Regression Discontinuity Designs with a Continuous Treatment

Yingying Dong[1], Ying-Ying Lee[1], Michael Gou[2]

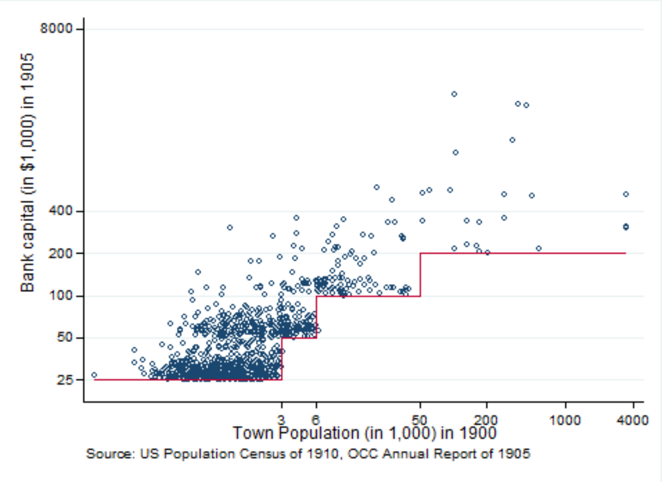
[1] University of California Irvine; [2] PWC

ASSA 2020

Motivating Example

- Do capital requirements promote bank stability?
  - Capital regulation is a principle tool for promoting bank stability.
- Take advantage of the discontinuity rule of the capital requirements in the early 20th century U.S.

Minimum capital requirements graded according to town population in the early 20th century.



Focus on the first threshold: Min Capital Req. $= \$25{,}000$ if Pop$<3{,}000$;

$= \$50{,}000$ if Pop>3000.

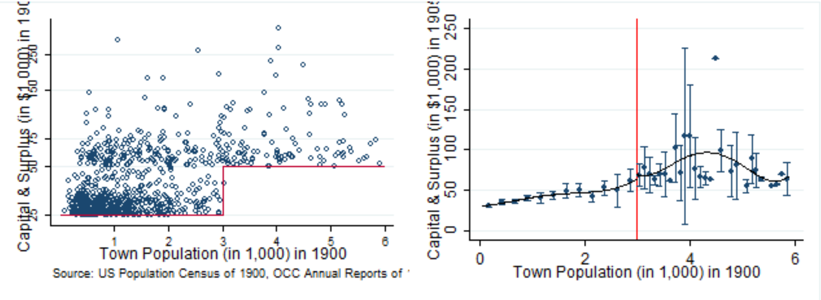


Figure: Scatter plot (left) and RD mean plot (right) around the first threshold of bank capital against town population

Motivating Example

Capital changes occur at the bottom 30% of the distribution; no significant change on average.

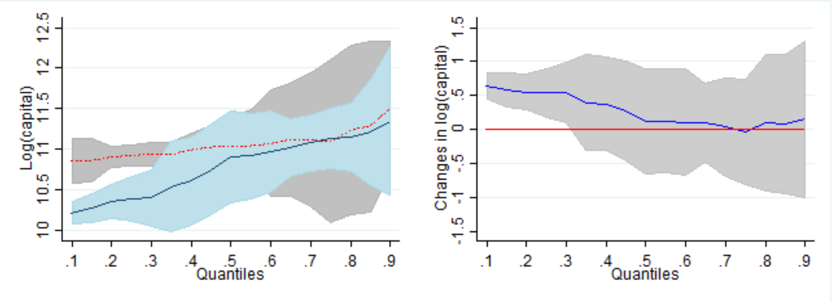


Figure: Quantile curves of bank capital below and above threshold (left) and quantile changes (right) with 95% CIs.

Motivation

- Notice the treatment, bank capital, is continuous.
- Tempted to apply the standard RD design (Hahn et al. 2001).
- Issues: 1. interpretation; 2. identification threat (little or no mean changes); 3. policy relevance.

Basic Idea

Let T be the treatment, R be the running variable and r_0 be a known policy threshold.

Write $T_i = a_i + b_i Z_i$, where $a_i \equiv T_{0i}$, $b_i \equiv T_{1i} - T_{0i}$, $Z_i = 1(R_i \geq r_0)$, T_{1i} (T_{0i}) is potential treatment above (below) the threshold.

E.g., T_{0i}: capital holdings right below the policy threshold;
T_{1i} : capital holdings right above.

- Quantiles of b_i capture exogenous distributional changes in the treatment T at r_0, under proper conditions.

Contributions

Provide new identification and inference theory for the class of RD designs with a continuous treatment:

- Identification relies on a distributional change in the first-stage, including the mean change as a special case.
- Can identify treatment effects at different treatment intensity.
- When the interest is in average effects, provide robust identification results:
  - identify a causal parameter under two alternative identifying assumptions
  - incorporate the standard RD estimand as a special case.

Contributions

- Provide bias-corrected robust inference (Calonico, Cattaneo, and Titiunik, 2014), along with AMSE optimal bandwidths (Imbens and Kalyanaraman, 2012).
- Apply the new estimators to estimate the impacts of bank capital, even though there is no exogenous mean change in bank capital.

Notation

- Y is the outcome of interest; can be either discrete or continuous.
- T is a continuous treatment;
- R is the continuous running variable that partly determines the treatment intensity.
- $Y = G(T, R, \varepsilon)$, where ε is the vector of (un)observables affecting outcome, and the dimension of ε is unrestricted.
- $T = q(R, U)$, where U embodies all relevant (un)observables affecting treatment other than R.

WLG, rewrite $T = q_1(R, U_1)Z + q_0(R, U_0)(1 - Z)$, where $Z = 1(R \geq r_0)$.

Identifying Assumptions

- Define $T_z \equiv q_z(R, U_z)$, $z = 0, 1$, so $T = T_1 Z + T_0(1 - Z)$.
- **Assumption 1** (Quantile representation) For $z = 0, 1$ and any $r \in \mathcal{R}$, the conditional distribution of T_z given $R = r$ is continuous with a strictly increasing CDF $F_{T_z|R}(T_z, r)$, and $q_z(r, u)$ is strictly monotonic in u.

Given Assumption 1, can let $U_z \equiv F_{T_z|R}(T_z, R) \sim Unif(0, 1)$ and so $q_z(R, U_z)$ is the quantile function of T_z.

Identifying Assumptions

Assumption 2 (Smoothness) $q_z(\cdot, u)$, $z = 0, 1$, is a continuous function for all $u \in (0, 1)$. $G(\cdot, \cdot, \cdot)$ is a continuous function. $f_{\varepsilon|U_z,R}(e, u, r)$ for all $e \in \mathcal{E}$ and $u \in (0, 1)$ is continuous in $r \in \mathcal{R}$.

Assumption 3 (Local treatment rank invariance or similarity)

1. $U_0 = U_1$, conditional on $R = r_0$, or more generally
2. $U_0 | (\varepsilon, R = r_0) \sim U_1 | (\varepsilon, R = r_0)$.

- 3.1 requires that banks stay at the same rank of the capital distribution, regardless of whether the town size is 2,999 or 3,000.
- 3.2 weakens 3.1 (see, e.g., Chernozhukov and Hansen, 2005, 2006).
- They have testable implications.

Control Variable and Reduced-form Effects

Lemma

Let $U \equiv U_1 Z + U_0 (1 - Z)$. *Under Assumptions 1 - 3,*
1) $T \perp \varepsilon | (U, R)$, *and*

2) for any $u \in (0,1)$ *and integrable function of* Y, $\Gamma(Y)$,

$$\lim_{r \to r_0^+} \mathbb{E}[\Gamma(Y) | U = u, R = r] - \lim_{r \to r_0^-} \mathbb{E}[\Gamma(Y) | U = u, R = r]$$

$$= \int (\Gamma(G(q_1(r_0, u), r_0, e)) - \Gamma(G(q_0(r_0, u), r_0, e))) \, dF_{\varepsilon | U,R}(e, u, r_0).$$

- The first part says U is a control variable. Cf. Imbens and Newey (2009).
- The second part says that conditioning on the observed treatment rank U, any changes in outcome at the RD threshold are causal.

Parameters of Interest

Parameter of interest 1: LATE at a given treatment quantile *(Q-LATE)*:

- For any $u \in \mathcal{U}$,

$$\tau(u) \equiv \mathbb{E}\left[\underbrace{\frac{G(T_1, r, e) - G(T_0, r, e)}{T_1 - T_0}}_{\text{Individual causal effect}} \middle| U = u, R = r_0\right]$$

$$= \frac{\mathbb{E}\left[Y \middle| U_1 = u, R = r_0\right] - \mathbb{E}\left[Y \middle| U_0 = u, R = r_0\right]}{q_1(r_0, u) - q_0(r_0, u)}.$$

- Useful since treatments effect may depend on treatment intensity.

Parameters of Interest

Parameter of interest 2: weighted average of Q-LATEs (WQ-LATE):

$$\pi(w) \equiv \int_{\mathcal{U}} \tau(u)\, w(u)\, du,$$

where $w(u) \geq 0$ and $\int_{\mathcal{U}} w(u)\, du = 1$, i.e., $w(u)$ is a properly defined weighting function.

- Can think of Q-LATE as a conditional LATE conditional on $U = u$, while WQ-LATE as an unconditional LATE at r_0.

First-Stage Assumption

Identification of Q-LATE and WQ-LATE

Assumption 4 (First-stage): $q_1(r_0, u) \neq q_0(r_0, u)$ for at least some $u \in (0, 1)$.

- Assumption 4 requires that treatment distribution changes at $R = r_0$.
  - e.g., the capital distribution changes at Pop=3,000.
- The standard RD design requires mean changes $\mathbb{E}[T_1 | R = r_0] \neq \mathbb{E}[T_0 | R = r_0]$.

Define $q^{\pm}(u) \equiv \lim_{r \to r_0^{\pm}} q(r,u)$. Let $m(t,r) = \mathbb{E}\left[Y \middle| T = t, R = r\right]$, and then define $m^{\pm}(u) \equiv \lim_{r \to r_0^{\pm}} m(q^{\pm}(u), r)$.

Theorem (Identification)

Under Assumptions 1–4, for any $u \in \mathcal{U}$, $\tau(u)$ is given by

$$\tau(u) = \frac{m^{+}(u) - m^{-}(u)}{q^{+}(u) - q^{-}(u)}. \tag{1}$$

Further $\pi(w) \equiv \int_{\mathcal{U}} \tau(u)\, w(u)\, du$ is identified for any known or estimable weighting function $w(u)$ such that $w(u) \geq 0$ and $\int_{\mathcal{U}} w(u)\, du = 1$.

Possible Weighting Functions

1. Equal weighting: $w^S(u) \equiv 1/\int_{\mathcal{U}} 1 du$.
2. The standard RD estimand can be expressed as a WQ-LATE using quantile change weighting:

$$\pi^{RD} = \int_{\mathcal{U}} \tau(u)\, w^{RD}(u)\, du,$$

where $w^{RD}(u) \equiv \frac{\Delta q(u)}{\int_{\mathcal{U}} \Delta q(u) u}$, and $\Delta q(u) \equiv q_1(r_0, u) - q_0(r_0, u)$.

Standard RD estimand

- Validity of the standard RD estimand π^{RD} in general requires the following monotonicity assumption.

Assumption 3' (Monotonicity): $\Pr\left(T_1\left(r_0\right) \geq T_0\left(r_0\right)\right) = 1$ or $\Pr\left(T_1\left(r_0\right) \leq T_0\left(r_0\right)\right) = 1$.

- Assumption 3' implies $\Delta q\left(u\right) \geq 0$ or $\Delta q\left(u\right) \leq 0$ for any $u \in (0, 1)$ and hence $w^{RD}\left(u\right) \geq 0$.

Robust Identification for WQ-LATE

Theorem (Robust identification)

Let Assumptions 1, 2 and 4 hold. Then under either Assumption 3 or 3′,

$$\pi^* = \int_{\mathcal{U}} \frac{m^+(u) - m^-(u)}{q^+(u) - q^-(u)} \frac{|\Delta q(u)|}{\int_{\mathcal{U}} |\Delta q(u)| du} du \tag{2}$$

identifies a weighted average effect of Y on T at $R = r_0$.

- π^* is valid under either monotonicity or rank similarity; useful as neither assumption implies the other.
  - If monotonicity holds, $\pi^* = \pi^{RD}$; otherwise if rank restriction holds, $\pi^* = \int_{\mathcal{U}} \tau(u) w^*(u) du$ for $w^*(u) \equiv \frac{|\Delta q(u)|}{\int_{\mathcal{U}} |\Delta q(u)| du}$ is a WQ-LATE.
  - Either way, π^* identifies a weighted average effect among those who change treatment at r_0.

Empirical Analysis

How do banks respond to increased min capital requirements?

- $T = \log(\text{capital})$.
- $R =$ town population.
- $Y = \log(\text{assets}), \log(\text{leverage})$, suspension (up to 24 years).
- Covariates (X) for validity check = bank age, black population (%), farmland (%), log (manufacturing output per capita).

Bias-corrected Estimates of Q-LATEs

Table 3 Effects of log(capital) on bank outcomes

Q-LATE	Quantile	Log(assets)		Log(leverage)		Suspension	
	0.10	0.987	(0.289)***	-0.013	(0.273)	-0.019	(0.122)
	0.12	1.003	(0.275)***	0.003	(0.265)	-0.034	(0.117)
	0.14	1.017	(0.269)***	0.017	(0.258)	-0.036	(0.117)
	0.16	0.991	(0.298)***	-0.009	(0.245)	-0.038	(0.124)
	0.18	0.942	(0.351)***	-0.058	(0.267)	-0.091	(0.140)
	0.20	0.946	(0.349)***	-0.054	(0.297)	-0.092	(0.139)
	0.22	0.948	(0.350)***	-0.052	(0.295)	-0.093	(0.139)
	0.24	0.950	(0.341)***	-0.050	(0.296)	-0.108	(0.131)
	0.26	0.916	(0.323)***	-0.084	(0.283)	-0.112	(0.130)
WQ-LATE		0.968	(0.384)***	-0.032	(0.349)	-0.073	(0.136)

Note: The first panel presents the bias-corrected estimates of Q-LATEs at equally spaced quantiles; The last row presents the bias-corrected estimates of WQ-LATEs; $h_R = 1,155$ and $h_T = 0.435$ for all estimation, which are the AMSE optimal bandwidths for the WQ-LATE estimator; The AMSE optimal bandwidth for the Q-LATE estimator h_R ranges from $1,167.97$ to $1,570.61$. $h/b = 0.517$; The trimming thresholds are determined by using a preliminary bandwidth for R equal to $3/4h_R = 866.25$; Standard errors are in the parentheses; ***Significant at the 1% level, **Significant at the 5% level.

Bias-corrected Estimates of Q-LATEs

- A 1% increase in capital leads to an almost 1% increase in assets among small banks (those targeted by policy)!
- Leverage is not significantly lowered; long-run suspension rate is not affected.

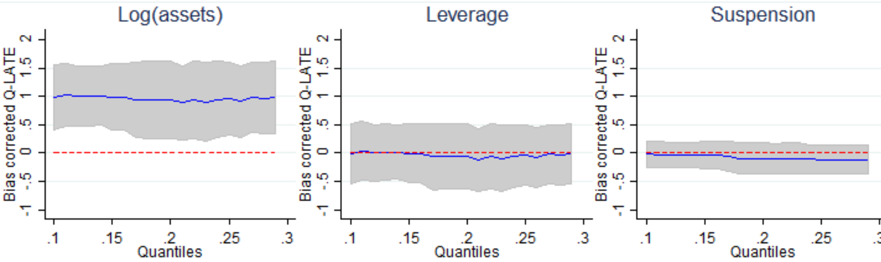


Figure: Estimated Q-LATEs at different quantiles with 95% CIs

Bias-corrected estimates of WQ-LATEs at different bandwidths

Results are robust to different bandwidths

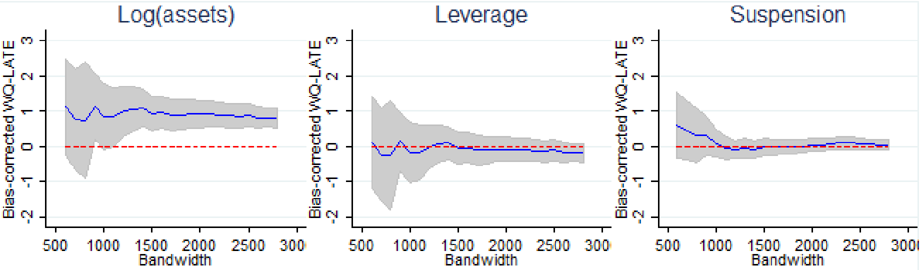


Figure: Estimated WQ-LATEs by different bandwidths

Conclusion

- Consider RD designs with a continuous treatment.
- Robust identification and inference utilizing any changes in the treatment distribution (mean changes are a special case).
    - Capital regulation focuses on small banks. The bottom of the capital distribution shifts up at the policy threshold.
- Identify what are likely to be the most policy relevant treatment effects by focusing on where the true treatment changes are.
    - Banks at the bottom of the capital distribution respond in ways that prevent the regulation from having intended effects!