

IV regression with distribution-valued outcomes

David Van Dijcke (U of Virginia) Kaspar Wüthrich (U of Michigan)

Midwestern Econometrics Group Conference 2026 Series

IV regression with distribution-valued outcomes

This paper. IV Fréchet regression (IVFR): IV when the outcome is an entire distribution

Setting.

- Units (groups): $j = 1, \dots, n$
- Outcome: $Y_j \in (\mathcal{Y}, W_2)$ (space of CDFs with Wasserstein distance W_2)
- Regressors: $X_j \in \mathbb{R}^p$ (includes treatment and covariates)
- Instruments: $Z_j \in \mathbb{R}^l, l \geq p$

Example. (Autor et al., 2013; Chetverikov et al., 2016)

- Unit: commuting zone (CZ)
- Outcome: distribution of wages in CZ
- Regressors: Chinese import competition (and other CZ characteristics)
- Instruments: Chinese import competition in same industry but other developed countries

Group-level IV in top-5 journals. About one in six top-5 IV papers 2016–2026 (prelim.)

Literature

	Scalar outcome	Distribution-valued outcome
Exogenous X	OLS / quantile regression (Koenker and Bassett, 1978)	Fréchet / Wasserstein regression (e.g., Petersen and Müller, 2019; Petersen et al., 2021; Lin et al., 2023; Kurisu et al., 2024; Van Dijkke, 2025; Song et al., 2026)
Endogenous X	2SLS / IVQR, LQTE (e.g., Abadie et al., 2002; Chernozhukov and Hansen, 2005; Lee, 2007; Frölich and Melly, 2013; Wüthrich, 2020)	This paper: IV Fréchet regression

Grouped and panel quantile models (e.g., Koenker, 2004; Canay, 2011; Galvao and Wang, 2015; Galvao and Kato, 2016; Gunsilius, 2023; Chiang et al., 2026)

Closest: grouped quantile IV (Chetverikov et al., 2016; Melly and Pons, 2025)

Functional VAR/LP (e.g., Chang et al., 2024; Huber et al., 2024; Bjørnland et al., 2026)

Distributions as time-series objects; here: IV with i.i.d. groups, Wasserstein geometry with no basis approximation

Model

Let $Q_{Y_j} \in \mathcal{Q}$ (space of quantile functions)

Linear in group covariates *at each quantile*. Coefficients are functions of u :

$$Q_{Y_j}(u) = \beta_0(u) + \beta_1(u)^\top (X_j - \mu_X) + \eta_j(u), \quad E[\eta_j(u)] = 0$$

Target structural quantile function: $q(x, u) := \beta_0(u) + \beta_1(u)^\top (x - \mu_X)$

IV exogeneity: $E[\tilde{Z}_j \eta_j(u)] = 0$, all u , where $\tilde{Z}_j = Z_j - \mu_Z$

Full rank: $\Sigma_{ZZ} \succ 0$, Σ_{ZX} full column rank

Background: Fréchet mean and Fréchet regression (Petersen and Müller, 2019)

Let (Ω, d) be a metric space, $Y_j \in \Omega$ a random object, and $X_j \in \mathbb{R}^p$ a vector of covariates. Denote the Euclidean distance by d_E .

Fréchet mean on (Ω, d) : $\omega_{\oplus} = \arg \min_{\omega} E[d^2(Y_j, \omega)]$

- In $(\mathcal{Y}, W_2)$: $W_2^2(\mu, \nu) = \int_0^1 (Q_{\mu}(u) - Q_{\nu}(u))^2 du$, mean has QF $E[Q_Y(u)]$

Fréchet regression: the linear conditional mean $m(x) = \alpha_0 + \alpha_1^T(x - \mu_X)$ obtains as

$$m(x) = \arg \min_m E[s(X, x) d_E^2(Y, m)], \quad s(X, x) = 1 + (X - \mu_X)^T \Sigma_{XX}^{-1} (x - \mu_X)$$

Idea of Fréchet regression: replace d_E with general d for metric-space-valued outcomes.

Requires *exogenous* $X \Rightarrow$ we provide IV analogue.

Motivating IVFR

Real-valued IV setup: $Y = \alpha_0 + \alpha_1^T(X - \mu_X) + \eta$, $E[\tilde{Z}\eta] = 0$. The 2SLS regression function can be written as

$$m(x) = \arg \min_m E[s(Z, x) d_E^2(Y, m)]$$

with IV weight

$$s(Z, x) = 1 + (x - \mu_X)^T (\Sigma_{ZX}^T \Sigma_{ZZ}^{-1} \Sigma_{ZX})^{-1} \Sigma_{ZX}^T \Sigma_{ZZ}^{-1} \tilde{Z}.$$

Distribution-valued case: replace d_E with W_2 ,

$$m^{\text{IVFR}}(x) = \arg \min_{m \in \mathcal{Y}} E[s(Z, x) W_2^2(Y, m)].$$

Key point: $s(Z, x)$ can be negative, so the IV-weighted average quantile curve need not be monotone.

Fréchet characterization and regression coefficients

IV-weighted quantile curve: $\psi_x(u) := E[s(Z, x) Q_Y(u)]$

L^2 projection onto $\mathcal{Q}$: $\Pi_{\mathcal{Q}}$

Proposition (Fréchet characterization).

$$Q_{m^{\text{IVFR}}(x)} = \Pi_{\mathcal{Q}}(\psi_x).$$

$Q_{m^{\text{IVFR}}(x)}$ represents the **IV-weighted (signed) Wasserstein barycenter** in quantile space.

Recover regression coefficients by population OLS:

$$\beta_0^{\text{IVFR}}(u) = E[Q_{m^{\text{IVFR}}(X)}(u)],$$

$$\beta_1^{\text{IVFR}}(u) = \Sigma_{XX}^{-1} E[(X - \mu_X) Q_{m^{\text{IVFR}}(X)}(u)].$$

Identification

By linearity of $s(Z, x)$ in x ,

$$\psi_x(u) = \beta_0^{\text{unc}}(u) + \beta_1^{\text{unc}}(u)^T(x - \mu_X),$$

where $\beta^{\text{unc}}(u)$ is exactly quantile-by-quantile population 2SLS.

Identification lemma. Under IV exogeneity, full rank, and the linear model:

$$Q_{m^{\text{IVFR}}(x)} = \Pi_{\mathcal{Q}}(\psi_x) = \psi_x = q(x, \cdot),$$

$$\beta^{\text{IVFR}}(u) = \beta^{\text{unc}}(u) = \beta(u).$$

Under correct specification: the population curve is already monotone, so projection is inactive and the coefficients recover the structural effects.

Under misspecification: project to a valid quantile function, then summarize it by OLS.

What does $\beta_1(u)$ identify?

With binary X and exogeneity,

$$\beta_1(u) = E[Q_{Y_j(1)}(u) - Q_{Y_j(0)}(u)] = \text{total effect on the group's } u\text{-quantile.}$$

Total effect can be decomposed into three channels. Relative to an individual-level model $Q_{ij}(u \mid W_{ij}, X_j) = W_{ij}^T \gamma(u) + X_j \delta(u) + \alpha_j(u)$,

$$\beta_1(u) = \underbrace{\delta(u)}_{\text{direct response}} + \underbrace{\text{composition}(u)}_{\text{who is in the group}} + \underbrace{\text{re-ranking}(u)}_{\text{group vs. within-type quantile}}$$

Relationship to Chetverikov et al. (2016) and Melly and Pons (2025):

- With individual covariates W_{ij} : they target the direct effect $\delta(u)$.
- Without them: they target the same total effect as unprojected IVFR.

The IVFR estimator

1. Estimate IV weights $\hat{s}_i(Z_i, x)$ using sample moments

2. Average quantile curves: $\hat{\psi}_{X_j}(u) = \frac{1}{n} \sum_{i=1}^n \hat{s}_i(Z_i, X_j) Q_{Y_i}(u)$

- If Q_{Y_i} is unknown, replace it by empirical QF (justified under high-level conditions).

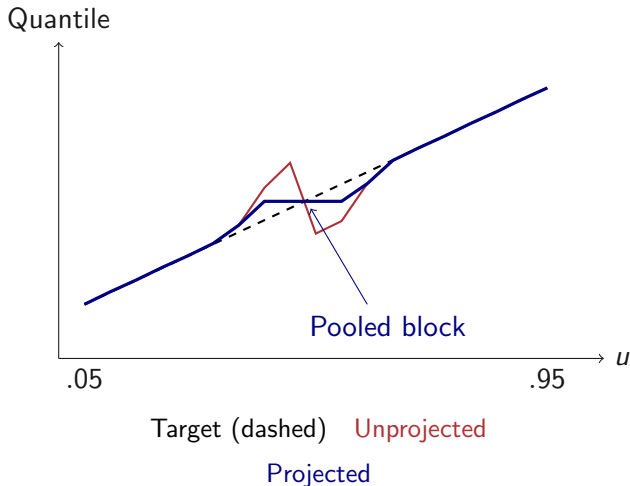
3. **Project to monotonicity:** $\hat{Q}(X_j, \cdot) = \Pi_{\mathcal{Q}}(\hat{\psi}_{X_j})$ (PAVA)

4. Regress projected curves on X :

$$\hat{\beta}_0^{\text{IVFR}}(u) = \frac{1}{n} \sum_{j=1}^n \hat{Q}(X_j, u),$$

$$\hat{\beta}_1^{\text{IVFR}}(u) = \hat{\Sigma}_{XX}^{-1} \frac{1}{n} \sum_{j=1}^n (X_j - \bar{X}) \hat{Q}(X_j, u).$$

What does the projection do?



- Negative IV weights can make the fit locally decrease
- PAVA pools adjacent violations and leaves the rest unchanged

Finite-sample improvement I: quantile function

Lemma (Robertson et al., 1988; Lin et al., 2019)

For any $\hat{Q}$ and $Q_0 \in \mathcal{Q}$:

$$\|\Pi_{\mathcal{Q}}\hat{Q} - Q_0\|_{L^p} \leq \|\hat{Q} - Q_0\|_{L^p}, \quad p \in [1, \infty].$$

- Applied to IVFR: $\hat{Q}(X_j, \cdot)$ closer to $q(X_j, \cdot)$ than $\hat{\psi}_{X_j}$
- Unlike projection, rearrangement (Chernozhukov et al., 2010) does not naturally arise in this framework and does not yield the (IV-weighted) barycenter interpretation.

Finite-sample improvement II: coefficients

Let $\tilde{\beta} = (\tilde{\beta}_0, \tilde{\beta}_1)$ denote quantile-by-quantile 2SLS without projection. For any target $b = (b_0, b_1)$ with $q_b(X_j, \cdot) \in \mathcal{Q}$:

Theorem (joint improvement)

$$\|\hat{\beta}_0^{\text{IVFR}} - b_0\|_{L^2}^2 + \|\hat{\beta}_1^{\text{IVFR}} - b_1\|_{\hat{\Sigma}_{XX}, L^2}^2 \leq \|\tilde{\beta}_0 - b_0\|_{L^2}^2 + \|\tilde{\beta}_1 - b_1\|_{\hat{\Sigma}_{XX}, L^2}^2.$$

- Under correct specification, set $b = \beta$.
- Holds without the linear model being true

Asymptotics: main assumptions

(A1) Sampling. $\{(Z_j, X_j, Y_j)\}$ iid *across* groups, within-group unrestricted

(A2) 4th moments. $\sup_u E[Q_Y(u)^4], E\|Z\|^4, E\|X\|^4 < \infty$

(A3) Bounded support. $\sup_x \|x - \mu_X\| \leq B$

(A4) Average Strictness. $\exists \kappa > 0: E[s(Z, x)\{Q_Y(d) - Q_Y(c)\}] \geq \kappa(d - c)$

(A5) Average Lipschitz. $\exists K: E[s(Z, x)\{Q_Y(d) - Q_Y(c)\}] \leq K(d - c)$

(A2)–(A3): implied by assumptions in Chetverikov et al. (2016) and Melly and Pons (2025)

(A4)–(A5): population IV-weighted curve has secant slopes in $[\kappa, K]$. Individual group distributions may have flats/atoms!

Functional CLTs and bootstrap validity

Theorem (FCLTs)

With full-rank IV and (A1)–(A5): $\sqrt{n}(\Pi_Q \hat{\psi}_x - \psi_x) \rightsquigarrow \mathbb{G}_x$ in $\ell^\infty([a, b])$ and $\sqrt{n}(\hat{\beta}^{\text{IVFR}} - \beta^{\text{IVFR}}) \rightsquigarrow \mathbb{G}_\beta$ in $\ell^\infty([a, b])^{p+1}$

- Under Average Strictness (A4) and Average Lipschitz (A5), monotonicity constraint is nonbinding in the limit $\Rightarrow$ same limit as quantile-by-quantile 2SLS.
- Binding constraints instead give directional (generally non-Gaussian) limits and require directional bootstrap (Fang and Santos, 2019).
- Holds under misspecification.
- **Pointwise inference:** analytical sandwich SEs
- **Uniform inference:** multiplier bootstrap, with PAVA re-applied in every draw.

Simulations: reduction in IMSE (in %) due to projection

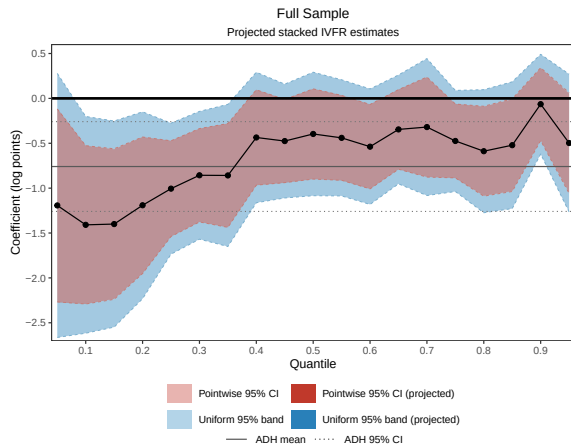
Median first-stage F	Scalar: $n = N = 50, p = 1$		Realistic: $n = 50, N = 25, p = 5$	
	Mean gain	Median gain	Mean gain	Median gain
≈ 1	42.1	7.3	53.5	14.1
≈ 3	52.8	1.6	58.3	4.8
≈ 5	47.6	0.5	74.8	2.1
≈ 10	49.6	0.0	45.8	0.9
≈ 20	0.2	0.0	1.6	0.3

- Projection protects against unstable fits and extreme losses.

► simulation design

Import competition and U.S. wages (Chetverikov et al., 2016)

$n = 1,444$, $F \approx 533$.



Full-sample estimates and bands nearly coincide with unprojected results.

Uniform significance:
10th–35th percentiles.

Projection lowers IMSE by up to 26.8% in small CZ subsamples (500 draws).

Conclusion

IV Fréchet regression: IV when the outcome is an entire distribution

- Quantile-by-quantile 2SLS + projection onto valid quantile functions
- Same target and same limit as quantile-by-quantile 2SLS under the linear model; never worse in finite samples, up to 75% lower IMSE
- Multiplier bootstrap for uniform bands; valid under misspecification

What else is in the paper?

- Functional Anderson–Rubin inference for weak instruments

Thank you!

Appendix: finite-sample improvements for individual coefficients

For a target b with monotone q_b , let r_{jk} be regressor k after partialling out the other centered regressors, with $\sum_j r_{jk}^2 > 0$.

Sufficient condition: removing the isolated slope error restores monotonicity

$$\underbrace{\hat{\psi}_{X_j}(\cdot) - r_{jk}(\tilde{\beta}_{1,k} - b_{1,k})(\cdot)}_{\text{fitted curve with the isolated coefficient set to its target}} \in \mathcal{Q} \quad \text{for every } j \text{ with } r_{jk} \neq 0.$$

Corollary (coefficient-wise gain). Under this condition, for any target b ,

$$\|\hat{\beta}_{1,k}^{\text{IVFR}} - b_{1,k}\|_{L^2}^2 \leq \|\tilde{\beta}_{1,k} - b_{1,k}\|_{L^2}^2.$$

Intuition: other estimation errors do not create violations on their own.

Appendix: improvements without sufficient condition

Residualize the k th centered regressor on the other regressors: r_{jk} , with $\hat{v}_k = n^{-1} \sum_j r_{jk}^2 > 0$, $J_k = \{j : r_{jk} \neq 0\}$.

$$\hat{\psi}_{X_j} = q_b(X_j, \cdot) + e_{jk} + r_{jk}(\tilde{\beta}_{1,k} - b_{1,k}), \quad D_{X_j} = \Pi_{\mathcal{Q}} \hat{\psi}_{X_j} - \hat{\psi}_{X_j}.$$

$$\begin{aligned} \|\hat{\beta}_{1,k}^{\text{IVFR}} - b_{1,k}\|_{L^2}^2 &\leq \|\tilde{\beta}_{1,k} - b_{1,k}\|_{L^2}^2 - \frac{\sum_{j \in J_k} \|D_{X_j}\|_{L^2}^2}{n\hat{v}_k} \\ &\quad - \frac{2 \sum_{j \in J_k} \langle D_{X_j}, e_{jk} \rangle_{L^2}}{n\hat{v}_k}. \end{aligned}$$

- e_{jk} collects intercept and nuisance-slope error. The cross-term can have either sign
- If $q_b(X_j, \cdot) + e_{jk} \in \mathcal{Q}$ for every $j \in J_k$, coefficient-wise improvement follows

Joint improvement does not require this additional condition.

Appendix: simulation design

DGP: CLP benchmark, omitting individual covariates (Chetverikov et al., 2016)

Individual outcomes within group j , and endogenous treatment:

$$Y_{ij} = b_0(U_{ij}) + \frac{U_{ij}}{2} + X_j \gamma(U_{ij}) + (\zeta_j - \frac{1}{2})U_{ij} + \sum_{k=1}^{p-1} W_{jk} h_k(U_{ij}), \quad X_j^0 = \{\pi_Z + \delta(\zeta_j - \frac{1}{2})\}Z_j + \zeta_j + \nu_j.$$

$U_{ij}, \zeta_j \sim U(0, 1)$, $Z_j, \nu_j \sim \exp(0.25 \mathcal{N}(0, 1))$, with independent primitive draws. The same ζ_j enters treatment and outcomes. π_Z controls relevance.

CLP special case: $b_0 = 0$, $\gamma(u) = \sqrt{u}$, $\pi_Z = 1$, $\delta = 0$, $p = 1$, $X_j = X_j^0$, so $Y_{ij} = X_j \sqrt{U_{ij}} + \zeta_j U_{ij}$.

Extensions: weaker instruments, controls, heterogeneous first stages, and nonmonotone effect curves. Bound treatment and scale b_0 to preserve valid group quantiles.

IMSE design

- 10,000 replications, 19 empirical group quantiles from .05 to .95, all draws retained
- Scalar: $n = N = 50$, $p = 1$. Realistic: $n = 50$, $N = 25$, $p = 5$, mixed controls, lognormal outcomes, curvature 0.2, $\delta = \pi_Z/1.3$
- Gains compare replication-level means or medians, not mean percentage gains
- Why mean gains can be large: at scalar $F \approx 5$, the largest 1% of unprojected losses give 99.7% of the MC mean

References I

- Abadie, A., Angrist, J., Imbens, G., 2002. Instrumental variables estimates of the effect of subsidized training on the quantiles of trainee earnings. *Econometrica* 70, 91–117. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/1468-0262.00270>, doi:<https://doi.org/10.1111/1468-0262.00270>, arXiv:<https://onlinelibrary.wiley.com/doi/pdf/10.1111/1468-0262.00270>.
- Autor, D.H., Dorn, D., Hanson, G.H., 2013. The China Syndrome: Local Labor Market Effects of Import Competition in the United States. *American Economic Review* 103, 2121–68. URL: <https://www.aeaweb.org/articles?id=10.1257/aer.103.6.2121>, doi:[10.1257/aer.103.6.2121](https://doi.org/10.1257/aer.103.6.2121).
- Bjørnland, H.C., Chang, Y., Cross, J.L., 2026. Oil and the stock market revisited: A mixed functional var approach. *Quantitative Economics* 17, 541–589. URL: <https://onlinelibrary.wiley.com/doi/abs/10.3982/QE2358>, doi:<https://doi.org/10.3982/QE2358>, arXiv:<https://onlinelibrary.wiley.com/doi/pdf/10.3982/QE2358>.
- Canay, I.A., 2011. A simple approach to quantile regression for panel data. *The Econometrics Journal* 14, 368–386. URL: <https://doi.org/10.1111/j.1368-423X.2011.00349.x>, doi:[10.1111/j.1368-423X.2011.00349.x](https://doi.org/10.1111/j.1368-423X.2011.00349.x), arXiv:<https://academic.oup.com/ectj/article-pdf/14/3/368/27670447/ectj0368.pdf>.

References II

- Chang, M., Chen, X., Schorfheide, F., 2024. Heterogeneity and aggregate fluctuations. *Journal of Political Economy* 132, 4021–4067. URL: <https://www.journals.uchicago.edu/doi/abs/10.1086/7314111>, doi:10.1086/7314111, arXiv:<https://www.journals.uchicago.edu/doi/pdf/10.1086/7314111>.
- Chernozhukov, V., Fernández-Val, I., Galichon, A., 2010. Quantile and probability curves without crossing. *Econometrica* 78, 1093–1125.
- Chernozhukov, V., Hansen, C., 2005. An IV model of quantile treatment effects. *Econometrica* 73, 245–261. doi:10.1111/j.1468-0262.2005.00570.x.
- Chetverikov, D., Larsen, B., Palmer, C., 2016. IV quantile regression for group-level treatments, with an application to the distributional effects of trade. *Econometrica* 84, 809–833.
- Chiang, H.D., Galvao, A.F., Wei, C.M., 2026. Panel quantile regression with common shocks. arXiv preprint arXiv:2602.19201 .
- Fang, Z., Santos, A., 2019. Inference on directionally differentiable functions. *The Review of Economic Studies* 86, 377–412. URL: <https://doi.org/10.1093/restud/rdy049>, doi:10.1093/restud/rdy049.
- Frölich, M., Melly, B., 2013. Unconditional quantile treatment effects under endogeneity. *Journal of Business & Economic Statistics* 31, 346–357. doi:10.1080/07350015.2013.803869.

References III

- Galvao, A.F., Kato, K., 2016. Smoothed quantile regression for panel data. *Journal of Econometrics* 193, 92–112. URL: <https://www.sciencedirect.com/science/article/pii/S0304407616300239>, doi:<https://doi.org/10.1016/j.jeconom.2016.01.008>.
- Galvao, A.F., Wang, L., 2015. Efficient minimum distance estimator for quantile regression fixed effects panel data. *Journal of Multivariate Analysis* 133, 1–26. URL: <https://www.sciencedirect.com/science/article/pii/S0047259X14001894>, doi:<https://doi.org/10.1016/j.jmva.2014.08.007>.
- Gunsilius, F.F., 2023. Distributional synthetic controls. *Econometrica* 91, 1105–1117.
- Huber, F., Marcellino, M., Tornese, T., 2024. The distributional effects of economic uncertainty. *arXiv* 2411.12655. URL: <https://arxiv.org/abs/2411.12655>.
- Koenker, R., 2004. Quantile regression for longitudinal data. *Journal of Multivariate Analysis* 91, 74–89. URL: <https://www.sciencedirect.com/science/article/pii/S0047259X04001113>, doi:<https://doi.org/10.1016/j.jmva.2004.05.006>. special Issue on Semiparametric and Nonparametric Mixed Models.
- Koenker, R., Bassett, G., 1978. Regression quantiles. *Econometrica: journal of the Econometric Society* , 33–50.

References IV

- Kurusu, D., Zhou, Y., Otsu, T., Müller, H.G., 2024. Geodesic causal inference. arXiv preprint arXiv:2406.19604 .
- Lee, S., 2007. Endogeneity in quantile regression models: A control function approach. *Journal of Econometrics* 141, 1131–1158. URL: <https://www.sciencedirect.com/science/article/pii/S0304407607000279>, doi:<https://doi.org/10.1016/j.jeconom.2007.01.014>.
- Lin, L., Thomas, B.S., Piegorsch, W.W., Scott, J., Carvalho, C., 2019. A projection approach for multiple monotone regression. arXiv preprint arXiv:1911.07553 .
- Lin, Z., Kong, D., Wang, L., 2023. Causal inference on distribution functions. *Journal of the Royal Statistical Society Series B: Statistical Methodology* 85, 378–398.
- Melly, B., Pons, M., 2025. Minimum distance estimation of quantile panel data models. arXiv preprint arXiv:2502.18242 .
- Petersen, A., Liu, X., Divani, A.A., 2021. Wasserstein F-tests and confidence bands for the Fréchet regression of density response curves. *The Annals of Statistics* 49, 590–611.
- Petersen, A., Müller, H.G., 2019. Fréchet regression for random objects with Euclidean predictors. *The Annals of Statistics* 47, 691 – 719. URL: <https://doi.org/10.1214/17-AOS1624>, doi:10.1214/17-AOS1624.

References V

- Robertson, T., Wright, F.T., Dykstra, R.L., 1988. Order Restricted Statistical Inference. John Wiley & Sons, New York.
- Song, W., Dubey, P., Müller, H.G., Petersen, A., 2026. Inference for Fréchet regression. arXiv preprint arXiv:2605.19519 .
- Van Dijcke, D., 2025. Regression discontinuity design with distribution-valued outcomes. arXiv preprint arXiv:2504.03992 Frozen October 2025 version at https://www.davidvandijcke.com/files/r3d_oct2025.pdf; referenced results refer to this version.
- Wüthrich, K., 2020. A comparison of two quantile models with endogeneity. Journal of Business & Economic Statistics 38, 443–456. URL: <https://doi.org/10.1080/07350015.2018.1514307>, doi:10.1080/07350015.2018.1514307, arXiv:<https://doi.org/10.1080/07350015.2018.1514307>.