathbb{W}_n = \sqrt{n}(\mathbb{H}_n^{-1} - \mathrm{I})$ and $B_n = \mathbb{1}\left\{\xi_{(1)} \leq x < \xi_{(n)}\right\}$. We have, by Fubini-Tonelli's theorem,

$$\begin{aligned}E[|R_{2n}|] &\leq \int_0^1 E[|\mathbb{W}_n \circ \mathbb{G}_n(x) - \mathbb{W}_n(x)| \times B_n] f_U(x) dF_Y^{-1}(x) \\ &\leq \int_0^1 E[(\mathbb{W}_n \circ \mathbb{G}_n(x) - \mathbb{W}_n(x))^2 \times B_n]^{1/2} f_U(x) dF_Y^{-1}(x).\end{aligned}$$

We apply the dominated convergence theorem to prove the result. First, note that for all $(i,j) \in \{1,...,n\}^2$,

$$|\mathbb{H}_n^{-1}(j/n) - \mathbb{H}_n^{-1}(i/n)| \sim \text{Beta}(|j-i|, n-|j-i|+1),$$

with the convention that the Beta$(0, n+1)$ is the Dirac distribution at 0. Hence, for any $k \in \{1, ..., n-1\}$,

$$\begin{aligned}
&E\left[(\mathbb{W}_n \circ \mathbb{G}_n(x) - \mathbb{W}_n(x))^2 \,|\mathbb{G}_n(x) = k/n\right] \\
&= n\left\{E\left[\left(\mathbb{H}_n^{-1}(k/n) - \mathbb{H}_n^{-1}(x) - (k/n - x)\right)^2\right]\right\} \\
&= n\left\{E\left[\left(\mathbb{H}_n^{-1}(k/n) - \mathbb{H}_n^{-1}(\lceil nx\rceil/n) - \frac{n}{n+1}(k/n - \lceil nx\rceil/n)\right.\right.\right. \\
&\quad \left.\left.\left.+\frac{n}{n+1}(\lceil nx\rceil/n - x) - \frac{1}{n+1}(k/n - x)\right)^2\right]\right\} \\
&= n\left\{V\left[\mathbb{H}_n^{-1}(k/n) - \mathbb{H}_n^{-1}(x)\right] + \frac{1}{(n+1)^2}(\lceil nx\rceil - nx - k/n + x)^2\right\} \\
&= \frac{n}{(n+1)^2(n+2)}\left|k - \lceil nx\rceil\right|(n+1-|k-\lceil nx\rceil|) + \frac{n}{(n+1)^2}(\lceil nx\rceil - nx - k/n + x)^2 \\
&\le \left|\frac{k}{n} - \frac{\lceil nx\rceil}{n}\right| + \frac{2}{n}\left[(\lceil nx\rceil - nx)^2 + \left(\frac{k}{n} - x\right)^2\right] \\
&\le \left|\frac{k}{n} - x\right| + \frac{1}{n}\left[3 + 2\left(\frac{k}{n} - x\right)^2\right],
\end{aligned} \tag{35}$$

where the first inequality follows by convexity and the last by the triangle inequality and because by definition, $|nx - \lceil nx\rceil| \le 1$. Now, remark that $B_n = 1$ iff $n\mathbb{G}_n(x) \in \{1, ..., n-1\}$. Then, by what precedes,

$$\begin{aligned}
E\left[(\mathbb{W}_n \circ \mathbb{G}_n(x) - \mathbb{W}_n(x))^2 \times B_n\right] &\le E\left[|\mathbb{G}_n(x) - x|\right] + \frac{1}{n}\left[3 + 2V(\mathbb{G}_n(x))\right] \\
&\to 0.
\end{aligned} \tag{36}$$

To apply the dominated convergence theorem, we show that there exists $q(.)$ such that for all $n \ge n_0$ and all $x \in [0,1]$,

$$E[(\mathbb{W}_n \circ \mathbb{G}_n(x) - \mathbb{W}_n(x))^2 \times B_n]^{1/2} \le q(x), \tag{37}$$

with $\int_0^1 q(x) f_U(x) dF_Y^{-1}(x) < \infty$. As above, we focus on a neighborhood of 0. If $x > 1/n$, we have, by (36) and (19),

$$E\left[(\mathbb{W}_n \circ \mathbb{G}_n(x) - \mathbb{W}_n(x))^2 \times B_n\right] \leq 7x.$$

Now suppose that $x < 1/n$. Remark that $E(B_n) \leq 1 - (1-x)^n \leq nx$. Then, integrating (35), we obtain

$$\begin{aligned} E\left[(\mathbb{W}_n \circ \mathbb{G}_n(x) - \mathbb{W}_n(x))^2 \times B_n\right] &\leq E\left[\mathbb{G}_n(x)\right] + \frac{1}{n}\left[3nx + 2V(\mathbb{G}_n(x))\right] \\ &\leq 7x. \end{aligned}$$

Then we can choose $q(x) = (7x)^{1/2}$ in (37). By Assumption 2 and Lemma 3, we have $\int_0^{1/2} q(x) f_U(x) dF_Y^{-1}(x) < \infty$. The same reasoning applies to the interval $[1/2, 1]$. The result follows.

Fifth step: $R_{3n} = o_p(1)$**.** Recall that R_{3n} is defined in (27). Let Λ denote the measure on $(0,1)$ such that $d\Lambda/dF_Y^{-1} = f_U$. Remark that $E[\mathbb{H}_n(x)] = \lceil nx \rceil/(n+1)$. Then

$$E[|R_{3n}|] \leq \int_0^1 [1 - x^n - (1-x)^n] \left| \frac{\lceil nx \rceil - (n+1)x}{(n+1)n^{-1/2}} \right| d\Lambda(x).$$

Let $f_n(x)$ denote the integrand. We have $\lim_{n\to\infty} f_n(x) = 0$. Moreover, using $1 - x^n - (1-x)^n \leq nx$, we obtain, when $x < 1/n$,

$$f_n(x) \leq 2n^{1/2}x \leq x^{1/2} \lesssim [x(1-x)]^{1/2}.$$

When $x \in [1/n, 1-1/n]$,

$$f_n(x) \leq \frac{2}{n^{1/2}} \lesssim [x(1-x)]^{1/2}.$$

Finally, when $x > 1 - 1/n$, using $1 - x^n \leq n(1-x)$,

$$f_n(x) \leq n(1-x)\frac{2}{(n+1)n^{-1/2}} \leq 2(1-x)^{1/2} \lesssim [x(1-x)]^{1/2}.$$

Moreover, $\int_0^1 [x(1-x)]^{1/2} d\Lambda < \infty$ by Lemma 3. Thus, by the dominated convergence theorem, $R_{4n} = o_p(1)$.

Sixth step: $R_{4n} = o_p(1)$**.** Recall that R_{4n} is defined in (28). We prove the stronger result that R_{4n} converges to 0 in L^1. By Fubini-Tonelli's theorem combined with Assumption 1 and Jensen's inequality, we have

$$E[|R_{4n}|] \leq \int_0^1 \sqrt{n} E\left[\left|\mathbb{1}\left\{x \in [\xi_{(1)}, \xi_{(n)}]\right\} - \mathbb{1}\left\{x \in [1/n, (n-1)/n]\right\}\right|\right] \\ \times E\left[\left(\mathbb{H}_n^{-1}(x) - \frac{\lceil nx \rceil}{(n+1)}\right)^2\right]^{1/2} d\Lambda(x). \tag{38}$$

Since $\mathbb{H}_n^{-1}(x) \sim$Beta$(\lceil nx \rceil, n+1-\lceil nx \rceil)$, we have

$$E\left[\left(\mathbb{H}_n^{-1}(x) - \frac{\lceil nx \rceil}{(n+1)}\right)^2\right]^{1/2} = \sqrt{\frac{\lceil nx \rceil (n+1-\lceil nx \rceil)}{(n+1)^2(n+2)}} \lesssim \sqrt{\frac{x(1-x)}{n}}.$$

Let $q_n(x)$ denote the first expectation in the integrand. By letting $p_n(x) := 1-x^n-(1-x)^n$, we have

$$\begin{aligned} q_n(x) &= \Pr(\xi_{(1)} \leq x \leq \xi_{(n)}, x < 1/n) + \Pr(\xi_{(1)} \leq x \leq \xi_{(n)}, x > (n-1)/n) \\ &\quad + \Pr(\xi_{(1)} > x \cup x > \xi_{(n)}, 1/n \leq x \leq (n-1)/n) \\ &= p_n(x)\left[\mathbb{1}\left\{x < 1/n\right\} + \mathbb{1}\left\{1-x < 1/n\right\}\right] + (1-p_n(x))\mathbb{1}\left\{1/n \leq x \leq (n-1)/n\right\}. \end{aligned}$$

Let $f_n(x)$ denote the integrand in the right-hand side of (38). For all $x \in (0,1)$, $\lim_{n\to\infty} p_n(x) = 1$ so from what precedes, $\lim_{n\to\infty} f_n(x) = 0$ for all $x \in [0,1]$. Moreover, using $q_n(x) \leq 1$, we get

$$f_n(x) \lesssim [x(1-x)]^{1/2},$$

with $\int_0^1 [x(1-x)]^{1/2} d\Lambda < \infty$ by Lemma 3. The result follows by the dominated convergence theorem.

Seventh step: asymptotic normality of J_{2n}**.** Let $\nu_i = F_Z(Z_i)$ and $I_{in} = [(i-1)/n, i/n)$. First, note that

$$-\sqrt{n} J_{2n} = \sum_{i=1}^n a_{in}\left(\nu_{(i)} - \frac{i}{(n+1)}\right), \tag{39}$$

where $a_{1n} = a_{nn} = 0$, and, for all $i \in \{2, \ldots, n-1\}$, $a_{in} = \sqrt{n}\Lambda\left(I_{in}\right)$. We now verify that the necessary and sufficient conditions given by Hecker (1976) for the asymptotic normality of the $L-$statistic in (39) hold in our case. Let us define

$$\sigma_n^2 = \frac{1}{n+2}\sum_{j=1}^{n}\sum_{k=1}^{n} a_{jn}a_{kn}\left[\left(\frac{j}{n+1}\wedge\frac{k}{n+1}\right) - \frac{jk}{(n+1)^2}\right].$$

We have to prove that

$$\lim_{n\to\infty}\frac{\max_{1\le i\le n}\left|\sum_{j=i}^{n} a_{jn}\right|}{n\sigma_n} = 0. \tag{40}$$

First, by Assumption 2 and Lemma 3, there exists $\delta < 1/2$ such that

$$\int_0^1 t^{\delta-b_1}(1-t)^{\delta-b_2}\, dF_Y^{-1}(t) < +\infty.$$

Now, because $a_{in} \ge 0$, we have, for all $n \ge 2$,

$$\begin{aligned}
\max_{1\le i\le n}\left|\sum_{j=i}^{n} a_{jn}\right| &= \sqrt{n}\sum_{j=2}^{n-1}\Lambda\left(I_{jn}\right) \\
&= \sqrt{n}\int_{1/n}^{(n-1)/n} f_U(t)\, dF_Y^{-1}(t) \\
&\le C_U\sqrt{n}\int_{1/n}^{(n-1)/n} t^{-b_1}(1-t)^{-b_2}\, dF_Y^{-1}(t) \\
&\le C_U 2^{\delta} n^{1/2+\delta}\int_{1/n}^{(n-1)/n} t^{\delta-b_1}(1-t)^{\delta-b_2}\, dF_Y^{-1}(t) \\
&\le C_U 2^{\delta} n^{1/2+\delta}\int_0^1 t^{\delta-b_1}(1-t)^{\delta-b_2}\, dF_Y^{-1}(t),
\end{aligned}$$

where the first inequality follows by Assumption 2 and the second uses the fact that $[t(1-t)]^{\delta} \ge 1/(2n)^{\delta}$ for all $t \in [1/n, 1-1/n]$. Therefore,

$$\max_{1\le i\le n}\left|\sum_{j=i}^{n} a_{jn}\right| = O(n^{1/2+\delta}). \tag{41}$$

Next, we have

$$\begin{aligned}
\sigma_n^2 &= \frac{n}{n+2}\sum_{j=2}^{n-1}\sum_{k=2}^{n-1}\Lambda\left(I_{jn}\right)\Lambda\left(I_{kn}\right)\left(\frac{j}{n+1}\wedge\frac{k}{n+1} - \frac{jk}{(n+1)^2}\right) \\
&= \frac{n}{n+2}\int_0^1\int_0^1 f_n(x,y)d\Lambda(x)d\Lambda(y),
\end{aligned}$$

where $f_n(x,y) = \frac{j}{n+1} \wedge \frac{k}{n+1} - \frac{jk}{(n+1)^2}$ when $(x,y) \in I_{jn} \times I_{kn}$, $1 < j \wedge k \leq j \vee k < n$, $f_n(x,y) = 0$ otherwise. For any $(x,y) \in (0,1)^2$, $f_n(x,y) \to f(x,y) := x \wedge y - xy$. Moreover, for any $(x,y) \in I_{jn} \times I_{kn}$, $1 < j \wedge k \leq j \vee k < n$,

$$\frac{j}{n+1} \wedge \frac{k}{n+1} \leq 2(x \wedge y),$$
$$1 - \frac{j}{n+1} \vee \frac{k}{n+1} \leq 2\left(1 - x \vee y\right).$$

Thus, $f_n(x,y) \leq 4f(x,y)$ for all $(x,y) \in [1/n, 1-1/n]^2$. This inequality also holds for $(x,y) \in [0,1]^2 \backslash [1/n, 1-1/n]^2$ since $f_n(x,y) = 0$ for such (x,y). Because $x \wedge y \leq (xy)^{1/2}$ and $1 - x \vee y \leq [(1-x)(1-y)]^{1/2}$, we have $f(x,y) \leq [x(1-x)y(1-y)]^{1/2}$. Moreover, by Lemma 3, $\int_0^1 [\mathrm{I}(1-\mathrm{I})]^{1/2} d\Lambda < \infty$. Thus, by the dominated convergence theorem,

$$\lim_{n\to\infty} \sigma_n^2 = \sigma^2 := \int_0^1 \int_0^1 (x \wedge y - xy) d\Lambda(x) d\Lambda(y) > 0. \tag{42}$$

Combined with (41), this implies (40). Thus, by Theorem 1 of Hecker (1976) and (42) again,

$$-\sqrt{n} J_{2n} \xrightarrow{d} \mathcal{N}(0, \sigma^2).$$

Eighth step: conclusion. By the previous steps and the proof of Theorem 1, we have

$$\sqrt{n}\left(\check{\theta}_n - \theta_0\right) = \frac{1}{\sqrt{n}} \sum_{i=1}^n (\eta_i + \varepsilon_i) + \sqrt{n} J_{2n} + o_p(1).$$

As shown in the proof of Theorem 1, the first term on the right-hand side is asymptotically normal. The second term is also asymptotically normal by the previous step. Moreover, by Assumption 3, J_{2n} is independent of the $(\eta_i, \varepsilon_i)_{i\geq 1}$. Therefore, the vector $(\sum_{i=1}^n (\eta_i + \varepsilon_i)/\sqrt{n}, \sqrt{n} J_{2n})$ converges jointly in distribution to two independent normal variables distributions. The result follows. $\square$

C Technical lemmas

In Theorems 1 and 2, we use the following lemma, which is established in Proposition 1 of Falkner and Teschl (2012).

Lemma 2 *Let g be some Borel measurable function on $[0,1]$, and F, Q be cdf's on $[0,1]$. Then, for any $0 \leq a \leq b \leq 1$,*

$$\int_{Q(a)}^{Q(b)} g \circ Q^{-1} \, dF = \int_a^b g \, dF \circ Q. \tag{43}$$

Lemma 3 *Suppose that Assumption 2 holds and that $a_1 > d_1$ and $a_2 > d_2$, then $\int_0^1 x^{a_1}(1-x)^{a_2} \, dF_Y^{-1}(x) < \infty$.*

Proof: first, we have

$$\int_0^1 x^{a_1}(1-x)^{a_2} \, dF_Y^{-1}(x) = \int_{\mathbb{R}} F_Y(u)^{a_1}(1-F_Y(u))^{a_2} \, du.$$

By Assumption 2 (ii), for all $u \in \mathbb{R}$:

$$|u| \leq C F_Y(u)^{-d_1}(1-F_Y(u))^{-d_2}.$$

Fix $\varepsilon > 0$. Then, for all $u \leq -1 \wedge F_Y^{-1}(\varepsilon)$, $F_Y(u) \leq C^{1/d_1}(1-\varepsilon)^{-d_2/d_1}|u|^{-1/d_1}$. Thus:

$$\int_{-\infty}^{-1 \wedge F_Y^{-1}(\varepsilon)} F_Y(u)^{a_1}(1-F_Y(u))^{a_2} \, du \leq C^{1/d_1}(1-\varepsilon)^{-d_2/d_1} \int_{-\infty}^{-1 \wedge F_Y^{-1}(\varepsilon)} |u|^{-a_1/d_1} \, du < \infty,$$

since $d_1 < a_1$. A similar reasoning shows that $\int_{1 \vee F_Y^{-1}(1-\varepsilon)}^{\infty} F_Y(u)^{a_1}(1-F_Y(u))^{a_2} \, du < \infty$, using $d_2 < a_2$. $\square$

We recall that $\overline{\mathbb{G}}_n$ is defined as $\overline{\mathbb{G}}_n(x) = \mathbb{G}_n(x) + \sum_{i=1}^{n-1} \mathbb{1}\left\{\xi_{(i)} < x < \xi_{(i+1)}\right\}/n$.

Lemma 4 *(Useful properties of $\mathbb{H}_n^{-1} \circ \overline{\mathbb{G}}_n$) There exists $\delta \in (0, 1/2)$ and $n_0 \in \mathbb{N}$ such that for all $0 < x < \delta$ and all $n \geq n_0$,*

$$E\left[\left|\mathbb{H}_n^{-1} \circ \overline{\mathbb{G}}_n(x) - x\right|\right] \lesssim x. \tag{44}$$

Moreover, for any $\eta > 0$, there exists n_1 such that for all $n \geq n_1$ and for all $0 < x < \delta$,

$$\Pr(\mathbb{H}_n^{-1} \circ \overline{\mathbb{G}}_n(x) > 1/2) \lesssim x^{1-\eta}. \tag{45}$$

Inequalities (44)-(45) hold if we replace x by $1-x$, using possibly another δ and n_0.

Proof: Let us define

$$\widetilde{\mathbb{G}_n}(x) = \mathbb{G}_n(x) + \frac{\mathbb{1}\{0 < \mathbb{G}_n(x) < 1\}}{n}. \tag{46}$$

Observe that for a given $x \in [0,1]$, we have, with probability one, $\overline{\mathbb{G}}_n(x) = \widetilde{\mathbb{G}_n}(x)$. Then, letting $p_n(x) = [1 - x^n - (1-x)^n]/n$,

$$E[\overline{\mathbb{G}}_n(x)] = x + p_n(x).$$

We now establish (44).

By the triangle inequality,

$$E\left[\left|\mathbb{H}_n^{-1} \circ \overline{\mathbb{G}}_n(x) - x\right|\right] \leq E\left[\left|\mathbb{H}_n^{-1} \circ \overline{\mathbb{G}}_n(x) - \frac{n}{n+1}\overline{\mathbb{G}}_n(x)\right|\right] + E\left[\left|\frac{n}{n+1}\overline{\mathbb{G}}_n(x) - x\right|\right]. \tag{47}$$

Consider the second term first. We have

$$\begin{aligned} E\left[\left|\frac{n}{n+1}\overline{\mathbb{G}}_n(x) - x\right|\right] &\leq \frac{n}{n+1}E\left[\left|\mathbb{G}_n(x) - x\right|\right] + \frac{np_n(x)}{n+1} + \frac{x}{n+1} \\ &\leq \frac{n}{n+1}E\left[\left|\mathbb{G}_n(x) - x\right|\right] + 2x \\ &\leq 4x, \end{aligned} \tag{48}$$

where the first inequality uses the triangle inequality and $\overline{\mathbb{G}}_n(x) = \widetilde{\mathbb{G}_n}(x)$ with probability one, the second follows by $p_n(x) \leq x$ and the third by (19).

Now, let us bound the first term of (47). Since $\mathbb{H}_n^{-1}(i/n) \sim$Beta$(i, n+1-i)$, we have $E\left[\mathbb{H}_n^{-1}(i/n)\right] = \frac{i}{n+1}$. By the law of iterated expectations and under Assumption 3, we have

$$E\left[\left|\mathbb{H}_n^{-1} \circ \overline{\mathbb{G}}_n(x) - \frac{n}{n+1}\overline{\mathbb{G}}_n(x)\right|\right] = \sum_{i=1}^{n} E\left[\left|\mathbb{H}_n^{-1}\left(\frac{i}{n}\right) - E\left[\mathbb{H}_n^{-1}\left(\frac{i}{n}\right)\right]\right|\right] \Pr\left(\overline{\mathbb{G}}_n(x) = \frac{i}{n}\right).$$

Then, using that $\overline{\mathbb{G}}_n = \widetilde{\mathbb{G}_n}$ with probability one, we obtain

$$\Pr\left(\overline{\mathbb{G}}_n(x) = \frac{i}{n}\right) = \mathbb{1}\{2 \leq i \leq n-1\}\binom{n}{i-1}x^{i-1}(1-x)^{n+1-i} + \mathbb{1}\{i = n\}x^{n-1}. \tag{49}$$

Let $B(\cdot,\cdot)$ denote the beta function and $Z \sim$Beta(a,b). Then, for all $(x,y) \in \mathbb{N}^{*2}$, we have

$$E\left[\left|Z - E[Z]\right|\right] = \frac{2a^a b^b}{B(a,b)(a+b)^{a+b+1}}, \tag{50}$$

$$\frac{x+y}{xyB(x,y)} = \binom{x+y}{x} \quad \forall (x,y) \in \mathbb{N}^*. \tag{51}$$

As a result,

$$
\begin{aligned}
&E\left[\left|\mathbb{H}_n^{-1}\circ\overline{\mathbb{G}}_n(x)-\frac{n}{n+1}\overline{\mathbb{G}}_n(x)\right|\right]\\
&\leq x^{n-1}+\frac{2}{(n+1)^{n+1}}\sum_{i=2}^{n-1}\frac{i^i(n+1-i)^{n+1-i}}{B(i,n+1-i)}\binom{n}{i-1}x^{i-1}(1-x)^{n+1-i}\\
&=x^{n-1}+2x(1-x)\sum_{j=0}^{n-3}\frac{(j+2)^{j+2}}{(n+1)^{n+1}(j+1)}\binom{n}{j}x^j(1-x)^{n-j}\\
&\leq 3x. \qquad (52)
\end{aligned}
$$

The first inequality follows using $E\left[\left|\mathbb{H}_n^{-1}(1)-E\left[\mathbb{H}_n^{-1}(1)\right]\right|\right]\leq 1$, (49) and (50). The equality is obtained by applying the change $j=i-2$ in the sum and (51). The last inequality uses $(j+2)^{j+2}\leq(n+1)^{n+1}(j+1)$ for all $j\in\{0,\ldots,n-3\}$. Given (47) and (48), (44) follows by (52).

We now turn to Equation (45). Because $(\overline{\mathbb{G}}_n(x),\mathbb{H}_n(1/2))\in\{0,1/n,...,1\}^2$ and $|\overline{\mathbb{G}}_n(x)-\mathbb{G}_n(x)|\leq 1/n$, $\overline{\mathbb{G}}_n(x)>\mathbb{H}_n(1/2)$ implies $\mathbb{G}_n(x)\geq\mathbb{H}_n(1/2)$. Moreover, $\mathbb{H}_n^{-1}(a)<b$ iff $a<\mathbb{H}_n(b)$. Then, by Kiefer's and Hoeffding's inequalities,

$$
\begin{aligned}
\Pr(\mathbb{H}_n^{-1}\circ\overline{\mathbb{G}}_n(x)>1/2)&=E\left[\Pr(\overline{\mathbb{G}}_n(x)>\mathbb{H}_n(1/2)|\mathbb{H}_n(1/2))\right]\\
&\leq E\left[\Pr(\mathbb{G}_n(x)\geq\mathbb{H}_n(1/2)|\mathbb{H}_n(1/2))\right]\\
&\leq E\left[(xe)^{n(\mathbb{H}_n(1/2)-x)^2}\right]\\
&\leq xe+\Pr\left(\mathbb{H}_n(1/2)-x<1/\sqrt{n}\right)\\
&\leq xe+\exp\left(-2(\sqrt{n}(x-1/2)+1)^2\right)\\
&=xe+\exp\left(-2n(x-1/2+1/\sqrt{n})^2\right).
\end{aligned}
$$

Let $\overline{\delta}\in(0,e^{-1}]$ and fix $\delta=\overline{\delta}/2$ and $n_0\geq(2/\overline{\delta})^2$. Then, for all $n\geq n_0$ and any $0<x\leq\delta$, we have

$$
\begin{aligned}
\left|x-1/2+\frac{1}{\sqrt{n}}\right|&=\frac{1}{2}-(x+1/\sqrt{n})\\
&\geq\frac{1}{2}-\overline{\delta}.
\end{aligned}
$$

Let $C = 2(1/2 - \overline{\delta})^2$ and suppose first that $x \geq \exp(A - Cn)$ for some A. Then some algebra shows that

$$\Pr(\mathbb{H}_n^{-1} \circ \overline{\mathbb{G}}_n(x) > 1/2) \lesssim x.$$

Now assume that $x < \exp(A - Cn)$. Then

$$\begin{aligned}
\Pr(\mathbb{H}_n^{-1} \circ \overline{\mathbb{G}}_n(x) > 1/2) \leq & \Pr(\mathbb{G}_n(x) \geq 1/n) \\
= & 1 - (1 - x)^n \\
\leq & nx \\
\leq & \frac{A - \ln x}{C} x.
\end{aligned}$$

For any $\eta > 0$, we have $-\ln x \lesssim x^{-\eta}$. Thus, $\Pr(\mathbb{H}_n^{-1} \circ \overline{\mathbb{G}}_n(x) > 1/2) \lesssim x^{1-\eta}$. □

Lemma 5 *(Bounds on moments involving F_U) Suppose that Assumption 2 holds and a random variable $Q_n(x)$ satisfies, for some $0 < \delta < 1/2$ and all $0 < x < \delta$, $E[|Q_n(x) - x|] \lesssim x$ and $\Pr(Q_n(x) > 1/2) \lesssim x^{1-b_1}$. Then, for such $x \in (0, \delta)$, $E[|F_U(Q_n(x)) - F_U(y)|] \lesssim x^{1-b_1}$. The latter inequality holds if we replace x by $1 - x$, using possibly another δ.*

Proof of Lemma 5: first, remark that for $x < 1/2$, $F_U(x) \lesssim x^{1-b_1}$. Then,

$$\begin{aligned}
E[|F_U(x) - F_U(Q_n(x))|] \leq & E[\mathbb{1}\{x > Q_n(x)\} |F_U(x) - F_U(Q_n(x))|] + \Pr(Q_n(x) > 1/2) \\
& + E\left[\mathbb{1}\{Q_n(x) \in [x, 1/2]\} |F_U(x) - F_U(Q_n(x))|\right] \\
\lesssim & F_U(x) + x^{1-b_1} + E\left[\mathbb{1}\{Q_n(x) \in [x, 1/2]\} |F_U(x) - F_U(Q_n(x))|\right] \\
\lesssim & x^{1-b_1} + E\left[\mathbb{1}\{Q_n(x) \in [x, 1/2]\} |F_U(x) - F_U(Q_n(x))|\right].
\end{aligned}$$

Now, if $Q_n(x) \in [x, 1/2)$, by the mean value theorem, there exists $X_n \in (x, 1/2)$ such that

$$F_U(x) - F_U(Q_n(x)) = f_U(X_n)(x - Q_n(x)).$$

Moreover, by Assumption 2 and $x < \delta$, $f_U(X_n) \lesssim x^{-b_1}$. Then, using $E[|Q_n(x) - x|] \lesssim x$,

$$E\left[\mathbb{1}\{Q_n(x) \in [x, 1/2]\} |F_U(x) - F_U(Q_n(x))|\right] \lesssim x^{1-b_1}.$$

The result follows. □

D Proofs details

Proof of the case $1-x$ in Lemma 4: We want to prove that there exists $\delta \in (1/2, 1)$ and $n_0 \in \mathbb{N}$ such that for all $\delta < x < 1$ and all $n \geq n_0$,

$$E\left[\left|\mathbb{H}_n^{-1} \circ \overline{\mathbb{G}}_n(1-x) - (1-x)\right|\right] \lesssim (1-x). \tag{53}$$

Moreover, for any $\eta > 0$, there exists n_1 such that for all $n \geq n_1$ and for all $\delta < x < 1$

$$\Pr(\mathbb{H}_n^{-1} \circ \overline{\mathbb{G}}_n(1-x) > 1/2) \lesssim (1-x)^{1-\eta}. \tag{54}$$

Proof: We first establish (53). Replace x by $1-x$ in the proof of Lemma 4 until Equation (48). The latter holds because Equation (19) can be obtained for $1-x$ by the same reasoning than before (use Cauchy-Schwarz inequality when $(1-x) \leq 1/n$ and the third part of Theorem 1 in Berend and Kontorovich (2013) when $(1-x) > 1/n$). By noticing that all the steps to show (52) remain valid when replacing x by $1-x$, this proves (53).

We now turn to Equation (54). Following the same reasoning than for the proof of (45), we obtain

$$\Pr(\mathbb{H}_n^{-1} \circ \overline{\mathbb{G}}_n(x) > 1/2) \leq (1-x)e + \exp\left(-2n(1-x-1/2+1/\sqrt{n})^2\right).$$

Let $\bar{\delta} \in (1-e^{-1}, 1)$ and fix $\delta = \bar{\delta}/2$ and $n_0 \geq (2/\bar{\delta})^2$. Then, for all $n \geq n_0$ and any $\delta < x < 1$, we have

$$\begin{aligned} \left|1 - x - 1/2 + \frac{1}{\sqrt{n}}\right| &= \frac{1}{2} - \left(1 - x + \frac{1}{\sqrt{n}}\right) \\ &\geq \frac{1}{2} - \bar{\delta}. \end{aligned}$$

All the remaining arguments hold by replacing x by $1-x$. This proves (54). $\square$