Supplementary Appendix

Distributional Effects with Two-Sided Measurement Error: An Application to Intergenerational Income Mobility*

Brantly Callaway[†] Tong Li[‡] Irina Murtazashvili[§] Emmanuel S. Tsyawo[¶]

September 20, 2025

This Supplementary Appendix contains proofs for several of the results from the main text as well as some supplementary results. Appendix SA contains some additional theoretical results as well as proofs for some of the results provided in the main text. Appendix SB provides conditions under which our main target parameters have a causal interpretation. Appendix SC explains how to extend our results in the presence of life-cycle measurement error. Finally, Appendix SD provides additional results for our application on intergenerational mobility.

SA Additional Theoretical Results and Proofs

The first part of this section contains the proof of Lemma 1 from the main text and then states and proves Lemma S1 which is used in the proof of Theorem 4. The second part of this section proves parts (b)-(f) of Theorem 4.

SA.1 Useful Lemmas

The following lemmas are supporting results used in the proofs of the results in the main text.

---

*The U.S. Securities and Exchange Commission disclaims responsibility for any private publication or statement of any SEC employee or Commissioner. This paper expresses the author's views and does not necessarily reflect those of the Commission, the Commissioners, or members of the staff.

[†]John Munro Godfrey, Sr. Department of Economics, University of Georgia, B422 Amos Hall, Athens, GA 30605. Email: brantly.callaway@uga.edu

[‡]Department of Economics, Vanderbilt University, VU Station B #351819, Nashville, TN 37235. Email: tong.li@vanderbilt.edu

[§]Division of Economic and Risk Analysis, U.S. Securities and Exchange Commission, 1617 John F Kennedy Blvd #520, Philadelphia, PA 19103. Email: murtazashviliirina@gmail.com

[¶]Department of Economics, Finance and Legal Studies, Culverhouse College of Business, University of Alabama. Email: estsyawo@gmail.com

Proof of Lemma 1. By the triangle inequality:

$$
\begin{aligned}
&\Big|Q_{n,\tau}\big(\beta(\tau),\sigma\big) - Q_{\tau}\big(\beta(\tau),\sigma\big)\Big| \\
&\leq \frac{1}{n}\sum_{i=1}^{n}\Big|\int_{\mathcal{U}}\rho_\tau\left(Y_i - u - X_i'\beta(\tau)\right)\Big(\widehat{f}^{(S)}_{U_{Y^*}|Y,X}(u\mid Y_i,X_i;\sigma) - f_{U_{Y^*}|Y,X}(u\mid Y_i,X_i;\sigma)\Big)du\Big| \\
&+\Big|\frac{1}{n}\sum_{i=1}^{n}\Big\{\int_{\mathcal{U}}\rho_\tau\left(Y_i - u - X_i'\beta(\tau)\right) f_{U_{Y^*}|Y,X}(u\mid Y_i,X_i;\sigma)\,du \\
&\qquad - \mathbb{E}\Big[\int_{\mathcal{U}}\rho_\tau\left(Y_i - u - X_i'\beta\right) f_{U_{Y^*}|Y,X}(u\mid Y_i,X_i;\sigma)\,du\Big]\Big\}\Big|.
\end{aligned}
$$

Define $\mathcal{E}_\tau := Y_i - X_i'\beta(\tau)$. Note that $\rho_\tau(w) \leq |w|$ and $\rho_\tau(w)$ is 1-Lipschitz.

First, by the Cauchy-Schwarz inequality,

$$
\begin{aligned}
&\Big|\int_{\mathcal{U}}\rho_\tau\left(Y_i - u - X_i'\beta(\tau)\right)\Big(\widehat{f}^{(S)}_{U_{Y^*}|Y,X}(u\mid Y_i,X_i;\sigma) - f_{U_{Y^*}|Y,X}(u\mid Y_i,X_i;\sigma)\Big)du\Big| \\
&= \Big|\int_{\mathcal{U}}\left(\rho_\tau\left(\mathcal{E}_\tau - u\right) - \rho_\tau(\mathcal{E}_\tau) + \rho_\tau(\mathcal{E}_\tau)\right)\Big(\widehat{f}^{(S)}_{U_{Y^*}|Y,X}(u\mid Y_i,X_i;\sigma) - f_{U_{Y^*}|Y,X}(u\mid Y_i,X_i;\sigma)\Big)\,du\Big| \\
&\leq \Big|\int_{\mathcal{U}}\left(\rho_\tau\left(\mathcal{E}_\tau - u\right) - \rho_\tau(\mathcal{E}_\tau)\right)\cdot\Big(\widehat{f}^{(S)}_{U_{Y^*}|Y,X}(u\mid Y_i,X_i;\sigma) - f_{U_{Y^*}|Y,X}(u\mid Y_i,X_i;\sigma)\Big)\,du\Big| \\
&+ \rho_\tau(\mathcal{E}_\tau)\Big|\int_{\mathcal{U}}\Big(\widehat{f}^{(S)}_{U_{Y^*}|Y,X}(u\mid Y_i,X_i;\sigma) - f_{U_{Y^*}|Y,X}(u\mid Y_i,X_i;\sigma)\Big)\,du\Big| \\
&= \Big|\int_{\mathcal{U}}\left(\rho_\tau\left(\mathcal{E}_\tau - u\right) - \rho_\tau(\mathcal{E}_\tau)\right)\cdot\Big(\widehat{f}^{(S)}_{U_{Y^*}|Y,X}(u\mid Y_i,X_i;\sigma) - f_{U_{Y^*}|Y,X}(u\mid Y_i,X_i;\sigma)\Big)\,du\Big|
\end{aligned}
$$

since $\int_{\mathcal{U}}\Big(\widehat{f}^{(S)}_{U_{Y^*}|Y,X}(u\mid Y_i,X_i;\sigma) - f_{U_{Y^*}|Y,X}(u\mid Y_i,X_i;\sigma)\Big)\,du = 0$. Let $\phi_\tau(u) := \rho_\tau(\mathcal{E}_\tau - u) - \rho_\tau(\mathcal{E}_\tau)$, and observe that since ρ_τ is 1–Lipschitz, $|\phi_\tau(u)| \leq |u|$; this upper bound holds uniformly in $\tau \in (0,1)$. Further, under Assumption 10(c), $\mathbb{E}[\phi_\tau^2(U(\sigma))|Y,X] \leq \int_{\mathcal{U}} u^2 f_{U_{Y^*}|Y,X}(u \mid Y,X;\sigma) \leq C\,a.s.$ uniformly in $\sigma \in \Gamma_\sigma$. Under Assumption 9,

$$
\begin{aligned}
&\int_{\mathcal{U}}\left(\rho_\tau\left(\mathcal{E}_\tau - u\right) - \rho_\tau(\mathcal{E}_\tau)\right)\cdot\Big(\widehat{f}^{(S)}_{U_{Y^*}|Y,X}(u\mid Y_i,X_i;\sigma) - f_{U_{Y^*}|Y,X}(u\mid Y_i,X_i;\sigma)\Big)\,du \\
&\qquad = \frac{1}{S}\sum_{s=1}^{S}\left(\phi_\tau(U_{is}(\sigma)) - \mathbb{E}[\phi_\tau(U_i(\sigma))\mid Y_i,X_i]\right) = O_p(S^{-1/2}) = o_p(n^{-1/2})
\end{aligned}
$$

by Prakasa Rao (2009, Theorem 10, eqn. 64), Chebyshev's inequality, that a β-mixing sequence is strongly mixing, and the condition $n/S = o(1)$.

Second, since ρ_τ is 1-Lipschitz, it follows from the triangle inequality, the Lyapunov inequality, the Schwarz inequality, Assumption 8, and Assumption 10 that

$$
\begin{aligned}
\mathbb{E}\Big[\int_{\mathcal{U}}\rho_\tau\left(Y - u - X'\beta(\tau)\right) f_{U_{Y^*}|Y,X}(u\mid Y,X;\sigma)\,du\Big] &= \mathbb{E}\Big[\mathbb{E}\big[\rho_\tau\big(\mathcal{E}_\tau - U(\sigma)\big)\mid Y,X\big]\Big] \\
&\leq \sup_{\tau\in\mathcal{T}}\mathbb{E}[|Y - X'\beta(\tau)|] + \mathbb{E}\big[\mathbb{E}[(|U(\sigma)|)|Y,X]\big] \\
&\leq \mathbb{E}[|Y|] + \mathbb{E}[\|X\|]\cdot\sup_{\tau\in\mathcal{T}}\|\beta(\tau)\| + \mathbb{E}\big[\big(\mathbb{E}[U(\sigma)^2|Y,X]\big)^{1/2}\big] \\
&< \infty.
\end{aligned}
$$

Conclude under Assumption 4 and the strong law of large numbers that

$$\frac{1}{n}\sum_{i=1}^{n}\Big\{\int_{\mathcal{U}}\rho_\tau\left(Y_i-u-X_i'\beta(\tau)\right)f_{U_{Y^*}|Y,X}(u\mid Y_i,X_i;\sigma)\,du \\ -\mathbb{E}\Big[\int_{\mathcal{U}}\rho_\tau\left(Y_i-u-X_i'\beta\right)f_{U_{Y^*}|Y,X}(u\mid Y_i,X_i;\sigma)\,du\Big]\Big\}\xrightarrow{a.s.}0.$$

Combining both parts above concludes the proof of the assertion as claimed.

□

Lemma S1. *Under Assumptions 1 to 5 and 7 to 10,* $\sqrt{n}\big(\widehat{\mathrm{F}}_{Y^*}^{-1}(r)-F_{Y^*}^{-1}(r)\big)$ *has the following representation:*

$$\sqrt{n}\big(\widehat{\mathrm{F}}_{Y^*}^{-1}(r)-F_{Y^*}^{-1}(r)\big) \\ =-\frac{1}{f_{Y^*}(\mathrm{F}_{Y^*}^{-1}(r))}\mathbb{E}\big[\mathbb{M}_{FY}^{L}\big(\mathrm{F}_{Y^*}^{-1}(r),X\big)\big]'\sqrt{n}\begin{bmatrix}\widehat{\boldsymbol{\beta}}_{Y^*}^{L}-\boldsymbol{\beta}_{Y^*}^{L}\\ \widehat{\sigma}_{Y^*}-\sigma_{Y^*}\end{bmatrix} \\ -\frac{1}{f_{Y^*}(\mathrm{F}_{Y^*}^{-1}(r))}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big(\mathrm{F}_{Y^*|X}(\mathrm{F}_{Y^*}^{-1}(r)\mid X_i)-\mathbb{E}[\mathrm{F}_{Y^*|X}(\mathrm{F}_{Y^*}^{-1}(r)\mid X_i)]\big)+o_p(1).$$

Proof. Let $y=\mathrm{F}_{Y^*}^{-1}(r)$, then $\widehat{\mathrm{F}}_{Y^*}(\widehat{y})=r=\mathrm{F}_{Y^*}(y)$, then by the MVT, $\widehat{\mathrm{F}}_{Y^*}(\widehat{y})-\mathrm{F}_{Y^*}(y)=0=\widehat{\mathrm{F}}_{Y^*}(y)-\mathrm{F}_{Y^*}(y)+\widehat{f}_{Y^*}(\bar{y})(\widehat{y}-y)$. Thus by the Law of Iterated Expectations (LIE),

$$\sqrt{n}\big(\widehat{\mathrm{F}}_{Y^*}^{-1}(r)-\mathrm{F}_{Y^*}^{-1}(r)\big)=\sqrt{n}(\widehat{y}-y) \\ =-\frac{1}{\widehat{f}_{Y^*}(\bar{y})}\sqrt{n}\big(\widehat{\mathrm{F}}_{Y^*}(y)-\mathrm{F}_{Y^*}(y)\big) \\ =-\frac{1}{\widehat{f}_{Y^*}(\bar{y})}\frac{1}{n}\sum_{i=1}^{n}\Big\{\sqrt{n}\big(\widehat{\mathrm{F}}_{Y^*|X}(y\mid X_i)-\mathrm{F}_{Y^*|X}(y\mid X_i)\big)+\sqrt{n}\big(\mathrm{F}_{Y^*|X}(y\mid X_i)-\mathbb{E}[\mathrm{F}_{Y^*|X}(y\mid X_i)]\big)\Big\} \\ =-\frac{1}{f_{Y^*}(\mathrm{F}_{Y^*}^{-1}(r))}\frac{1}{n}\sum_{i=1}^{n}\Big\{\sqrt{n}\big(\widehat{\mathrm{F}}_{Y^*|X}(\mathrm{F}_{Y^*}^{-1}(r)\mid X_i)-\mathrm{F}_{Y^*|X}(\mathrm{F}_{Y^*}^{-1}(r)\mid X_i)\big)\Big\} \\ -\frac{1}{f_{Y^*}(\mathrm{F}_{Y^*}^{-1}(r))}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big(\mathrm{F}_{Y^*|X}(\mathrm{F}_{Y^*}^{-1}(r)\mid X_i)-\mathbb{E}[\mathrm{F}_{Y^*|X}(\mathrm{F}_{Y^*}^{-1}(r)\mid X_i)]\big)+o_p(1).$$

In addition to Assumption 7 and Corollary 1,

$$\sqrt{n}\big(\widehat{\mathrm{F}}_{Y^*|X}(\mathrm{F}_{Y^*}^{-1}(r)\mid X_i)-\mathrm{F}_{Y^*|X}(\mathrm{F}_{Y^*}^{-1}(r)\mid X_i)\big)=\mathbb{M}_{FY}^{L}\big(\mathrm{F}_{Y^*}^{-1}(r),X_i\big)'\sqrt{n}\begin{bmatrix}\widehat{\boldsymbol{\beta}}_{Y^*}^{L}-\boldsymbol{\beta}_{Y^*}^{L}\\ \widehat{\sigma}_{Y^*}-\sigma_{Y^*}\end{bmatrix}+o_p(1)$$

whence

$$\frac{1}{n}\sum_{i=1}^{n}\Big\{\sqrt{n}\big(\widehat{\mathrm{F}}_{Y^*|X}(\mathrm{F}_{Y^*}^{-1}(r)\mid X_i)-\mathrm{F}_{Y^*|X}(\mathrm{F}_{Y^*}^{-1}(r)\mid X_i)\big)\Big\}=\mathbb{E}\big[\mathbb{M}_{FY}^{L}\big(\mathrm{F}_{Y^*}^{-1}(r),X\big)\big]'\sqrt{n}\begin{bmatrix}\widehat{\boldsymbol{\beta}}_{Y^*}^{L}-\boldsymbol{\beta}_{Y^*}^{L}\\ \widehat{\sigma}_{Y^*}-\sigma_{Y^*}\end{bmatrix}+o_p(1)$$

thanks to the Weak Law of Large Numbers (WLLN), the CMT and Theorem 2. From the foregoing,

$$\sqrt{n}\big(\widehat{\mathrm{F}}_{Y^*}^{-1}(r)-F_{Y^*}^{-1}(r)\big)$$

$$= -\frac{1}{f_{Y^*}(\mathrm{F}_{Y^*}^{-1}(r))}\mathbb{E}\big[\mathbb{M}_{FY}^{L}\big(\mathrm{F}_{Y^*}^{-1}(r), X\big)\big]'\sqrt{n}\begin{bmatrix}\widehat{\boldsymbol{\beta}}_{Y^*}^{L} - \boldsymbol{\beta}_{Y^*}^{L}\\ \widehat{\sigma}_{Y^*} - \sigma_{Y^*}\end{bmatrix}$$
$$- \frac{1}{f_{Y^*}(\mathrm{F}_{Y^*}^{-1}(r))}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big(\mathrm{F}_{Y^*|X}(\mathrm{F}_{Y^*}^{-1}(r) \mid X_i) - \mathbb{E}[\mathrm{F}_{Y^*|X}(\mathrm{F}_{Y^*}^{-1}(r) \mid X_i)]\big) + o_p(1).$$

□

SA.2 Proof of Theorem 4

The proofs of parts (b) through (f) of Theorem 4 are organized in the following lemmas.

SA.2.1 Proof of Theorem 4(b)

The following lemma provides a proof of part (b) of Theorem 4 from the main text.

Lemma S2. *Suppose Assumptions 1 to 10 hold, then*

$$\sqrt{n}\big(\widehat{\mathrm{F}}_{Y^*|T^*X}(y|t,x) - \mathrm{F}_{Y^*|T^*X}(y|t,x)\big) \xrightarrow{d} \mathcal{N}\big(0,\, \sigma(\mathrm{F}_{Y^*|T^*X}(y|t,x))\big).$$

Proof. Recall $\mathrm{F}_{Y^*|T^*X}(y|t,x) = C_{2|X}\big(\mathrm{F}_{Y^*|X}(y \mid x), \mathrm{F}_{T^*|X}(t \mid x)\big) =: C_2\big(\mathrm{F}_{Y^*|X}(y \mid x), \mathrm{F}_{T^*|X}(t \mid x);\, \delta\big)$ with $C_{2|X}(r,s) = \frac{\partial C_{Y^*T^*|X}(r,s)}{\partial s}$. The estimator is given by $\widehat{\mathrm{F}}_{Y^*|T^*X}(y|t,x) = C_2\big(\widehat{\mathrm{F}}_{Y^*|X}(y \mid x), \widehat{\mathrm{F}}_{T^*|X}(t \mid x);\, \widehat{\delta}\big)$ whence

$$\begin{aligned}\sqrt{n}\big(\widehat{\mathrm{F}}_{Y^*|T^*X}(y|t,x) &- \mathrm{F}_{Y^*|T^*X}(y|t,x)\big)\\ &= \sqrt{n}\Big(C_2\big(\widehat{\mathrm{F}}_{Y^*|X}(y \mid x), \widehat{\mathrm{F}}_{T^*|X}(t \mid x);\, \widehat{\delta}\big) - C_2\big(\mathrm{F}_{Y^*|X}(y \mid x), \mathrm{F}_{T^*|X}(t \mid x);\, \delta\big)\Big)\\ &=: \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\psi_i^{(b)} + o_p(1)\end{aligned}$$

where $\sigma(\mathrm{F}_{Y^*|T^*X}(y|t,x)) = \mathbb{E}\big[(\psi^{(b)})^2\big]$ following arguments analogous to that of Lemma 2. □

SA.2.2 Proof of Theorem 4(c)

The following lemma provides a proof of part (c) of Theorem 4 from the main text.

Lemma S3. *Suppose Assumptions 1 to 10 hold, then* $\sqrt{n}\big(\widehat{Q}_{Y^*|T^*X}(\tau \mid t,x) - Q_{Y^*|T^*X}(\tau \mid t,x)\big) \xrightarrow{d} \mathcal{N}\big(0,\, \sigma(Q_{Y^*|T^*X}(\tau \mid t,x))\big)$.

Proof. Recall $Q_{Y^*|T^*X}(\tau|t,x) = Q_{Y^*|X}\Big(C_{2;1|X}^{-1}\big(\tau; \mathrm{F}_{T^*|X}(t \mid x) \mid x\big)\Big) = x'\beta_{Y^*}\Big(C_{2;1|X}^{-1}\big(\tau,\, \mathrm{F}_{T^*|X}(t \mid x);\, \delta\big)\Big)$ under Assumption 2 where $C_{2;1|X}^{-1}(\cdot;\cdot)$ is the inverse of $C_{2|X}$ with respect to its first argument. The estimator is given by $\widehat{Q}_{Y^*|T^*X}(\tau|t,x) = x'\widehat{\beta}_{Y^*}\Big(C_{2;1|X}^{-1}\big(\tau,\, \widehat{\mathrm{F}}_{T^*|X}(t \mid x);\, \widehat{\delta}\big)\Big)$. Consider the following decomposition

$$\begin{aligned}\sqrt{n}\big(\widehat{Q}_{Y^*|T^*X}(\tau|t,x) &- Q_{Y^*|T^*X}(\tau|t,x)\big)\\ &= \sqrt{n}x'\Big(\widehat{\beta}_{Y^*}\big(C_{2;1|X}^{-1}\big(\tau; \widehat{\mathrm{F}}_{T^*|X}(t \mid x);\, \widehat{\delta}\big)\big) - \beta_{Y^*}\big(C_{2;1|X}^{-1}\big(\tau; \mathrm{F}_{T^*|X}(t \mid x);\, \delta\big)\big)\Big)\\ &= \sqrt{n}x'\big(\widehat{\beta}_{Y^*}(\widehat{\tau}_Q) - \widehat{\beta}_{Y^*}(\tau_Q)\big) + \sqrt{n}x'\big(\widehat{\beta}_{Y^*}(\tau_Q) - \beta_{Y^*}(\tau_Q)\big)\\ &= x'\widehat{\beta}_{Y^*}^{\partial}(\bar{\tau}_Q)\sqrt{n}(\widehat{\tau}_Q - \tau_Q) + \sqrt{n}x'\big(\widehat{\beta}_{Y^*}(\tau_Q) - \beta_{Y^*}(\tau_Q)\big) \qquad \text{(S1)}\end{aligned}$$

where $\tau_Q := C^{-1}_{2;1|X}(\tau;\, \tau_t),\ \tau_t := \mathrm{F}_{T^*|X}(t \mid x)$, and notice that $C_{2|X}(\tau_Q,\, \tau_t;\, \delta) = \tau = C_{2|X}(\widehat{\tau}_Q,\, \widehat{\tau}_t;\, \widehat{\delta})$. Thanks to the MVT,

$$
\begin{aligned}
0 &= C_{2|X}(\widehat{\tau}_Q,\, \widehat{\tau}_t;\, \widehat{\delta}) - C_{2|X}(\tau_Q,\, \tau_t;\, \delta) \\
&= \big(C_{2|X}(\widehat{\tau}_Q,\, \widehat{\tau}_t;\, \widehat{\delta}) - C_{2|X}(\tau_Q,\, \widehat{\tau}_t;\, \widehat{\delta})\big) + \big(C_{2|X}(\tau_Q,\, \widehat{\tau}_t;\, \widehat{\delta}) - C_{2|X}(\tau_Q,\, \tau_t;\, \widehat{\delta})\big) \\
&+ \big(C_{2|X}(\tau_Q,\, \tau_t;\, \widehat{\delta}) - C_{2|X}(\tau_Q,\, \tau_t;\, \delta)\big) \\
&= \partial_1 C_{2|X}(\bar{\tau}_Q,\, \widehat{\tau}_t;\, \widehat{\delta})(\widehat{\tau}_Q - \tau_Q) + \partial_2 C_{2|X}(\tau_Q,\, \bar{\tau}_t;\, \widehat{\delta})(\widehat{\tau}_t - \tau_t) + \partial_\delta C_{2|X}(\tau_Q,\, \tau_t;\, \bar{\delta})(\widehat{\delta} - \delta) \\
&= c_{Y^*T^*|X}(\bar{\tau}_Q,\, \widehat{\tau}_t;\, \widehat{\delta})(\widehat{\tau}_Q - \tau_Q) + \partial_2 C_{2|X}(\tau_Q,\, \bar{\tau}_t;\, \widehat{\delta})(\widehat{\tau}_t - \tau_t) + \partial_\delta C_{2|X}(\tau_Q,\, \tau_t;\, \bar{\delta})(\widehat{\delta} - \delta)
\end{aligned}
$$

by the definition of the conditional copula density, i.e., $c_{Y^*T^*|X}(r,\, s;\, \cdot) = \partial_1(\partial_2 C_{Y^*T^*|X}(r,\, s;\, \cdot)) = \partial_1(C_{2|X}(r,\, s;\, \cdot))$.

From the foregoing and that the conditional copula is strictly positive (Assumption 6), apply the CMT and Theorem 2 to obtain

$$
\begin{aligned}
\sqrt{n}(\widehat{\tau}_Q - \tau_Q) &= -\big(c_{Y^*T^*|X}(\bar{\tau}_Q,\, \widehat{\tau}_t;\, \widehat{\delta})\big)^{-1}\Big(\partial_2 C_{2|X}(\tau_Q,\, \bar{\tau}_t;\, \widehat{\delta})\sqrt{n}(\widehat{\tau}_t - \tau_t) + \partial_\delta C_{2|X}(\tau_Q,\, \tau_t;\, \bar{\delta})'\sqrt{n}(\widehat{\delta} - \delta)\Big) \\
&= -\big(c_{Y^*T^*|X}(\tau_Q,\, \tau_t;\, \delta)\big)^{-1}\partial_2 C_{2|X}(\tau_Q,\, \tau_t;\, \delta)\sqrt{n}(\widehat{\mathrm{F}}_{T^*|X}(t \mid x) - \mathrm{F}_{T^*|X}(t \mid x)) \\
&\quad - \big(c_{Y^*T^*|X}(\tau_Q,\, \tau_t;\, \delta)\big)^{-1}\partial_\delta C_{2|X}(\tau_Q,\, \tau_t;\, \delta)'\sqrt{n}(\widehat{\delta} - \delta) + o_p(1).
\end{aligned}
$$

Substituting terms into (S1) using the representation in (A10), it follows from Theorem 2 and the CMT that

$$
\begin{aligned}
&\sqrt{n}\big(\widehat{Q}_{Y^*|T^*X}(\tau|t,x) - Q_{Y^*|T^*X}(\tau|t,x)\big) \\
&\quad= \frac{1}{f_{Y^*|X}\big(C^{-1}_{2;1|X}(\tau;\, \mathrm{F}_{T^*|X}(t|x)) \mid x\big)}\sqrt{n}(\widehat{\tau}_Q - \tau_Q) \\
&\quad+ \sum_{\ell=1}^{L} \omega_{\ell,\beta}(\tau_Q)\, x'\sqrt{n}\big(\widehat{\beta}_{Y^*}(\tau_\ell) - \beta_{Y^*}(\tau_\ell)\big) + o_p(1) \\
&\quad= \underbrace{-\frac{\big(c_{Y^*T^*|X}(\tau_Q,\, \tau_t;\, \delta)\big)^{-1}}{f_{Y^*|X}\big(C^{-1}_{2;1|X}(\tau;\, \mathrm{F}_{T^*|X}(t|x)) \mid x\big)}\partial_2 C_{2|X}(\tau_Q,\, \tau_t;\, \delta)}_{\mathbb{M}^L_{Q,F}(\tau,t,x)}\sqrt{n}(\widehat{\mathrm{F}}_{T^*|X}(t \mid x) - \mathrm{F}_{T^*|X}(t \mid x)) \\
&\quad\underbrace{-\frac{\big(c_{Y^*T^*|X}(\tau_Q,\, \tau_t;\, \delta)\big)^{-1}}{f_{Y^*|X}\big(C^{-1}_{2;1|X}(\tau;\, \mathrm{F}_{T^*|X}(t|x)) \mid x\big)}\partial_\delta C_{2|X}(\tau_Q,\, \tau_t;\, \delta)'}_{\mathbb{M}^L_{Q,\delta}(\tau,t,x)'}\sqrt{n}(\widehat{\delta} - \delta) \\
&\quad+ \underbrace{\Big[\omega_{1,\beta}(\tau_Q)x', \ldots, \omega_{L,\beta}(\tau_Q)x',\, 0\Big]}_{\mathbb{M}^L_{Q,\beta}(\tau,t,x)'}\sqrt{n}\begin{bmatrix}\widehat{\boldsymbol{\beta}}^L_{Y^*} - \boldsymbol{\beta}^L_{Y^*} \\ \widehat{\sigma}_{Y^*} - \sigma_{Y^*}\end{bmatrix} + o_p(1) \\
&\quad= \mathbb{M}^L_{Q,F}(\tau,t,x)\sqrt{n}(\widehat{\mathrm{F}}_{T^*|X}(t \mid x) - \mathrm{F}_{T^*|X}(t \mid x)) + \mathbb{M}^L_{Q,\delta}(\tau,t,x)'\sqrt{n}(\widehat{\delta} - \delta) \\
&\quad+ \mathbb{M}^L_{Q,\beta}(\tau,t,x)'\sqrt{n}\begin{bmatrix}\widehat{\boldsymbol{\beta}}^L_{Y^*} - \boldsymbol{\beta}^L_{Y^*} \\ \widehat{\sigma}_{Y^*} - \sigma_{Y^*}\end{bmatrix} + o_p(1) \\
&\quad=: \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\psi_i^{(c)} + o_p(1) \xrightarrow{d} \mathcal{N}\big(0, \sigma(Q_{Y^*|T^*X}(\tau|t,x))\big)
\end{aligned}
$$

where $\sigma(Q_{Y^*|T^*X}(\tau|t,x)) = \mathbb{E}\big[(\psi^{(c)})^2\big]$. Since a linear combination of asymptotically linear quantities is asymptotically linear, the conclusion follows from Corollary 1 (applied to T analogously), Proposition 2, Theorem 3, and Slutsky's Theorem. □

SA.2.3 Proof of Theorem 4(d)

The following lemma provides a proof of part (d) of Theorem 4 from the main text.

Lemma S4. *Suppose Assumptions 1 to 10 hold, then*

$$\sqrt{n}\big(\widehat{\theta}_{TM}(r_1,r_2,s_1,s_2) - \theta_{TM}(r_1,r_2,s_1,s_2)\big) \xrightarrow{d} \mathcal{N}\big(0,\, \sigma(\theta_{TM}(r_1,r_2,s_1,s_2))\big).$$

Proof. The unconditional joint CDF can be obtained from the conditional copula:

$$\begin{aligned} F_{Y^*T^*}(y,\,t) &= \int_{\mathcal{X}} \mathrm{F}_{Y^*T^*|X}(y,\,t \mid x) d\mathrm{F}_X(x) \\ &= \int_{\mathcal{X}} C_{Y^*T^*|X}\big(\mathrm{F}_{Y^*|X}(y \mid x), \mathrm{F}_{T^*|X}(t \mid x);\, \delta\big)\, d\mathrm{F}_X(x). \end{aligned}$$

Define, for $r = \mathrm{F}_{Y^*}(y)$ and $s = \mathrm{F}_{T^*}(t)$, the induced unconditional copula

$$\begin{aligned} C_{Y^*T^*}(r,s) &= \mathrm{F}_{Y^*T^*}\big(\mathrm{F}_{Y^*}^{-1}(r),\, \mathrm{F}_{T^*}^{-1}(s)\big) \\ &= \int_{\mathcal{X}} C_{Y^*T^*|X}\big(\mathrm{F}_{Y^*|X}(\mathrm{F}_{Y^*}^{-1}(r) \mid x), \mathrm{F}_{T^*|X}(\mathrm{F}_{T^*}^{-1}(s) \mid x);\, \delta\big)\, d\mathrm{F}_X(x) \\ &= \mathbb{E}\Big[C_{Y^*T^*|X}\big(\mathrm{F}_{Y^*|X}(\mathrm{F}_{Y^*}^{-1}(r) \mid X), \mathrm{F}_{T^*|X}(\mathrm{F}_{T^*}^{-1}(s) \mid X);\, \delta\big)\Big]. \end{aligned}$$

Recall $\theta_{TM}(r_1,r_2,s_1,s_2) = \dfrac{C_{Y^*T^*}(r_2,s_2) - C_{Y^*T^*}(r_1,s_2) - C_{Y^*T^*}(r_2,s_1) + C_{Y^*T^*}(r_1,s_1)}{s_2 - s_1}$, and the estimator is given by

$$\widehat{\theta}_{TM}(r_1,r_2,s_1,s_2) = \frac{\widehat{C}_{Y^*T^*}(r_2,s_2;\,\widehat{\delta}) - \widehat{C}_{Y^*T^*}(r_1,s_2;\,\widehat{\delta}) - \widehat{C}_{Y^*T^*}(r_2,s_1;\,\widehat{\delta}) + \widehat{C}_{Y^*T^*}(r_1,s_1;\,\widehat{\delta})}{s_2 - s_1}$$

where

$$\widehat{C}_{Y^*T^*}(r,s) := \frac{1}{n}\sum_{i=1}^{n} C_{Y^*T^*|X}\big(\widehat{\mathrm{F}}_{Y^*|X}(\widehat{\mathrm{F}}_{Y^*}^{-1}(r) \mid X_i), \widehat{\mathrm{F}}_{T^*|X}(\widehat{\mathrm{F}}_{T^*}^{-1}(s) \mid X_i);\, \widehat{\delta}\big).$$

Let $y := \mathrm{F}_{Y^*}^{-1}(r)$, $\hat{y} := \widehat{\mathrm{F}}_{Y^*}^{-1}(r)$, $t := \mathrm{F}_{T^*}^{-1}(s)$, $\hat{t} := \widehat{\mathrm{F}}_{T^*}^{-1}(s)$, $u_i := \mathrm{F}_{Y^*|X}(y \mid X_i)$, $\hat{u}_i := \widehat{\mathrm{F}}_{Y^*|X}(\hat{y} \mid X_i)$, $v_i := \mathrm{F}_{T^*|X}(t \mid X_i)$, and $\hat{v}_i := \widehat{\mathrm{F}}_{T^*|X}(\hat{t} \mid X_i)$. Also, $C(u,v;\delta) := C_{Y^*T^*|X}(u,v;\delta)$ and its partials are $C_1 = \partial C/\partial u,\ C_2 = \partial C/\partial v,\ C_\delta = \partial C/\partial \delta$. Then, by Assumption 6 and the MVT, $C(\hat{u}_i,\hat{v}_i;\hat{\delta}) - C(u_i,v_i;\delta) = \underbrace{C_1(\tilde{u}_i,\hat{v}_i;\hat{\delta})}_{\omega_{FY}(X_i;r,s)}(\hat{u}_i - u_i) + \underbrace{C_2(u_i,\tilde{v}_i;\hat{\delta})}_{\omega_{FT}(X_i;r,s)}(\hat{v}_i - v_i) + \underbrace{C_\delta(u_i,v_i;\tilde{\delta}_i)}_{\omega_\delta(X_i;r,s)}{}'(\hat{\delta} - \delta)$.

Thanks to the MVT and the decomposition in Lemma 2, the following representation holds:

$$\begin{aligned} \sqrt{n}\big(\widehat{C}_{Y^*T^*}(r,s;\widehat{\delta}) - C_{Y^*T^*}(r,s;\delta)\big) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Big\{&C_{Y^*T^*|X}\big(\widehat{\mathrm{F}}_{Y^*|X}(\widehat{\mathrm{F}}_{Y^*}^{-1}(r) \mid X_i), \widehat{\mathrm{F}}_{T^*|X}(\widehat{\mathrm{F}}_{T^*}^{-1}(s) \mid X_i);\, \widehat{\delta}\big) \\ &- \mathbb{E}\Big[C_{Y^*T^*|X}\big(\mathrm{F}_{Y^*|X}(\mathrm{F}_{Y^*}^{-1}(r) \mid X), \mathrm{F}_{T^*|X}(\mathrm{F}_{T^*}^{-1}(s) \mid X);\, \delta\big)\Big]\Big\} \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{n}\sum_{i=1}^{n}\omega_{FY}(X_i; r, s)\,\sqrt{n}\big(\widehat{\mathrm{F}}_{Y^*|X} - \mathrm{F}_{Y^*|X}\big)(\mathrm{F}_{Y^*}^{-1}(r) \mid X_i) \\
&\quad + \frac{1}{n}\sum_{i=1}^{n}\omega_{FT}(X_i; r, s)\,\sqrt{n}\big(\widehat{\mathrm{F}}_{T^*|X} - \mathrm{F}_{T^*|X}\big)(\mathrm{F}_{T^*}^{-1}(s) \mid X_i) \\
&\quad + \Big(\frac{1}{n}\sum_{i=1}^{n}\omega_{QY}(X_i; r, s)\Big)\,\sqrt{n}\big(\widehat{\mathrm{F}}_{Y^*}^{-1}(r) - \mathrm{F}_{Y^*}^{-1}(r)\big) \\
&\quad + \Big(\frac{1}{n}\sum_{i=1}^{n}\omega_{QT}(X_i; r, s)\Big)\,\sqrt{n}\big(\widehat{\mathrm{F}}_{T^*}^{-1}(s) - \mathrm{F}_{T^*}^{-1}(s)\big) \\
&\quad + \Big(\frac{1}{n}\sum_{i=1}^{n}\omega_{\delta}(X_i; r, s)\Big)'\,\sqrt{n}(\hat{\delta} - \delta) \\
&\quad + \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Big\{C(\mathrm{F}_{Y^*|X}(\mathrm{F}_{Y^*}^{-1}(r) \mid X_i), \mathrm{F}_{T^*|X}(\mathrm{F}_{T^*}^{-1}(s) \mid X_i); \delta) \\
&\qquad\qquad - \mathbb{E}\big[C(\mathrm{F}_{Y^*|X}(\mathrm{F}_{Y^*}^{-1}(r) \mid X_i), \mathrm{F}_{T^*|X}(\mathrm{F}_{T^*}^{-1}(s) \mid X_i); \delta)\big]\Big\}.
\end{aligned}
$$

where

$$
\omega_{QY}(X_i; r, s) := \omega_{FY}(X_i; r, s)\,\widehat{f}_{Y^*|X}(\bar{y}_i \mid X_i), \text{ and } \qquad \omega_{QT}(X_i; r, s) := \omega_{FT}(X_i; r, s)\,\widehat{f}_{T^*|X}(\bar{t}_i \mid X_i).
$$

First, by Assumption 7 and Corollary 1(a),

$$
\sqrt{n}\big(\widehat{\mathrm{F}}_{Y^*|X} - \mathrm{F}_{Y^*|X}\big)(\mathrm{F}_{Y^*}^{-1}(r) \mid X_i) = \mathbb{M}_{FY}^{L}(\mathrm{F}_{Y^*}^{-1}(r), X_i)'\sqrt{n}\begin{bmatrix}\widehat{\boldsymbol{\beta}}_{Y^*}^{L} - \boldsymbol{\beta}_{Y^*}^{L} \\ \widehat{\sigma}_{Y^*} - \sigma_{Y^*}\end{bmatrix} + o_p(1).
$$

Second, by Lemma S1 and Assumption 7,

$$
\begin{aligned}
&\sqrt{n}\big(\widehat{\mathrm{F}}_{Y^*}^{-1}(r) - F_{Y^*}^{-1}(r)\big) \\
&\quad = -\frac{1}{f_{Y^*}(\mathrm{F}_{Y^*}^{-1}(r))}\mathbb{E}\big[\mathbb{M}_{FY}^{L}\big(\mathrm{F}_{Y^*}^{-1}(r), X\big)\big]'\sqrt{n}\begin{bmatrix}\widehat{\boldsymbol{\beta}}_{Y^*}^{L} - \boldsymbol{\beta}_{Y^*}^{L} \\ \widehat{\sigma}_{Y^*} - \sigma_{Y^*}\end{bmatrix} \\
&\qquad - \frac{1}{f_{Y^*}(\mathrm{F}_{Y^*}^{-1}(r))}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big(\mathrm{F}_{Y^*|X}(\mathrm{F}_{Y^*}^{-1}(r) \mid X_i) - \mathbb{E}[\mathrm{F}_{Y^*|X}(\mathrm{F}_{Y^*}^{-1}(r) \mid X_i)]\big) + o_p(1).
\end{aligned}
$$

Third, by Proposition 2,

$$
\sqrt{n}(\widehat{\delta} - \delta) = -\mathcal{H}_{\delta}^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathcal{S}_i(\delta) - \mathcal{H}_{\delta}^{-1}\mathbb{M}_{\Delta Y}^{L}\sqrt{n}\begin{bmatrix}\widehat{\boldsymbol{\beta}}_{Y^*}^{L} - \boldsymbol{\beta}_{Y^*}^{L} \\ \widehat{\sigma}_{Y^*} - \sigma_{Y^*}\end{bmatrix} - \mathcal{H}_{\delta}^{-1}\mathbb{M}_{\Delta T}^{L}\sqrt{n}\begin{bmatrix}\widehat{\boldsymbol{\beta}}_{T^*}^{L} - \boldsymbol{\beta}_{T^*}^{L} \\ \widehat{\sigma}_{T^*} - \sigma_{T^*}\end{bmatrix} + o_p(1)
$$

Lastly, the summands

$$
\Big\{C(\mathrm{F}_{Y^*|X}(\mathrm{F}_{Y^*}^{-1}(r) \mid X_i), \mathrm{F}_{T^*|X}(\mathrm{F}_{T^*}^{-1}(s) \mid X_i); \delta) - \mathbb{E}\big[C(\mathrm{F}_{Y^*|X}(\mathrm{F}_{Y^*}^{-1}(r) \mid X_i), \mathrm{F}_{T^*|X}(\mathrm{F}_{T^*}^{-1}(s) \mid X_i); \delta)\big]\Big\}_{i=1}^{n}
$$

are *i.i.d*, mean-zero, with bounded second moments uniformly in $(r, s, \delta')' \in (0, 1)^2 \times \Gamma_\delta$ by the definition of a copula function.

In sum, $\sqrt{n}\big(\widehat{C}_{Y^*T^*}(r, s; \widehat{\delta}) - C_{Y^*T^*}(r, s; \delta)\big)$ can be expressed as a weighted sum of asymptotically normal quantities up to a $o_p(1)$ term:

$$
\begin{aligned}
&\sqrt{n}\big(\widehat{C}_{Y^*T^*}(r,s;\widehat{\delta}) - C_{Y^*T^*}(r,s;\delta)\big) \\
&\quad = \mathbb{M}^L_{CY}(r,s)'\sqrt{n}\begin{bmatrix}\widehat{\boldsymbol{\beta}}^L_{Y^*} - \boldsymbol{\beta}^L_{Y^*} \\ \widehat{\sigma}_{Y^*} - \sigma_{Y^*}\end{bmatrix} \;+\; \mathbb{M}^L_{CT}(r,s)'\sqrt{n}\begin{bmatrix}\widehat{\boldsymbol{\beta}}^L_{T^*} - \boldsymbol{\beta}^L_{T^*} \\ \widehat{\sigma}_{T^*} - \sigma_{T^*}\end{bmatrix} \\
&\qquad + \mathbb{M}^L_{CF^{-1}_{Y^*}}(r,s)\frac{1}{\sqrt{n}}\sum_{i=1}^n \big(\mathrm{F}_{Y^*|X}(\mathrm{F}^{-1}_{Y^*}(r) \mid X_i) - \mathbb{E}[\mathrm{F}_{Y^*|X}(\mathrm{F}^{-1}_{Y^*}(r) \mid X_i)]\big) \\
&\qquad + \mathbb{M}^L_{CF^{-1}_{T^*}}(r,s)\frac{1}{\sqrt{n}}\sum_{i=1}^n \big(\mathrm{F}_{T^*|X}(\mathrm{F}^{-1}_{T^*}(s) \mid X_i) - \mathbb{E}[\mathrm{F}_{T^*|X}(\mathrm{F}^{-1}_{T^*}(s) \mid X_i)]\big) \\
&\qquad + \frac{1}{\sqrt{n}}\sum_{i=1}^n \Big\{C(\mathrm{F}_{Y^*|X}(\mathrm{F}^{-1}_{Y^*}(r) \mid X_i), \mathrm{F}_{T^*|X}(\mathrm{F}^{-1}_{T^*}(s) \mid X_i);\delta) \\
&\qquad\qquad - \mathbb{E}\big[C(\mathrm{F}_{Y^*|X}(\mathrm{F}^{-1}_{Y^*}(r) \mid X_i), \mathrm{F}_{T^*|X}(\mathrm{F}^{-1}_{T^*}(s) \mid X_i);\delta)\big]\Big\} \\
&\qquad + \mathbb{M}^L_{C\delta}(r,s)'\frac{1}{\sqrt{n}}\sum_{i=1}^n \mathcal{S}_i(\delta) + o_p(1) \\
&\quad =: \frac{1}{\sqrt{n}}\sum_{i=1}^n \psi_i^{(d)}(r,s) + o_p(1)
\end{aligned} \tag{S2}
$$

where the representation in the last line follows because the linear combination of asymptotically linear quantities is asymptotically linear. The conclusion then follows in addition Theorem 3, Proposition 2, and Slutsky's Theorem noting that

$$
\begin{aligned}
&\sqrt{n}\big(\widehat{\theta}_{TM}(r_1,r_2,s_1,s_2) - \theta_{TM}(r_1,r_2,s_1,s_2)\big) = (s_2-s_1)^{-1}\sqrt{n}\big(\widehat{C}_{Y^*T^*}(r_2,s_2;\widehat{\delta}) - C_{Y^*T^*}(r_2,s_2;\delta)\big) \\
&\; -(s_2-s_1)^{-1}\sqrt{n}\big(\widehat{C}_{Y^*T^*}(r_1,s_2;\widehat{\delta}) - C_{Y^*T^*}(r_1,s_2;\delta)\big) - (s_2-s_1)^{-1}\sqrt{n}\big(\widehat{C}_{Y^*T^*}(r_2,s_1;\widehat{\delta}) - C_{Y^*T^*}(r_2,s_1;\delta)\big) \\
&\; +(s_2-s_1)^{-1}\sqrt{n}\big(\widehat{C}_{Y^*T^*}(r_1,s_1;\widehat{\delta}) - C_{Y^*T^*}(r_1,s_1;\delta)\big) \\
&\quad =: (s_2-s_1)^{-1}\Big(\frac{1}{\sqrt{n}}\sum_{i=1}^n \psi_i^{(d)}(r_2,s_2) - \frac{1}{\sqrt{n}}\sum_{i=1}^n \psi_i^{(d)}(r_1,s_2) - \frac{1}{\sqrt{n}}\sum_{i=1}^n \psi_i^{(d)}(r_2,s_1) + \frac{1}{\sqrt{n}}\sum_{i=1}^n \psi_i^{(d)}(r_1,s_1)\Big) + o_p(1) \\
&\quad =: \frac{1}{\sqrt{n}}\sum_{i=1}^n \psi_i^{(d)} + o_p(1)
\end{aligned}
$$

where $\sigma(\theta_{TM}(r_1,r_2,s_1,s_2)) = \mathbb{E}\big[(\psi^{(d)})^2\big]$. □

SA.2.4 Proof of Theorem 4(e)

The following lemma provides a proof of part (e) of Theorem 4 from the main text.

Lemma S5. *Under Assumptions 1 to 10,* $\sqrt{n}(\widehat{\rho}_S - \rho_S) \xrightarrow{d} \mathcal{N}\big(0, \sigma(\rho_S)\big)$.

Proof. Recall $\rho_S = 12\int_0^1\int_0^1 C_{Y^*T^*}(r,s)\,dr\,ds - 3$ where

$$
C_{Y^*T^*}(r,s) = \mathbb{E}\Big[C_{Y^*T^*|X}\big(\mathrm{F}_{Y^*|X}(\mathrm{F}^{-1}_{Y^*}(r) \mid X),\; \mathrm{F}_{T^*|X}(\mathrm{F}^{-1}_{T^*}(s) \mid X);\, \delta\big)\Big].
$$

From (S2) in the proof of Lemma S4,

$$
\begin{aligned}
\sqrt{n}(\widehat{\rho}_S - \rho_S)/12 &= \sqrt{n}\int_0^1\int_0^1 \big(\widehat{C}_{Y^*T^*}(r,s) - C_{Y^*T^*}(r,s)\big)\,dr\,ds \\
&= \Big(\int_0^1\int_0^1 \mathbb{M}^L_{CY}(r,s)\,dr\,ds\Big)'\sqrt{n}\begin{bmatrix}\widehat{\boldsymbol{\beta}}^L_{Y^*} - \boldsymbol{\beta}^L_{Y^*}\\ \widehat{\sigma}_{Y^*} - \sigma_{Y^*}\end{bmatrix} \;+\; \Big(\int_0^1\int_0^1 \mathbb{M}^L_{CT}(r,s)\,dr\,ds\Big)'\sqrt{n}\begin{bmatrix}\widehat{\boldsymbol{\beta}}^L_{T^*} - \boldsymbol{\beta}^L_{T^*}\\ \widehat{\sigma}_{T^*} - \sigma_{T^*}\end{bmatrix} \\
&\quad + \frac{1}{\sqrt{n}}\sum_{i=1}^n \Big\{\Big(\int_0^1\int_0^1 \mathbb{M}^L_{CF^{-1}_{Y^*}}(r,s)\mathrm{F}_{Y^*|X}(\mathrm{F}^{-1}_{Y^*}(r)\mid X_i)\,dr\,ds\Big) \\
&\qquad\qquad - \mathbb{E}\Big[\int_0^1\int_0^1 \mathbb{M}^L_{CF^{-1}_{Y^*}}(r,s)\mathrm{F}_{Y^*|X}(\mathrm{F}^{-1}_{Y^*}(r)\mid X_i)\,dr\,ds\Big]\Big\} \\
&\quad + \frac{1}{\sqrt{n}}\sum_{i=1}^n \Big\{\Big(\int_0^1\int_0^1 \mathbb{M}^L_{CF^{-1}_{T^*}}(r,s)\mathrm{F}_{T^*|X}(\mathrm{F}^{-1}_{T^*}(s)\mid X_i)\,dr\,ds\Big) \\
&\qquad\qquad - \mathbb{E}\Big[\int_0^1\int_0^1 \mathbb{M}^L_{CF^{-1}_{T^*}}(r,s)\mathrm{F}_{T^*|X}(\mathrm{F}^{-1}_{T^*}(s)\mid X_i)\,dr\,ds\Big]\Big\} \\
&\quad + \frac{1}{\sqrt{n}}\sum_{i=1}^n \Big\{\int_0^1\int_0^1 C(\mathrm{F}_{Y^*|X}(\mathrm{F}^{-1}_{Y^*}(r)\mid X_i), \mathrm{F}_{T^*|X}(\mathrm{F}^{-1}_{T^*}(s)\mid X_i);\delta)\,dr\,ds \\
&\qquad\qquad - \mathbb{E}\Big[\int_0^1\int_0^1 C(\mathrm{F}_{Y^*|X}(\mathrm{F}^{-1}_{Y^*}(r)\mid X_i), \mathrm{F}_{T^*|X}(\mathrm{F}^{-1}_{T^*}(s)\mid X_i);\delta)\,dr\,ds\Big]\Big\} \\
&\quad + \Big(\int_0^1\int_0^1 \mathbb{M}^L_{C\delta}(r,s)\,dr\,ds\Big)'\frac{1}{\sqrt{n}}\sum_{i=1}^n \mathcal{S}_i(\delta) + o_p(1) \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^n \Big\{\int_0^1\int_0^1 \psi_i^{(d)}(r,s)\,dr\,ds\Big\} + o_p(1) \\
&=: \frac{1}{\sqrt{n}}\sum_{i=1}^n \psi_i^{(e)} + o_p(1) \xrightarrow{d} \mathcal{N}(0, \sigma(\rho_S))
\end{aligned}
$$

where $\sigma(\rho_S) = \mathbb{E}\big[(\psi_i^{(e)})^2\big]$. The conclusion follows from Lemma S4. □

SA.2.5 Proof of Theorem 4(f)

The following lemma provides a proof of part (f) of Theorem 4 from the main text.

Lemma S6. *Suppose Assumptions 1 to 10 hold, then* $\sqrt{n}\big(\widehat{\theta}_U(\Delta, s_1, s_2) - \theta_U(\Delta, s_1, s_2)\big) \xrightarrow{d} \mathcal{N}\big(0,\, \sigma(\theta_U(\Delta, s_1, s_2))\big)$.

Proof. Recall $\theta_U(\Delta, s_1, s_2) := \dfrac{\mathrm{P}\big(\mathrm{F}_{Y^*}(Y^*) > \mathrm{F}_{T^*}(T^*) + \Delta,\ s_1 \le \mathrm{F}_{T^*}(T^*) \le s_2\big)}{s_2 - s_1}$. A first step is to obtain the joint unconditional pdf $f_{Y^*T^*}(\cdot,\,\cdot)$ from the conditional copula density:

$$
\begin{aligned}
f_{Y^*T^*}(y,\, t) &= \int_{\mathcal{X}} f_{Y^*T^*|X}(y,\, t \mid x)\,d\mathrm{F}_X(x) \\
&= \int_{\mathcal{X}} c_{Y^*T^*|X}\big(\mathrm{F}_{Y^*|X}(y\mid x), \mathrm{F}_{T^*|X}(t\mid x);\, \delta\big) f_{Y^*|X}(y\mid x) f_{T^*|X}(t\mid x)\,d\mathrm{F}_X(x).
\end{aligned}
$$

Applying a change of variables $r = \mathrm{F}_{Y^*}(y)$ and $s = \mathrm{F}_{T^*}(t)$, then $y = \mathrm{F}^{-1}_{Y^*}(r)$, $t = \mathrm{F}^{-1}_{T^*}(s)$, $dy = dr/f_{Y^*}(y) = dr/f_{Y^*}(\mathrm{F}^{-1}_{Y^*}(r))$, and $dt = ds/f_{T^*}(t) = ds/f_{T^*}(\mathrm{F}^{-1}_{T^*}(s))$. Thus

$$
\begin{aligned}
\mathrm{P}\big(\mathrm{F}_{Y^*}(Y^*) > \mathrm{F}_{T^*}(T^*) + \Delta,\ s_1 \le \mathrm{F}_{T^*}(T^*) \le s_2\big)&\\
&= \iint \mathbb{1}\{\mathrm{F}_{Y^*}(y) > \mathrm{F}_{T^*}(t) + \Delta\}\mathbb{1}\{s_1 \le \mathrm{F}_{T^*}(t) \le s_2\} f_{Y^*T^*}(y,t)\,dy\,dt\\
&= \int_0^1\!\!\int_0^1 \mathbb{1}\{r > s+\Delta\}\,\mathbb{1}\{s_1 \le s \le s_2\}\,\frac{f_{Y^*T^*}\big(\mathrm{F}_{Y^*}^{-1}(r),\,\mathrm{F}_{T^*}^{-1}(s)\big)}{f_{Y^*}\big(\mathrm{F}_{Y^*}^{-1}(r)\big)\,f_{T^*}\big(\mathrm{F}_{T^*}^{-1}(s)\big)}\,dr\,ds.
\end{aligned}
$$

The plug-in estimator via numerical integration is given by

$$
\widehat{\theta}_U(\Delta, s_1, s_2) := \int_0^1\!\!\int_0^1 \frac{\mathbb{1}\{r > s+\Delta\}\,\mathbb{1}\{s_1 \le s \le s_2\}}{(s_2 - s_1)}\,\frac{\widehat{f}_{Y^*T^*}\big(\widehat{\mathrm{F}}_{Y^*}^{-1}(r),\,\widehat{\mathrm{F}}_{T^*}^{-1}(s)\big)}{\widehat{f}_{Y^*}\big(\widehat{\mathrm{F}}_{Y^*}^{-1}(r)\big)\,\widehat{f}_{T^*}\big(\widehat{\mathrm{F}}_{T^*}^{-1}(s)\big)}\,dr\,ds.
$$

Consider the decomposition:

$$
\begin{aligned}
&\frac{\widehat{f}_{Y^*T^*}\big(\widehat{\mathrm{F}}_{Y^*}^{-1}(r),\,\widehat{\mathrm{F}}_{T^*}^{-1}(s)\big)}{\widehat{f}_{Y^*}\big(\widehat{\mathrm{F}}_{Y^*}^{-1}(r)\big)\,\widehat{f}_{T^*}\big(\widehat{\mathrm{F}}_{T^*}^{-1}(s)\big)} - \frac{f_{Y^*T^*}\big(\mathrm{F}_{Y^*}^{-1}(r),\,\mathrm{F}_{T^*}^{-1}(s)\big)}{f_{Y^*}\big(\mathrm{F}_{Y^*}^{-1}(r)\big)\,f_{T^*}\big(\mathrm{F}_{T^*}^{-1}(s)\big)} = \frac{\widehat{f}_{Y^*T^*}(y,t) - f_{Y^*T^*}(y,t)}{f_{Y^*}(y)f_{T^*}(t)}\\
&\quad + \frac{\partial_y \widehat{f}_{Y^*T^*}(\bar{y}_1,\bar{t}_1)\,(\widehat{y}-y) + \partial_t \widehat{f}_{Y^*T^*}(\bar{y}_2,\bar{t}_2)\,(\widehat{t}-t)}{f_{Y^*}(y)f_{T^*}(t)}\\
&\quad + \frac{\widehat{f}_{Y^*T^*}(\widehat{y},\widehat{t})}{f_{Y^*}(y)f_{T^*}(t)\,\widehat{f}_{Y^*}(\widehat{y})\widehat{f}_{T^*}(\widehat{t})}\\
&\qquad \times \Big\{\big[f_{Y^*}(y) - \widehat{f}_{Y^*}(y)\big]\,f_{T^*}(t) + f_{Y^*}(y)\,\big[f_{T^*}(t) - \widehat{f}_{T^*}(t)\big] - \widehat{f}_{T^*}(\widehat{t})\,\widehat{f}'_{Y^*}(\bar{y})\,(\widehat{y}-y) - \widehat{f}_{Y^*}(y)\,\widehat{f}'_{T^*}(\bar{t})\,(\widehat{t}-t)\Big\}\\
&= \omega_f(y,t)\,\big(\widehat{f}_{Y^*T^*}(y,t) - f_{Y^*T^*}(y,t)\big) \;+\; \widehat{\omega}_{fy}(\widehat{y},\widehat{t},y)\,\big(\widehat{f}_{Y^*}(y) - f_{Y^*}(y)\big) \;+\; \widehat{\omega}_{ft}(\widehat{y},\widehat{t},t)\,\big(\widehat{f}_{T^*}(t) - f_{T^*}(t)\big)\\
&\qquad + \;\widehat{\omega}_y(\widehat{y},\widehat{t},y,t)\,(\widehat{y}-y) \;+\; \widehat{\omega}_t(\widehat{y},\widehat{t},y,t)\,(\widehat{t}-t)
\end{aligned}
$$

with weights given by $\omega_f(y,t) = \frac{1}{f_{Y^*}(y)f_{T^*}(t)}$, $\widehat{\omega}_{fy}(\widehat{y},\widehat{t},y) = -\dfrac{\widehat{f}_{Y^*T^*}(\widehat{y},\widehat{t})}{f_{Y^*}(y)\,\widehat{f}_{Y^*}(\widehat{y})\,\widehat{f}_{T^*}(\widehat{t})}$, $\widehat{\omega}_{ft}(\widehat{y},\widehat{t},t) = -\dfrac{\widehat{f}_{Y^*T^*}(\widehat{y},\widehat{t})}{f_{T^*}(t)\,\widehat{f}_{Y^*}(\widehat{y})\,\widehat{f}_{T^*}(\widehat{t})}$, $\widehat{\omega}_y(\widehat{y},\widehat{t},y,t) = \dfrac{\partial_y \widehat{f}_{Y^*T^*}(\bar{y}_1,\bar{t}_1)}{f_{Y^*}(y)f_{T^*}(t)} - \dfrac{\widehat{f}_{Y^*T^*}(\widehat{y},\widehat{t})\,\widehat{f}_{T^*}(\widehat{t})}{f_{Y^*}(y)f_{T^*}(t)\,\widehat{f}_{Y^*}(\widehat{y})\,\widehat{f}_{T^*}(\widehat{t})}\,\widehat{f}'_{Y^*}(\bar{y})$, and $\widehat{\omega}_t(\widehat{y},\widehat{t},y,t) = \dfrac{\partial_t \widehat{f}_{Y^*T^*}(\bar{y}_2,\bar{t}_2)}{f_{Y^*}(y)f_{T^*}(t)} - \dfrac{\widehat{f}_{Y^*T^*}(\widehat{y},\widehat{t})\,\widehat{f}_{Y^*}(y)}{f_{Y^*}(y)f_{T^*}(t)\,\widehat{f}_{Y^*}(\widehat{y})\,\widehat{f}_{T^*}(\widehat{t})}\,\widehat{f}'_{T^*}(\bar{t})$.

Thus,

$$
\begin{aligned}
&\sqrt{n}\big(\widehat{\theta}_U(\Delta,s_1,s_2) - \theta_U(\Delta,s_1,s_2)\big)\\
&\quad = \int_0^1\!\!\int_0^1 \frac{\mathbb{1}\{r > s+\Delta\}\,\mathbb{1}\{s_1 \le s \le s_2\}}{(s_2 - s_1)}\left(\frac{\widehat{f}_{Y^*T^*}\big(\widehat{\mathrm{F}}_{Y^*}^{-1}(r),\,\widehat{\mathrm{F}}_{T^*}^{-1}(s)\big)}{\widehat{f}_{Y^*}\big(\widehat{\mathrm{F}}_{Y^*}^{-1}(r)\big)\,\widehat{f}_{T^*}\big(\widehat{\mathrm{F}}_{T^*}^{-1}(s)\big)} - \frac{f_{Y^*T^*}\big(\mathrm{F}_{Y^*}^{-1}(r),\,\mathrm{F}_{T^*}^{-1}(s)\big)}{f_{Y^*}\big(\mathrm{F}_{Y^*}^{-1}(r)\big)\,f_{T^*}\big(\mathrm{F}_{T^*}^{-1}(s)\big)}\right)dr\,ds\\
&\quad = \int_0^1\!\!\int_0^1 \frac{\mathbb{1}\{r > s+\Delta\}\,\mathbb{1}\{s_1 \le s \le s_2\}}{(s_2 - s_1)}\,\omega_f(y,t)\,\sqrt{n}\big(\widehat{f}_{Y^*T^*}(\mathrm{F}_{Y^*}^{-1}(r), F_{T^*}^{-1}(s)) - f_{Y^*T^*}(\mathrm{F}_{Y^*}^{-1}(r), F_{T^*}^{-1}(s))\big)\,dr\,ds\\
&\quad + \int_0^1\left(\int_0^1 \frac{\mathbb{1}\{r > s+\Delta\}\,\mathbb{1}\{s_1 \le s \le s_2\}}{(s_2 - s_1)}\,\widehat{\omega}_{fy}(\widehat{y},\widehat{t},y)\,ds\right)\sqrt{n}\big(\widehat{f}_{Y^*}(\mathrm{F}_{Y^*}^{-1}(r)) - f_{Y^*}(\mathrm{F}_{Y^*}^{-1}(r))\big)\,dr\\
&\quad + \int_0^1\left(\int_0^1 \frac{\mathbb{1}\{r > s+\Delta\}\,\mathbb{1}\{s_1 \le s \le s_2\}}{(s_2 - s_1)}\,\widehat{\omega}_{ft}(\widehat{y},\widehat{t},t)\,dr\right)\sqrt{n}\big(\widehat{f}_{T^*}(\mathrm{F}_{T^*}^{-1}(s)) - f_{T^*}(\mathrm{F}_{T^*}^{-1}(s))\big)\,ds\\
&\quad + \int_0^1\left(\int_0^1 \frac{\mathbb{1}\{r > s+\Delta\}\,\mathbb{1}\{s_1 \le s \le s_2\}}{(s_2 - s_1)}\,\widehat{\omega}_y(\widehat{y},\widehat{t},y,t)\;ds\right)\sqrt{n}\big(\widehat{\mathrm{F}}_{Y^*}^{-1}(r) - \mathrm{F}_{Y^*}^{-1}(r)\big)\,dr
\end{aligned}
$$

$$+\int_0^1 \left(\int_0^1 \frac{\mathbb{1}\{r > s + \Delta\}\, \mathbb{1}\{s_1 \leq s \leq s_2\}}{(s_2 - s_1)}\, \widehat{\omega}_t(\widehat{y}, \widehat{t}, y, t)\, dr \right) \sqrt{n}\big(\widehat{\mathrm{F}}_{T^*}^{-1}(s) - \mathrm{F}_{T^*}^{-1}(s)\big)\, ds.$$

The summands are studied in turn.

First, by the MVT,

$$\begin{aligned}
&\sqrt{n}\big(\widehat{f}_{Y^*T^*}(y,t) - f_{Y^*T^*}(y,t)\big) \\
&= \frac{1}{n}\sum_{i=1}^{n} \Big[c_{Y^*T^*|X}\big(\widehat{\mathrm{F}}_{Y^*|X}(y \mid X_i), \widehat{\mathrm{F}}_{T^*|X}(t \mid X_i);\, \widehat{\delta}\big)\, \widehat{f}_{Y^*|X}(y \mid X_i) \widehat{f}_{T^*|X}(t \mid X_i) \Big] \\
&\quad - \mathbb{E}\Big[c_{Y^*T^*|X}\big(\mathrm{F}_{Y^*|X}(y \mid X), \mathrm{F}_{T^*|X}(t \mid X);\, \delta\big)\, f_{Y^*|X}(y \mid X) f_{T^*|X}(t \mid X) \Big] \\
&= \frac{1}{n}\sum_{i=1}^{n} \Big\{ c_{Y^*T^*|X}\big(\widehat{\mathrm{F}}_{Y^*|X}(y \mid X_i), \widehat{\mathrm{F}}_{T^*|X}(t \mid X_i);\, \widehat{\delta}\big)\, \widehat{f}_{Y^*|X}(y \mid X_i) \widehat{f}_{T^*|X}(t \mid X_i) \\
&\qquad - c_{Y^*T^*|X}\big(\mathrm{F}_{Y^*|X}(y \mid X_i), \mathrm{F}_{T^*|X}(t \mid X_i);\, \delta\big)\, f_{Y^*|X}(y \mid X_i) f_{T^*|X}(t \mid X_i) \Big\} \\
&\quad + \frac{1}{n}\sum_{i=1}^{n} \Big\{ c_{Y^*T^*|X}\big(\mathrm{F}_{Y^*|X}(y \mid X_i), \mathrm{F}_{T^*|X}(t \mid X_i);\, \delta\big)\, f_{Y^*|X}(y \mid X_i) f_{T^*|X}(t \mid X_i) \\
&\qquad - \mathbb{E}\Big[c_{Y^*T^*|X}\big(\mathrm{F}_{Y^*|X}(y \mid X_i), \mathrm{F}_{T^*|X}(t \mid X_i);\, \delta\big)\, f_{Y^*|X}(y \mid X_i) f_{T^*|X}(t \mid X_i) \Big] \Big\}
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{n}\sum_{i=1}^{n} \omega_{FY}(X_i; y, t) \sqrt{n}\big(\widehat{\mathrm{F}}_{Y^*|X}(y \mid X_i) - \mathrm{F}_{Y^*|X}(y \mid X_i)\big) \\
&\quad + \frac{1}{n}\sum_{i=1}^{n} \omega_{FT}(X_i; y, t) \sqrt{n}\big(\widehat{\mathrm{F}}_{T^*|X}(t \mid X_i) - \mathrm{F}_{T^*|X}(t \mid X_i)\big) \\
&\quad + \frac{1}{n}\sum_{i=1}^{n} \omega_{fY}(X_i; y, t) \sqrt{n}\big(\widehat{f}_{Y^*|X}(y \mid X_i) - f_{Y^*|X}(y \mid X_i)\big) \\
&\quad + \frac{1}{n}\sum_{i=1}^{n} \omega_{fT}(X_i; y, t) \sqrt{n}\big(\widehat{f}_{T^*|X}(t \mid X_i) - f_{T^*|X}(t \mid X_i)\big) \\
&\quad + \frac{1}{n}\sum_{i=1}^{n} \omega_{\delta}(X_i; y, t)' \sqrt{n}\big(\widehat{\delta} - \delta\big) \\
&\quad + \frac{1}{\sqrt{n}}\sum_{i=1}^{n} \Big\{ c_{Y^*T^*|X}\big(\mathrm{F}_{Y^*|X}(y \mid X_i), \mathrm{F}_{T^*|X}(t \mid X_i);\, \delta\big)\, f_{Y^*|X}(y \mid X_i) f_{T^*|X}(t \mid X_i) \\
&\qquad - \mathbb{E}\Big[c_{Y^*T^*|X}\big(\mathrm{F}_{Y^*|X}(y \mid X_i), \mathrm{F}_{T^*|X}(t \mid X_i);\, \delta\big)\, f_{Y^*|X}(y \mid X_i) f_{T^*|X}(t \mid X_i) \Big] \Big\}
\end{aligned}$$

where the weights are given by

$$\begin{aligned}
\omega_{FY}(X_i; y, t) &= c_1(\bar{u}_i, \widehat{v}_i; \widehat{\delta})\, \widehat{f}_{Y^*|X}(y \mid X_i) \widehat{f}_{T^*|X}(t \mid X_i), \quad \omega_{FT}(X_i; y, t) = c_2(u_i, \bar{v}_i; \widehat{\delta})\, \widehat{f}_{Y^*|X}(y \mid X_i) \widehat{f}_{T^*|X}(t \mid X_i), \\
\omega_{fY}(X_i; y, t) &= c(u_i, v_i; \widehat{\delta})\, \widehat{f}_{T^*|X}(t \mid X_i), \quad \omega_{fT}(X_i; y, t) = c(u_i, v_i; \widehat{\delta})\, f_{Y^*|X}(y \mid X_i), \\
\omega_{\delta}(X_i; y, t) &= c_{\delta}(u_i, v_i; \bar{\delta}_i)\, f_{Y^*|X}(y \mid X_i) f_{T^*|X}(t \mid X_i),\ u_i := \mathrm{F}_{Y^*|X}(y \mid X_i),\ \hat{u}_i := \widehat{\mathrm{F}}_{Y^*|X}(y \mid X_i), \\
v_i &:= \mathrm{F}_{T^*|X}(t \mid X_i), \quad \text{and } \hat{v}_i := \widehat{\mathrm{F}}_{T^*|X}(t \mid X_i).
\end{aligned}$$

The summands of

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Big\{c_{Y^*T^*|X}\big(\mathrm{F}_{Y^*|X}(y\mid X_i),\mathrm{F}_{T^*|X}(t\mid X_i);\,\delta\big)f_{Y^*|X}(y\mid X_i)f_{T^*|X}(t\mid X_i)$$
$$-\,\mathbb{E}\Big[c_{Y^*T^*|X}\big(\mathrm{F}_{Y^*|X}(y\mid X_i),\mathrm{F}_{T^*|X}(t\mid X_i);\,\delta\big)f_{Y^*|X}(y\mid X_i)f_{T^*|X}(t\mid X_i)\Big]\Big\}$$

are $i.i.d.$ under Assumption 4, mean-zero, with bounded second moments under Assumptions 6 and 10 — the CLT applies.

Second, under the conditions of Lemma 3(a),

$$\begin{aligned}
\sqrt{n}\big(\widehat{f}_{Y^*}(\mathrm{F}^{-1}_{Y^*}(r))-f_{Y^*}(\mathrm{F}^{-1}_{Y^*}(r))\big) &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big(\widehat{f}_{Y^*|X}(\mathrm{F}^{-1}_{Y^*}(r)\mid X_i)-\mathbb{E}\big[f_{Y^*|X}(\mathrm{F}^{-1}_{Y^*}(r)\mid X_i)\big]\big)\\
&= \frac{1}{n}\sum_{i=1}^{n}\sqrt{n}\big(\widehat{f}_{Y^*|X}(\mathrm{F}^{-1}_{Y^*}(r)\mid X_i)-f_{Y^*|X}(\mathrm{F}^{-1}_{Y^*}(r)\mid X_i)\big)\\
&\quad+\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big(f_{Y^*|X}(\mathrm{F}^{-1}_{Y^*}(r)\mid X_i)-\mathbb{E}\big[f_{Y^*|X}(\mathrm{F}^{-1}_{Y^*}(r)\mid X_i)\big]\big)\\
&= \sum_{\ell=1}^{L}\Big\{\frac{1}{n}\sum_{i=1}^{n}\widehat{\mathbb{R}}_{\ell,f_{Y^*|X}}(\mathrm{F}^{-1}_{Y^*}(r),X_i)\Big\}'\sqrt{n}\big(\widehat{\beta}_{Y^*}(\tau_\ell)-\beta_{Y^*}(\tau_\ell)\big)\\
&\quad+\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big(f_{Y^*|X}(\mathrm{F}^{-1}_{Y^*}(r)\mid X_i)-\mathbb{E}\big[f_{Y^*|X}(\mathrm{F}^{-1}_{Y^*}(r)\mid X_i)\big]\big)\\
&= \mathbb{E}\big[\mathbb{M}^{L}_{fY}\big(\mathrm{F}^{-1}_{Y^*}(r),X\big)\big]'\sqrt{n}\begin{bmatrix}\widehat{\boldsymbol{\beta}}^{L}_{Y^*}-\boldsymbol{\beta}^{L}_{Y^*}\\ \widehat{\sigma}_{Y^*}-\sigma_{Y^*}\end{bmatrix}\\
&\quad+\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big(f_{Y^*|X}(\mathrm{F}^{-1}_{Y^*}(r)\mid X_i)-\mathbb{E}\big[f_{Y^*|X}(\mathrm{F}^{-1}_{Y^*}(r)\mid X_i)\big]\big)+o_p(1)
\end{aligned}$$

where

$$\mathbb{M}^{L}_{fY}(y,x):=\operatorname*{plim}_{n\to\infty}\Big[\widehat{\mathbb{R}}_{1,f_{Y^*|X}}(y,x)',\ldots,\widehat{\mathbb{R}}_{L,f_{Y^*|X}}(y,x)',\,0\Big]'.$$

Under Assumption 4 and Assumption 10(d), the summands $\big\{f_{Y^*|X}(\mathrm{F}^{-1}_{Y^*}(r)\mid X_i)-\mathbb{E}\big[f_{Y^*|X}(\mathrm{F}^{-1}_{Y^*}(r)\mid X_i)\big]\big\}_{i=1}^{n}$ are $i.i.d.$, mean-zero with bounded second moments.

Putting together the foregoing, it follows from Corollary 1, lemmas S1 and 3, and proposition 2 that $\sqrt{n}\big(\widehat{\theta}_U(\Delta,s_1,s_2)-\theta_U(\Delta,s_1,s_2)\big)$ has the representation of a weighted sum of asymptotically normally linear and asymptotically normally distributed quantities:

$$\begin{aligned}
&\sqrt{n}\big(\widehat{\theta}_U(\Delta,s_1,s_2)-\theta_U(\Delta,s_1,s_2)\big)\\
&\quad=\Big(\int_0^1\!\!\int_0^1\mathbb{M}^{L}_{\theta Y}(r,s)\,dr\,ds\Big)'\sqrt{n}\begin{bmatrix}\widehat{\boldsymbol{\beta}}^{L}_{Y^*}-\boldsymbol{\beta}^{L}_{Y^*}\\ \widehat{\sigma}_{Y^*}-\sigma_{Y^*}\end{bmatrix}\;+\;\Big(\int_0^1\!\!\int_0^1\mathbb{M}^{L}_{\theta T}(r,s)\,dr\,ds\Big)'\sqrt{n}\begin{bmatrix}\widehat{\boldsymbol{\beta}}^{L}_{T^*}-\boldsymbol{\beta}^{L}_{T^*}\\ \widehat{\sigma}_{T^*}-\sigma_{T^*}\end{bmatrix}\\
&\quad+\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Big\{\int_0^1\!\!\int_0^1\mathbb{M}^{L}_{\theta f}(r,s)c_{Y^*T^*|X}\big(\mathrm{F}_{Y^*|X}(\mathrm{F}^{-1}_{Y^*}(r)\mid X_i),\mathrm{F}_{T^*|X}(\mathrm{F}^{-1}_{T^*}(s)\mid X_i);\,\delta\big)
\end{aligned}$$

$$
\begin{aligned}
&\quad \times f_{Y^*|X}(\mathrm{F}_{Y^*}^{-1}(r) \mid X_i) f_{T^*|X}(\mathrm{F}_{T^*}^{-1}(s) \mid X_i)\, dr\, ds \\
&\quad - \mathbb{E}\Big[\int_0^1 \int_0^1 \mathbb{M}_{\theta f}^L(r,s) c_{Y^*T^*|X}\big(\mathrm{F}_{Y^*|X}(\mathrm{F}_{Y^*}^{-1}(r) \mid X_i), \mathrm{F}_{T^*|X}(\mathrm{F}_{T^*}^{-1}(s) \mid X_i);\, \delta\big) \\
&\quad \times f_{Y^*|X}(\mathrm{F}_{Y^*}^{-1}(r) \mid X_i) f_{T^*|X}(\mathrm{F}_{T^*}^{-1}(s) \mid X_i)\, dr\, ds\Big]\Big\} \\
&+ \frac{1}{\sqrt{n}} \sum_{i=1}^n \Big\{ \int_0^1 \int_0^1 \mathbb{M}_{\theta f_{Y^*|X}}^L(r,s) f_{Y^*|X}(\mathrm{F}_{Y^*}^{-1}(r) \mid X_i)\, dr\, ds \\
&\quad - \mathbb{E}\Big[\int_0^1 \int_0^1 \mathbb{M}_{\theta f_{Y^*|X}}^L(r,s) f_{Y^*|X}(\mathrm{F}_{Y^*}^{-1}(r) \mid X_i)\, dr\, ds\Big]\Big\} \\
&+ \frac{1}{\sqrt{n}} \sum_{i=1}^n \Big\{ \int_0^1 \int_0^1 \mathbb{M}_{\theta f_{T^*|X}}^L(r,s) f_{T^*|X}(\mathrm{F}_{T^*}^{-1}(s) \mid X_i)\, dr\, ds \\
&\quad - \mathbb{E}\Big[\int_0^1 \int_0^1 \mathbb{M}_{\theta f_{T^*|X}}^L(r,s) f_{T^*|X}(\mathrm{F}_{T^*}^{-1}(s) \mid X_i)\, dr\, ds\Big]\Big\} \\
&+ \frac{1}{\sqrt{n}} \sum_{i=1}^n \Big\{ \int_0^1 \int_0^1 \mathbb{M}_{\theta F_{Y^*}^{-1}}^L(r,s) \mathrm{F}_{Y^*|X}(\mathrm{F}_{Y^*}^{-1}(r) \mid X_i)\, dr\, ds - \mathbb{E}\Big[\int_0^1 \int_0^1 \mathbb{M}_{\theta F_{Y^*}^{-1}}^L(r,s) \mathrm{F}_{Y^*|X}(\mathrm{F}_{Y^*}^{-1}(r) \mid X_i)\, dr\, ds\Big]\Big\} \\
&+ \frac{1}{\sqrt{n}} \sum_{i=1}^n \Big\{ \int_0^1 \int_0^1 \mathbb{M}_{\theta F_{T^*}^{-1}}^L(r,s) \mathrm{F}_{T^*|X}(\mathrm{F}_{T^*}^{-1}(s) \mid X_i)\, dr\, ds - \mathbb{E}\Big[\int_0^1 \int_0^1 \mathbb{M}_{\theta F_{T^*}^{-1}}^L(r,s) \mathrm{F}_{T^*|X}(\mathrm{F}_{T^*}^{-1}(s) \mid X_i)\, dr\, ds\Big]\Big\} \\
&+ \Big(\int_0^1 \int_0^1 \mathbb{M}_{\theta\delta}^L(r,s)\, dr\, ds\Big)' \frac{1}{\sqrt{n}} \sum_{i=1}^n \mathcal{S}_i(\delta) + o_p(1) \\
&=: \frac{1}{\sqrt{n}} \sum_{i=1}^n \psi_i^{(f)} + o_p(1) \xrightarrow{d} \mathcal{N}\big(0,\, \sigma(\theta_U(\Delta, s_1, s_2))\big)
\end{aligned}
$$

where $\sigma(\theta_U(\Delta, s_1, s_2)) = \mathbb{E}\big[(\psi^{(f)})^2\big]$. □

SB Causal Effects

Our framework is also related to the literature on continuous treatment effects (see, for example, Hirano and Imbens (2004), Flores (2007), Flores, Flores-Lagunes, Gonzalez, and Neumann (2012), and Galvao and Wang (2015)).[1] Like that literature, we consider the case with one particular continuous "treatment" variable whose effect on the outcome is of primary interest, but with other covariates that need to be controlled for. Let $Y^*(t)$ denote an individual's "potential" outcome if they experience treatment t; note that this is well-defined regardless of what treatment level a particular individual actually experiences, but (in the absence of measurement error) only $Y^* = Y^*(T^*)$ is observed. Next, we show that several of our parameters of interest in Section 2 correspond to the parameters of interest in the continuous treatment effect literature under the commonly invoked assumption of unconfoundedness.

Assumption U (Unconfoundedness)**.** *For all $t \in$ support(T^*),*

$$Y^*(t) \perp\!\!\!\perp T^* | X$$

Assumption U says that, after conditioning on covariates X, the amount/intensity of the treatment is as

[1] See Haavelmo (1944) (in particular, the discussion about "potential influence" in Section 7) for a much earlier discussion of causality with continuous treatments that is conceptually similar to the references above.

good as randomly assigned. This is a strong assumption; it seems unlikely that it would hold in the context of intergenerational mobility — to be clear, we do not maintain this assumption in our application but note it here to show that our approach could deliver causally interpretable parameters if Assumption U holds in a given application. That being said, Assumption U is the leading assumption in the continuous treatment effects literature (as well as being a leading assumption in the binary treatment effects literature), and a similar assumption has been made in each of the continuous treatment effect papers cited above. Under Assumption U, it is straightforward to show that

$$\mathrm{P}\big(Y^*(t) \leq y | X = x\big) = \mathrm{F}_{Y^*|T^*X}(y|t, x)$$

which implies that the distribution of potential outcomes is the same as the distribution in Equation (1) in the main text. Thus, Assumption U serves to give terms like those in Equation (1) a causal interpretation as in the next example.

Example S1 (Dose-Response Functions and Distributional Treatment Effects)**.**

The treatment effects mentioned above depend on particular values of the covariates, x. Often, researchers would like to report a summary measure that integrates out the covariates. For a continuous treatment, researchers often report these dose-response functions.[2] In particular, under Assumption U, the distributional dose-response function is given by

$$\mathrm{P}\big(Y^*(t) \leq y\big) = \mathbb{E}\big[\mathrm{F}_{Y^*|T^*X}(y|t, X)\big] \tag{S3}$$

Interestingly, this is exactly the same transformation of the conditional distribution as for the counterfactual distribution in Equation (3) in the main text. Moreover, under Assumption U, an unconditional distributional treatment effect of moving from treatment level t to treatment level t' is given by

$$DTE(t, t') = \mathrm{P}\big(Y^*(t) \leq y\big) - \mathrm{P}\big(Y^*(t') \leq y\big)$$

One can analogously define unconditional quantile treatment effects by the change in particular quantiles when moving from one treatment level to another. Under Assumption U, each of these treatment effect parameters has a causal interpretation.

SC Life-Cycle Measurement Error

In this appendix, we show how to adapt the life-cycle measurement error framework of Jenkins (1987), Haider and Solon (2006), and Nybom and Stuhler (2016) (among others) to our setting by building on work in Hu and Sasaki (2015) and An, Wang, and Xiao (2020).

We define $Y_{a_Y,i}$ to be child's income if they were observed at age a_Y; likewise, we define $T_{a_T,i}$ to be parents' income if they were observed at age a_T. Life-cycle measurement error models imply that we can

[2]Dose-response functions are closely related to unconditional quantile treatment effects that are commonly reported in the case of a binary treatment. For example, in the case of intergenerational income mobility, having an income at a high conditional quantile (e.g., conditional on parents' education) indicates that a child has a relatively high income conditional on their parents' income and their parents' education. If child's income tends to be increasing in parents' education, then it could be the case that a child of highly educated parents might be in a low conditional quantile but a middle or upper unconditional quantile. Unconditional quantiles correspond to being in the lower or upper part of the overall income distribution (Frolich and Melly (2013) and Powell (2016) contain good discussions of the difference between conditional and unconditional quantiles).

write

$$Y_{a_Y,i} = \lambda^Y_{a_Y} Y^*_i + U^Y_{a_Y,i}$$
$$T_{a_Y,i} = \lambda^T_{a_T} T^*_i + U^Y_{a_T,i}$$

where, for conciseness, the notation for the measurement error terms is slightly modified relative to what it was in the main text. Importantly, this allows for systematically different observed incomes depending on the age at which an individual is observed. For example, it is often argued that $\lambda^Y_{a_Y}$ and $\lambda^Y_{a_T}$ are less than one for those observed during their early adult years.

Let A_Y denote the age at which child's income is actually observed, and let A_T denote the age at which parents' income is actually observed. Thus, the observed $Y_i = Y_{A_Y,i}$ and $T_i = T_{A_T,i}$. Moreover, we can write

$$Y_{A_Y,i} = \lambda^Y_{A_Y} Y^*_i + U^Y_{A_Y,i}$$
$$Y_{A_T,i} = \lambda^T_{A_T} T^*_i + U^T_{A_T,i}$$

We use $\mathcal{A}$ to denote the set of all possible ages (for simplicity, we'll assume that this is the same for parents and children), and we denote the vector of all measurement errors (i.e., across all possible ages) by $U^Y_{\vec{a}_Y}$, for children, and by $U^T_{\vec{a}_T}$ for parents. We make the following assumptions:

Assumption S1. $(Y^*, T^*, U^Y_{\vec{a}_Y}, U^T_{\vec{a}_T}, X) \perp\!\!\!\perp (A_Y, A_T)$

Assumption S2. *For all* $(a_Y, a_T) \in \mathcal{A}^2$, $(U^Y_{a_Y}, U^T_{a_T}) \perp\!\!\!\perp (Y^*, T^*, X)$

Assumption S3. *For all* $(a_Y, a_T) \in \mathcal{A}^2$, $U^Y_{a_Y} \perp\!\!\!\perp U^T_{a_T}$ *and* $(U^Y_{a_Y}/\lambda^Y_{a_Y}) \sim \mathrm{F}_{\tilde{U}_Y}$

Assumption S4. *There exist a known* $a^*_Y \in \mathcal{A}$ *and a known* $a^*_T \in \mathcal{A}$ *such that* $\lambda^Y_{a^*_Y} = \lambda^T_{a^*_T} = 1$.

Assumption S5. *For all* $(a_Y, a_T) \in \mathcal{A}^2$, $\mathbb{E}[U^Y_{a_Y}] = \mathbb{E}[U^T_{a_T}] = 0$

Assumption S1 says that child's permanent income, parents' permanent income, covariates, and the vector of measurement errors across ages are independent of the age at which income is observed. It coincides with the idea that, in terms of permanent incomes, measurement error, and covariates, differences between cohorts at different ages are due to their being observed at different points in the life-cycle. Assumption S2 says that child's permanent income, parents' permanent income, and covariates are independent of the measurement errors at any age. This is the analog of Assumption 1(ii) in the main text. Assumption S2 allows for the distributions of $U^Y_{a_Y}$ and $U^T_{a_T}$ to vary across a_Y and a_T, so that, for example, the variance of the measurement error for parents' income could change at different ages. It also allows for serial correlation in the measurement error so that shocks to income can have persistent effects. Assumption S3 says that the measurement error for child's income is independent of the measurement error for parents' income. This is analogous to Assumption 1(iii) in the main text. The second part says that, after dividing the measurement error for child's income at age a_Y by $\lambda^Y_{a_Y}$, that term follows the same distribution over time. In other words, after accounting for the life-cycle measurement error from $\lambda^Y_{a_Y}$, the distribution of measurement error is the same across ages for children. Alternatively, this assumption could be imposed on the measurement error for parents' income. Assumption S4 says that there exist known ages where the mean of Y_{a_Y} is equal to the mean of Y^* and the mean of Y_{a_T} is equal to the mean of T^* — and, therefore, in these periods the classical measurement error conditions hold. This condition is a required normalization in the life-cycle measurement error literature (see, for example, Assumption 1 in An, Wang, and Xiao (2020)), and it is typically thought to be satisfied for individuals in their early to mid-thirties (Haider and Solon (2006) and Nybom and Stuhler (2016)); and, in some cases, it appears to be satisfied for older ages as well. Assumption S5 says that, for all ages, the mean of the measurement errors is equal to 0. Combined with Assumption S1, it implies that $\mathbb{E}[U^Y_{a_Y}|A_Y] = \mathbb{E}[U^T_{a_T}|A_T] = 0$ for all $(a_Y, a_T) \in \mathcal{A}^2$.

Proposition S1. *Under Assumptions S1 to S5, for all $a_Y \in \mathcal{A}$ and $a_T \in \mathcal{A}$,*

$$\lambda^Y_{a_Y} = \frac{\mathbb{E}[Y|A_Y = a_Y]}{\mathbb{E}[Y|A_Y = a^*_Y]} \quad \text{and} \quad \lambda^T_{a_T} = \frac{\mathbb{E}[T|A_T = a_T]}{\mathbb{E}[T|A_T = a^*_T]}$$

Proof. We prove the result for $\lambda^Y_{a_Y}$ and note that the result for $\lambda^T_{a_T}$ holds using analogous arguments. For any $a_Y \in \mathcal{A}$,

$$\begin{aligned}\mathbb{E}[Y|A_Y = a_Y] &= \lambda^Y_{a_Y}\mathbb{E}[Y^*|A_Y = a_Y] \\ &= \lambda^Y_{a_Y}\mathbb{E}[Y^*] \end{aligned} \tag{S4}$$

where the first equality holds by Assumptions S1 and S5, and the second equality holds by Assumption S1. Equation (S4) combined with Assumption S4, implies that

$$\mathbb{E}[Y|A_Y = a^*_Y] = \mathbb{E}[Y^*] \tag{S5}$$

Dividing Equation (S4) by Equation (S5) implies the result. □

Given the result in Proposition S1 that $\lambda^Y_{a_Y}$ and $\lambda^T_{a_T}$ are identified, we next define

$$\begin{aligned}\tilde{Y}_i &:= \frac{Y_{A_Y,i}}{\lambda^Y_{A_Y}} = Y^*_i + \tilde{U}^Y_{A_Y,i} \\ \tilde{T}_i &:= \frac{T_{A_T,i}}{\lambda^T_{A_T}} = T^*_i + \tilde{U}^T_{A_T,i}\end{aligned}$$

where $\tilde{U}^Y_{A_Y,i} := \frac{U^Y_{A_Y,i}}{\lambda^Y_{A_Y}}$ and $\tilde{U}^T_{A_T,i} := \frac{U^T_{A_T,i}}{\lambda^T_{A_T}}$. For children, $\tilde{Y}_i$ comes from dividing observed child's income by $\lambda^Y_{a_Y}$ corresponding to the age at which child's income is actually observed; likewise, for parents, $\tilde{T}_i$ comes from dividing observed parents' incomes by $\lambda^T_{a_T}$ corresponding to the age at which parents' income is actually observed.

In the next result, we show that, under the conditions in this section, $\tilde{U}^Y_{A_Y}$ and $\tilde{U}^T_{A_T}$ (the transformed versions of the measurement errors) meet the same requirements imposed on the measurement errors as in the main text in Assumption 1. This implies that, in the context of life-cycle measurement error, $\tilde{Y}_i$ and $\tilde{T}_i$ can be used in place of Y_i and T_i to estimate the nonlinear measures of intergenerational mobility considered in the main text.

Proposition S2. *Under Assumptions S1 to S5,*

$$(\tilde{U}^Y_{A_Y}, \tilde{U}^T_{A_T}) \perp\!\!\!\perp (Y^*, T^*, X) \quad \text{and} \quad \tilde{U}^Y_{A_Y} \perp\!\!\!\perp \tilde{U}^T_{A_T}$$

Proof. For any two functions g and h such that the expectation exists, notice that

$$\begin{aligned}
&\mathbb{E}\big[g(\tilde{U}^Y_{A_Y}, \tilde{U}^T_{A_T})h(Y^*, T^*, X)\big] \\
&\quad = \sum_{a_Y \in \mathcal{A}} \sum_{a_T \in \mathcal{A}} \mathbb{E}\left[g\left(\frac{U^Y_{a_Y}}{\lambda^Y_{a_y}}, \frac{U^T_{a_T}}{\lambda^T_{a_T}} \right) h(Y^*, T^*, X) \Big| A_Y = a_Y, A_T = a_T \right] \mathrm{P}(A_Y = a_y, A_T = a_t) \\
&\quad = \sum_{a_Y \in \mathcal{A}} \sum_{a_T \in \mathcal{A}} \mathbb{E}\left[g\left(\frac{U^Y_{a_Y}}{\lambda^Y_{a_y}}, \frac{U^T_{a_T}}{\lambda^T_{a_T}} \right) h(Y^*, T^*, X) \right] \mathrm{P}(A_Y = a_y, A_T = a_t) \\
&\quad = \mathbb{E}\big[h(Y^*, T^*, X)\big] \sum_{a_Y \in \mathcal{A}} \sum_{a_T \in \mathcal{A}} \mathbb{E}\left[g\left(\frac{U^Y_{a_Y}}{\lambda^Y_{a_y}}, \frac{U^T_{a_T}}{\lambda^T_{a_T}} \right) \right] \mathrm{P}(A_Y = a_y, A_T = a_t) \\
&\quad = \mathbb{E}\big[h(Y^*, T^*, X)\big] \sum_{a_Y \in \mathcal{A}} \sum_{a_T \in \mathcal{A}} \mathbb{E}\left[g\left(\frac{U^Y_{a_Y}}{\lambda^Y_{a_y}}, \frac{U^T_{a_T}}{\lambda^T_{a_T}} \right) \Big| A_Y = a_Y, A_T = a_T \right] \mathrm{P}(A_Y = a_y, A_T = a_t) \\
&\quad = \mathbb{E}\big[h(Y^*, T^*, X)\big] \mathbb{E}\left[g\left(\tilde{U}^Y_{A_Y}, \tilde{U}^T_{A_T} \right) \right]
\end{aligned}$$

where the first equality holds by the law of iterated expectations, the second equality holds by Assumption S1, the third equality holds by Assumption S2 and because $\mathbb{E}\big[h(Y^*, T^*, X)\big]$ does not depend on a_Y or a_T, the fourth equality holds by Assumption S1, and the last equality holds by the law of iterated expectations. This proves the first part of the result.

For the second part of the result, for any two functions g and h such that the expectation exists, notice that

$$\begin{aligned}
\mathbb{E}\left[g\left(\tilde{U}^Y_{A_Y} \right) h\left(\tilde{U}^T_{A_T} \right) \right] &= \sum_{a_Y \in \mathcal{A}} \sum_{a_T \in \mathcal{A}} \mathbb{E}\left[g\left(\frac{U^Y_{a_Y}}{\lambda^Y_{a_y}} \right) h\left(\frac{U^T_{a_T}}{\lambda^T_{a_T}} \right) \Big| A_Y = a_Y, A_T = a_T \right] \mathrm{P}(A_Y = a_y, A_T = a_T) \\
&= \sum_{a_Y \in \mathcal{A}} \sum_{a_T \in \mathcal{A}} \mathbb{E}\left[g\left(\frac{U^Y_{a_Y}}{\lambda^Y_{a_y}} \right) h\left(\frac{U^T_{a_T}}{\lambda^T_{a_T}} \right) \right] \mathrm{P}(A_Y = a_y, A_T = a_T) \\
&= \sum_{a_Y \in \mathcal{A}} \sum_{a_T \in \mathcal{A}} \mathbb{E}\left[g\left(\frac{U^Y_{a_Y}}{\lambda^Y_{a_y}} \right) \right] \mathbb{E}\left[h\left(\frac{U^T_{a_T}}{\lambda^T_{a_T}} \right) \right] \mathrm{P}(A_Y = a_y, A_T = a_T) \\
&= \sum_{a_Y \in \mathcal{A}} \sum_{a_T \in \mathcal{A}} \mathbb{E}\left[g\left(\tilde{U}^Y_{A_Y} \right) \right] \mathbb{E}\left[h\left(\frac{U^T_{a_T}}{\lambda^T_{a_T}} \right) \right] \mathrm{P}(A_Y = a_y, A_T = a_T) \\
&= \mathbb{E}\left[g\left(\tilde{U}^Y_{A_Y} \right) \right] \sum_{a_T \in \mathcal{A}} \mathbb{E}\left[h\left(\frac{U^T_{a_T}}{\lambda^T_{a_T}} \right) \right] \mathrm{P}(A_T = a_T) \\
&= \mathbb{E}\left[g\left(\tilde{U}^Y_{A_Y} \right) \right] \sum_{a_T \in \mathcal{A}} \mathbb{E}\left[h\left(\frac{U^T_{a_T}}{\lambda^T_{a_T}} \right) \Big| A_T = a_T \right] \mathrm{P}(A_T = a_T) \\
&= \mathbb{E}\left[g\left(\tilde{U}^Y_{A_Y} \right) \right] \mathbb{E}\left[h\left(\tilde{U}^T_{A_T} \right) \right]
\end{aligned}$$

where the first equality holds by the law of iterated expectations, the second equality holds by Assumption S1, the third equality holds by the first part of Assumption S3, the fourth equality holds by the second part of Assumption S3, the fifth equality holds because $\mathbb{E}\big[g(\tilde{U}^Y_{A_Y})\big]$ can be moved outside of the summations and from marginalizing the joint distribution of A_Y and A_T, the sixth equality holds by Assumption S1, and the last equality holds by the law of iterated expectations. □

SD Additional Empirical Results

This appendix contains additional results for the application. In particular, we provide the same results as in the main text but (i) for an alternative copula and (ii) for an alternative set of covariates.

Gaussian Copula

This section reports the same results as in the main text, but for a Gaussian conditional copula rather than a Clayton conditional copula. First, in this case, we estimate the rank-rank correlation of son's and father's permanent incomes to be 0.481 (s.e.=0.130), which is somewhat larger in magnitude (indicating less intergenerational income mobility) than in the case with a Clayton copula.

Table S1: Transition Matrix using a Gaussian Copula

		Father's Income Quartile			
		1	2	3	4
Son's Income Quartile	4	0.076	0.165	0.275	0.483
		(0.043)	(0.029)	(0.007)	(0.068)
	3	0.168	0.255	0.298	0.279
		(0.030)	(0.007)	(0.023)	(0.007)
	2	0.279	0.294	0.261	0.166
		(0.007)	(0.023)	(0.006)	(0.030)
	1	0.477	0.285	0.166	0.072
		(0.069)	(0.008)	(0.030)	(0.041)

Notes: The table provides estimates of a transition matrix allowing for measurement error as in the main text, but using a Gaussian copula rather than a Clayton copula. The columns are organized by quartiles of father's income; e.g., columns labeled "1" use data from fathers whose income is in the first quartile. Similarly, rows are organized by quartiles of son's income. Standard errors are computed using the bootstrap.

Next, Table S1 reports our estimate of the transition matrix using a Gaussian copula. These results are broadly similar to the ones under a Clayton copula and, if anything, indicate somewhat lower overall estimates of intergenerational mobility using a Gaussian copula rather than a Clayton copula.

Table S2: Upward Mobility using a Gaussian Copula

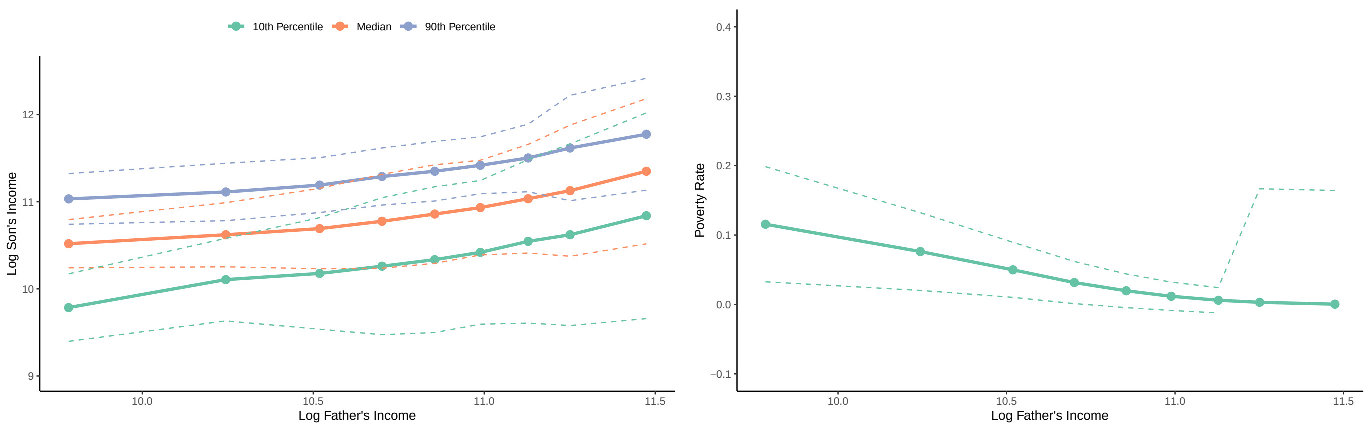
	Father's Income Quartile			
	1	2	3	4
Measurement Error				
	0.757	0.566	0.424	0.245
	(0.034)	(0.014)	(0.016)	(0.035)

Notes: The table provides estimates of upward mobility allowing for measurement error as in the main text but using a Gaussian copula rather than a Clayton copula. The columns are organized by quartiles of father's income; e.g., columns labeled "1" use data from fathers whose income is in the first quartile. Standard errors are computed using the bootstrap.

Next, Table S2 provides estimates of upward mobility by quartiles of father's permanent income. Again, these results are broadly similar to those arising from using a Clayton copula. Relative to the estimates in the main text, using a Gaussian copula, sons whose fathers were in the second quartile of the permanent income distribution are somewhat less likely to have higher permanent income ranks than their fathers, while sons whose fathers were in the top quartile of the permanent income distribution are somewhat more likely to have higher permanent income ranks than their fathers. Otherwise, the estimates are very similar.

Finally, Figure S1 provides estimates of the conditional quantiles of son's permanent income and the poverty rate, both as a function of father's permanent income. Relative to the results in the main text, using a Gaussian copula, the estimates of conditional quantiles of son's permanent income as a function of father's permanent income are broadly similar — in this case, the slopes of the conditional quantiles are somewhat larger than in the main text, and the standard errors are also somewhat larger. Using a Gaussian copula, the estimates of the poverty rate for sons whose fathers were in the lower and middle parts of the permanent income distribution are somewhat higher (though in both cases, the standard errors are quite large); for sons whose fathers were in the upper part of the permanent income distribution, the estimates of the poverty rate are similar to those in the main text. At especially high values of father's permanent income, although the estimated poverty rate is similar to that reported in the main text, the reported confidence intervals are wide. This seems to arise because, after adjusting for measurement error, the 80th and 90th percentiles of father's permanent income are close to the top of the distribution of father's observed income, which results in numerically unstable results conditional on high values of father's permanent income in some bootstrap iterations.

Figure S1: Quantiles and Poverty Rates for Sons using Gaussian Copula



(a) Quantiles of Son's Income as a Function of Father's Income

(b) Poverty Rate as a Function of Father's Income

Notes: The figure provides (i) estimates of quantiles of son's income as a function of father's income and (ii) estimates of the poverty rate of sons as a function of father's income. In both cases, estimates are provided for the 10th, 20th, . . . , and 90th percentiles of father's income. Both estimates are conditional on son's age and father's age being equal to their averages in the sample and account for measurement error using the approach suggested in the paper. The difference relative to the estimates in the main text is that the copula is Gaussian rather than a Clayton copula. Standard errors are computed using the bootstrap.

Additional Covariates

Finally, we provide results that additionally include son's race and father's years of education as additional covariates in the first step quantile regressions. In this section, we only report results for the copula-type parameters, which are directly comparable to the estimates in the main text; the conditional distribution-type parameters are not comparable because they condition on different sets of covariates, and, therefore, we do not report them in this section. Besides using a different set of covariates, the estimates in this section come from the same estimation procedure in the main text; most notably, as in the main text, we use a Clayton copula in this section.

First, our estimate of the rank-rank correlation in this case is 0.248 (s.e.=0.102). This is somewhat smaller than the corresponding estimate in the main text and is between that estimate and our estimate when we ignore measurement error.

Table S3: Transition Matrix including Additional Covariates

		Father's Income Quartile			
		1	2	3	4
Son's Income Quartile	4	0.155	0.218	0.269	0.358
		(0.034)	(0.013)	(0.011)	(0.037)
	3	0.214	0.247	0.266	0.274
		(0.029)	(0.006)	(0.015)	(0.014)
	2	0.262	0.267	0.253	0.218
		(0.010)	(0.016)	(0.006)	(0.016)
	1	0.369	0.268	0.212	0.150
		(0.069)	(0.009)	(0.029)	(0.035)

Notes: The table provides estimates of a transition matrix allowing for measurement error as in the main text, but where the first step quantile regressions additionally include son's race and father's education as covariates. The columns are organized by quartiles of father's income; e.g., columns labeled "1" use data from fathers whose income is in the first quartile. Similarly, rows are organized by quartiles of son's income. Standard errors are computed using the bootstrap.

Next, Table S3 reports a transition matrix. These results are similar to the ones reported in the main text. These results possibly indicate slightly more intergenerational mobility than the corresponding estimates in the main text. The main difference is that we estimate that sons whose father's permanent income is in the first quartile are somewhat less likely to stay in the first quartile of the permanent income distribution (though more likely to stay in the first quartile than in the case where we ignore measurement error). The other estimates, particularly for sons whose father's permanent income is in the top quartile, are quite similar to the results in the main text.

Finally, Table S4 reports upward mobility estimates as a function of father's permanent income quartile. These estimates are very similar to the estimates reported in the main text.

Table S4: Upward Mobility including Additional Covariates

	Father's Income Quartile			
	1	2	3	4
Measurement Error				
	0.811	0.597	0.401	0.183
	(0.033)	(0.007)	(0.018)	(0.017)

Notes: The table provides estimates of upward mobility allowing for measurement error as in the main text, but where the first step quantile regressions additionally include son's race and father's education as covariates. The columns are organized by quartiles of father's income; e.g., columns labeled "1" use data from fathers whose income is in the first quartile. Standard errors are computed using the bootstrap.

References

An, Yonghong, Le Wang, and Ruli Xiao (2020). "A nonparametric nonclassical measurement error approach to estimating intergenerational mobility elasticities". *Journal of Business & Economic Statistics*, pp. 1–17.

Flores, Carlos A (2007). "Estimation of dose-response functions and optimal doses with a continuous treatment". Working Paper.

Flores, Carlos A, Alfonso Flores-Lagunes, Arturo Gonzalez, and Todd C Neumann (2012). "Estimating the effects of length of exposure to instruction in a training program: The case of Job Corps". *Review of Economics and Statistics* 94.1, pp. 153–171.

Frolich, Markus and Blaise Melly (2013). "Unconditional quantile treatment effects under endogeneity". *Journal of Business & Economic Statistics* 31.3, pp. 346–357.

Galvao, Antonio F and Liang Wang (2015). "Uniformly semiparametric efficient estimation of treatment effects with a continuous treatment". *Journal of the American Statistical Association* 110.512, pp. 1528–1542.

Haavelmo, Trygve (1944). "The probability approach in econometrics". *Econometrica*, pp. iii–115.

Haider, Steven and Gary Solon (2006). "Life-cycle variation in the association between current and lifetime earnings". *American Economic Review* 96.4, pp. 1308–1320.

Hirano, Keisuke and Guido W Imbens (2004). "The propensity score with continuous treatments". *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives* 226164, pp. 73–84.

Hu, Yingyao and Yuya Sasaki (2015). "Closed-form estimation of nonparametric models with non-classical measurement errors". *Journal of Econometrics* 185.2, pp. 392–408.

Jenkins, Stephen (1987). "Snapshots versus movies: 'Lifecycle biases' and the estimation of intergenerational earnings inheritance". *European Economic Review* 31.5, pp. 1149–1158.

Nybom, Martin and Jan Stuhler (2016). "Heterogeneous income profiles and lifecycle bias in intergenerational mobility estimation". *Journal of Human Resources* 51.1, pp. 239–268.

Powell, David (2016). "Quantile regression with nonadditive fixed effects". Working Paper.

Prakasa Rao, Bhagavatula LS (2009). "Conditional independence, conditional mixing and conditional association". *Annals of the Institute of Statistical Mathematics* 61.2, pp. 441–460.