Regression Discontinuity Designs with Sample Selection

Yingying Dong

University of California Irvine

Motivation: Recent Examples

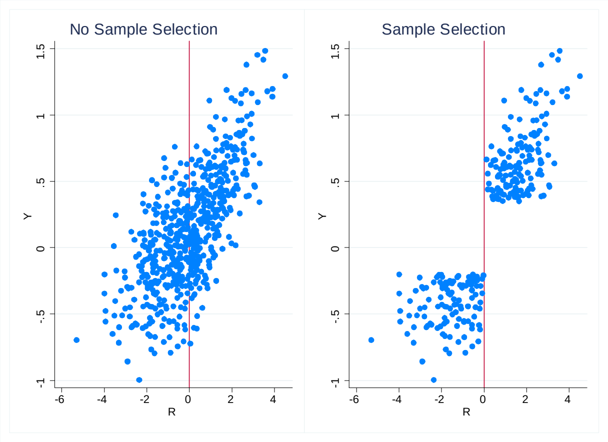
One frequently encountered issue in empirical applications of RD design – Sample selection!

Recent Examples:

- McCrary and Royer (2011) investigate effects of women's education on fertility and *infant health* (observed only if giving birth).
- Kim (2012) estimates effects of taking remedial courses on *performance in subsequent main courses* (observed only if completing subsequent courses).
- Other: Martorell and McFarlin (2011), Sampaio et al. (2013)...

Motivation: Illustration of the Issue

- Sample selection –> incomparability of observations above/ below the RD cutoff.
- Standard RD design (Hahn et al. 2001) is invalid.



What Does This Paper Do?

Extends the standard RD design to allow for sample selection or missing outcomes.

- Point identifies extensive and intensive margins of the RD treatment effect.
  - Extensive margin: effect on the participation probability.
  - Intensive margin: effect on the observed outcome distribution conditional on participation.
- Provides subgroup treatment effect bounds.
  - Also point identification of characteristics of subgroups (e.g. always participating vs. quitting compliers) - useful for policy.

What Does This Paper Do?

Identification here

- does not require exclusion restrictions in the selection equation.
  - Standard sample selection correction requires exclusion restrictions or functional form/distributional assumptions: Heckman (1979, 1990), Ahn and Powell (1993), Lee (1994), Andrews and Schafgans (1998), Das, Newey and Vella (2003) and Lewbel (2007) etc.
- does not require specifying the selection mechanism.
  - Sample selection could be caused by non-participation, dropout, survey nonresponse, or other reasons (e.g., censoring by death).

Applies these results to examine effects of academic probation on college persistence and GPA.

Notation

T =Treatment; $T \in \{0, 1\}$.

S =Sample selection indicator; $S \in \{0, 1\}$.

Y^* =(Partially observed) outcome.

Observe $Y = Y^*$ if $S = 1$; missing if $S = 0$.

Y_t^*, $t = 1, 0$ is potential outcome under treatment or no treatment.

S_t, $t = 1, 0$ is potential sample selection under treatment or no treatment.

Notation

- Let $T_1(r)$ and $T_0(r)$ be potential treatment status above or below the RD cutoff, respectively.

Let $T = h(R, V)$ for unobservables V, which can be a vector.
WLG, rewrite $T = h_1(R, V) Z + h_0(R, V)(1 - Z)$, where $Z = 1(R \geq r_0)$.
Then define $T_1(r) \equiv h_1(r, V)$ and $T_0(r) \equiv h_0(r, V)$.

- Individual types

Always takers (A): $T_1(r) = T_0(r) = 1$, Never takers (N): $T_1(r) = T_0(r) = 0$, Compliers (C): $T_1(r) > T_0(r)$, Defiers (D): $T_1(r) < T_0(r)$ (Angrist, Imbens, and Rubin, 1996).

Definitions

Two Margins of the RD treatment effect

- Extensive margin $\equiv E\left[S_1 - S_0 \middle| R = r_0, C\right]$.
  - e.g. change in the dropout rate with or without the probation policy.
- Intensive margin
  $\equiv E\left[Y_1^* \middle| S_1 = 1, R = r_0, C\right] - E\left[Y_0^* \middle| S_0 = 1, R = r_0, C\right]$.
  - e.g., how the quality (training) of college graduates differs with or without the probation policy, regardless of composition.
  - causal only from the distributional point of view.

Identifying Assumption

ASSUMPTION 1: The following assumptions hold for $r \in (r_0 - \varepsilon, r_0 + \varepsilon)$ for some small $\varepsilon > 0$.

A1. (Discontinuity): $\lim_{r \uparrow r_0} E[T|R = r] \neq \lim_{r \downarrow r_0} E[T|R = r]$.

A2. (Monotonicity): $\Pr(T_1 \geq T_0) = 1$.

A3. (Smoothness): $F_{Y_t^*, S_t|R,\Theta}(y, s|r)$ for $s, t \in \{0, 1\}$ is continuous at r_0. The probability of each type $\Pr[\Theta|R = r]$ for $\Theta \in \{A, N, C\}$, is continuous at r_0. The density of R is continuous and strictly positive at r_0.

Point Identification of Extensive vs. Intensive Margin

THEOREM 1 Let $I_t \equiv 1(T = t)$. Given Assumption 1, for $t = 0, 1$

$$E[S_t | R = r_0, C] = \frac{\lim_{r \downarrow r_0} E[SI_t | R = r] - \lim_{r \uparrow r_0} E[SI_t | R = r]}{\lim_{r \downarrow r_0} E[I_t | R = r] - \lim_{r \uparrow r_0} E[I_t | R = r]}, \quad (1)$$

$$E[g(Y_t^*) | S_t = 1, R = r_0, C] \quad (2)$$
$$= \frac{\lim_{r \downarrow r_0} E[g(Y^*) SI_t | R = r] - \lim_{r \uparrow r_0} E[g(Y^*) SI_t | R = r]}{\lim_{r \downarrow r_0} E[SI_t | R = r] - \lim_{r \uparrow r_0} E[SI_t | R = r]}.$$

- $p_t \equiv E[S_t | R = r_0, C]$; *Extensive margin* $= p_1 - p_0$.
- Setting $g(Y^*) = 1(Y^* \leq y)$ identifies $F_{Y_t^* | S_t = 1, R = r_0, C}(y)$; Setting $g(Y^*) = Y^*$ identifies $\mathbb{E}_t \equiv E[Y_t^* | S_t = 1, R = r_0, C]$ and *Intensive margin*$= \mathbb{E}_1 - \mathbb{E}_0$.
- See, e.g., Abadie (2003) and Frandsen et al. (2012).

General Bounds without Additional Assumptions

Composition of $S_1 = 1$ is not the same as that of $S_0 = 1$. Can derive bounds on always participating ($S_0 = S_1 = 1$) compliers.

- Let $p_{11} \equiv E\left[S_1 = 1, S_0 = 1 \middle| R = r_0, C\right]$
    - $p_{11} \in \mathcal{P} \equiv (0,1] \cap \left[p_0 + p_1 - 1, \min\left\{p_0, p_1\right\}\right]$;
- Let $F_t\left(y\right) \equiv F_{Y_t^* | S_t = 1, R = r_0, C}\left(y\right)$ and $Q_t\left(\tau\right) \equiv F_t^{-1}(\tau)$ for $t = 0, 1$ and $\tau \in (0, 1)$.

By Horowitz and Manski (1995), the worst-case (best-case) scenario bounds are

$$\Delta^{LB} = \min_{p_{11} \in \mathcal{P}} \left(\frac{p_1}{p_{11}} \int_{-\infty}^{Q_1(p_{11}/p_1)} y dF_1\left(y\right) - \frac{p_0}{p_{11}} \int_{Q_0(1-p_{11}/p_0)}^{+\infty} y dF_0\left(y\right) \right),$$

$$\Delta^{UB} = \max_{p_{11} \in \mathcal{P}} \left(\frac{p_1}{p_{11}} \int_{Q_1(1-p_{11}/p_1)}^{+\infty} y dF_1\left(y\right) - \frac{p_0}{p_{11}} \int_{-\infty}^{Q_0(p_{11}/p_0)} y dF_0\left(y\right) \right).$$

Bounds under Monotonic Selection

ASSUMPTION 2 (Monotonic Selection): $\Pr(S_0 \geq S_1) = 1$.

- Treatment can only affect sample selection in "one direction" (Lee, 2009, Kim 2012, Blanco et al, 2013, Chen and Flores, 2014).
- Consistent with a latent index sample selection model with an additively separable latent error (Heckman 1979, 1990 and Vytlacil 2002).

Bounds under Monotonic Selection

- Given monotonic selection, $S_1 = 1$ consists of only always participants $(S_1 = 1, S_0 = 1)$, while $S_0 = 1$ consists of always participants and quitters $(S_1 = 0, S_0 = 1)$.

- Can then identify the fraction of quitters among $S_0 = 1$:

$$
\begin{aligned}
q &\equiv \frac{\Pr\left(S_1 = 0, S_0 = 1 | R = r_0, C\right)}{\Pr\left(S_0 = 1 | R = r_0, C\right)} \\
&= \frac{\lim_{r \downarrow r_0} E\left[S | R = r\right] - \lim_{r \uparrow r_0} E\left[S | R = r\right]}{\lim_{r \downarrow r_0} E\left[S\left(1 - T\right) | R = r\right] - \lim_{r \uparrow r_0} E\left[S\left(1 - T\right) | R = r\right]}.
\end{aligned}
$$

- "Trimming" the lower or upper q fraction of observations in the distribution of $Y_0^* | S_0 = 1, R = r_0, C$ yields the worst-case scenario bounds (Horowitz and Manski, 1995).

Bounds under Monotonic Selection

THEOREM 2 Under Assumptions 1 and 2, the upper and lower bounds of $E\left[Y_1^* - Y_0^* | S_1 = 1, S_0 = 1, R = r_0, C\right]$ are given by

$$\begin{aligned}
\Delta_m^{UB} &= \mathbb{E}_1 - \frac{1}{1-q}\int_{-\infty}^{Q_0(1-q)} y dF_{Y_0^*|S_0=1,R=r_0,C}(y) \\
&= \mathbb{E}_1 - \frac{E\left[1\left(Y_0^* \le Q_0(1-q)\right) Y_0^* | S_0 = 1, R = r_0, C\right]}{1-q}, \text{ and}
\end{aligned}$$

$$\begin{aligned}
\Delta_m^{LB} &= \mathbb{E}_1 - \frac{1}{1-q}\int_{Q_0(q)}^{+\infty} y dF_{Y_0^*|S_0=1,R=r_0,C}(y) \\
&= \mathbb{E}_1 - \frac{E\left[1\left(Y_0^* \ge Q_0(q)\right) Y_0^* | S_0 = 1, R = r_0, C\right]}{1-q},
\end{aligned}$$

respectively, where $\mathbb{E}_1 \equiv E\left[Y_1^* | S_1 = 1, R = r_0, C\right]$, the quantile $Q_0(\tau) \equiv \inf\{y: F_{Y_0^*|S_0=1,R=r_0,C}(y) \ge \tau\}$ for $\tau = 1-q,\ q$.

Bounds on QTE under Monotonic Selection

- Can also construct bounds on the quantile treatment effect (QTE).
- Define $QTE(\tau) \equiv F^{-1}_{Y_1^*|S_0=1,S_1=1,R=r_0,C}(\tau) - F^{-1}_{Y_0^*|S_0=1,S_1=1,R=r_0,C}(\tau)$ for $\tau \in (0,1)$.
- Suppose Assumptions 1 and 2 hold. The upper and lower bounds of $QTE(\tau)$ are

$$\begin{aligned} QTE^{UB}(\tau) &= Q_1(\tau) - Q_0(\tau(1-q)) \quad \text{and} \\ QTE^{LB}(\tau) &= Q_1(\tau) - Q_0(1-(1-\tau)(1-q)). \end{aligned}$$

Identify Subgroup Characteristics

- Can identify characteristics of always participating or quitting compliers, given monotonic selection.
- Let X be some pre-determined covariate.
  - 1. Replacing Y^* with X in Theorem 1 identifies $F_{X|S_1=1,R=r_0,C}(x)$, $F_{X|S_0=1,R=r_0,C}(x)$.
  - 2. $F_{X|S_1=1,R=r_0,C}(x)$ is for always participants: $F_{X|S_1=1,R=r_0,C}(x) = F_{X|S_1=1,S_0=1,R=r_0,C}(x)$.
  - 3. $F_{X|S_0=1,R=r_0,C}(x)$ is for $1-q$ fraction of always participants and q fraction of quitters.

Testable Implications of Monotonic Selection / Identify Subgroup Characteristics

COROLLARY 2 Assume that A1, A2 hold. Assume further that A3 holds after replacing Y_t^* with X. Under Assumption 2, we have

$$F_{X|S_1=1,S_0=1,R=r_0,C}(x) = F_{X|S_1=1,R=r_0,C}(x),$$
$$F_{X|S_1=0,S_0=1,R=r_0,C}(x) = \frac{1}{q}F_{X|S_0=1,R=r_0,C}(x) - \frac{1-q}{q}F_{X|S_1=1,R=r_0,C}(x).$$

- Monotonic sample selection implies

$$1 \geq \frac{1}{q}F_{X|S_0=1,R=r_0,C}(x) - \frac{1-q}{q}F_{X|S_1=1,R=r_0,C}(x) \geq 0 \text{ for all } x \in \mathcal{X}.$$

- Can *verify monotonic selection* by a one-sided t test for the above.

Bounds under Stochastic Dominance

ASSUMPTION 3 (Stochastic Dominance):
$F_{Y_1^*|S_0=1,S_1=1,R=r_0,C}(y) \leq F_{Y_1^*|S_0=0,S_1=1,R=r_0,C}(y)$ and
$F_{Y_0^*|S_0=1,S_1=1,R=r_0,C}(y) \leq F_{Y_0^*|S_0=1,S_1=0,R=r_0,C}(y)$ for all $y \in \mathcal{Y}$.

The distribution of Y_1^* (Y_0^*) for the always participating compliers weakly stochastically dominates that of the quitting (newly participating) compliers.

Bounds under Stochastic Dominance

THEOREM 3 Assume that $p_0 + p_1 > 1$. Given Assumptions 1 and 3, the upper and lower bounds of $E\left[Y_1^* - Y_0^* | S_1 = 1, S_0 = 1, R = r_0, C\right]$ are given by

$$\begin{aligned}
\Delta_s^{UB} &= E\left[Y_1^* | S_1 = 1, Y_1^* \geq Q_1\left(\frac{1-p_0}{p_1}\right), R = r_0, C\right] - \mathbb{E}_0 \\
&= \frac{p_1 E\left[1\left(Y_1^* \geq Q_1\left(\frac{1-p_0}{p_1}\right)\right) Y_1^* | S_1 = 1, R = r_0, C\right]}{p_0 + p_1 - 1} - \mathbb{E}_0
\end{aligned}$$

$$\begin{aligned}
\Delta_s^{LB} &= \mathbb{E}_1 - E\left[Y_0^* | S_0 = 1, Y_0^* \geq Q_0\left(\frac{1-p_1}{p_0}\right), R = r_0, C\right] \\
&= \mathbb{E}_1 - \frac{p_0 E\left[1\left(Y_0^* \geq Q_0\left(\frac{1-p_1}{p_0}\right)\right) Y_0^* | S_0 = 1, R = r_0, C\right]}{p_0 + p_1 - 1}.
\end{aligned}$$

where $\mathbb{E}_t \equiv E\left[Y_t^* | S_t = 1, R = r_0, C\right]$, $t = 0, 1$.

Bounds on QTE under Stochastic Dominance

Suppose Assumptions 1 and 3 hold. The upper and lower bounds of $QTE(\tau)$ are

$$
\begin{aligned}
QTE_s^{UB}(\tau) &= Q_1\left(1 - \frac{(1-\tau)(p_0 + p_1 - 1)}{p_1}\right) - Q_0(\tau) \quad \text{and} \\
QTE_s^{LB}(\tau) &= Q_1(\tau) - Q_0\left(1 - \frac{(1-\tau)(p_0 + p_1 - 1)}{p_0}\right).
\end{aligned}
$$

Combining Monotonic Selection and Stochastic Dominance

Note that Assumptions 2 and 3 may be changed and combined, depending on their plausibility in a particular empirical application.

E.g., Can tighten the bounds using both assumptions:

$$\begin{aligned}
\Delta_{ms}^{LB} &= \Delta_{m}^{LB} \\
&= E\left[Y_1^* | S_1 = 1, R = r_0, C\right] - \frac{1}{1-q} \int_{Q_0(q)}^{+\infty} y dF_{Y_0^* | S_0 = 1, R = r_0, C}(y), \\
\Delta_{ms}^{UB} &= E\left[Y_1^* | S_1 = 1, R = r_0, C\right] - E\left[Y_0^* | S_0 = 1, R = r_0, C\right]. \\
&= \mathbb{E}_1 - \mathbb{E}_0.
\end{aligned}$$

The intensive margin effect is the upper bound!

Academic Probation and Gender Differences in Responses

Nearly all colleges and universities in the US have the academic probation policy.

Surprisingly little empirical evidence.

$Y^* =$ cumulative GPA.

$S = 1$ if one does not drop out and 0 otherwise.

$T =$ ever being placed on academic probation.

$R =$ first semester GPA; On probation if GPA falls below $r_0 = 2.0$.

Empirical Application: Data Description

Confidential data from Texas Higher Education Opportunity Project (THEOP).

- Both admission records and transcript data are available.
- Sample is representative of the entire population of the first-time freshmen cohorts between 1992 and 2002 from a large public university in Texas.
- Sample consists of 64,310 students for whom there are complete records.

Empirical Application: Sample Summary Statistics

Note that the sample size is much smaller for final GPA, indicating serious sample selection or attrition!

Table 1 Sample Descriptive Statistics

	Ever on probation		Never on probation		
	N	Mean (SD)	N	Mean (SD)	Difference
	II: 1st semester GPA=2.0±0.5				
Final GPA	4,607	2.565	7,901	2.808	-0.243
		(0.324)		(0.323)	(0.006)***
College completion	8,512	0.541	9,351	0.845	-0.304
		(0.498)		(0.362)	(0.006)***
Male	8,512	0.565	9,357	0.465	0.100
		(0.496)		(0.499)	(0.007)***
White	8,512	0.746	9,351	0.806	-0.059
		(0.435)		(0.396)	(0.006)***
SAT score	8,497	1,111	9,336	1,124	-12.43
		(127.2)		(120.4)	(1.855)***
Top 25% of HS class	8,512	0.706	9,351	0.778	-0.073
		(0.456)		(0.415)	(0.007)***
HS NHS member	8,512	0.265	9,351	0.273	-0.008
		(0.442)		(0.446)	(0.007)

First-stage Figures

Ever placement on probation

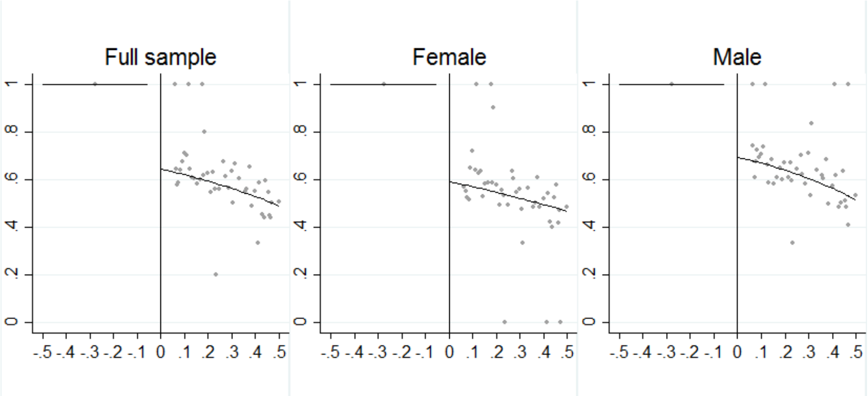
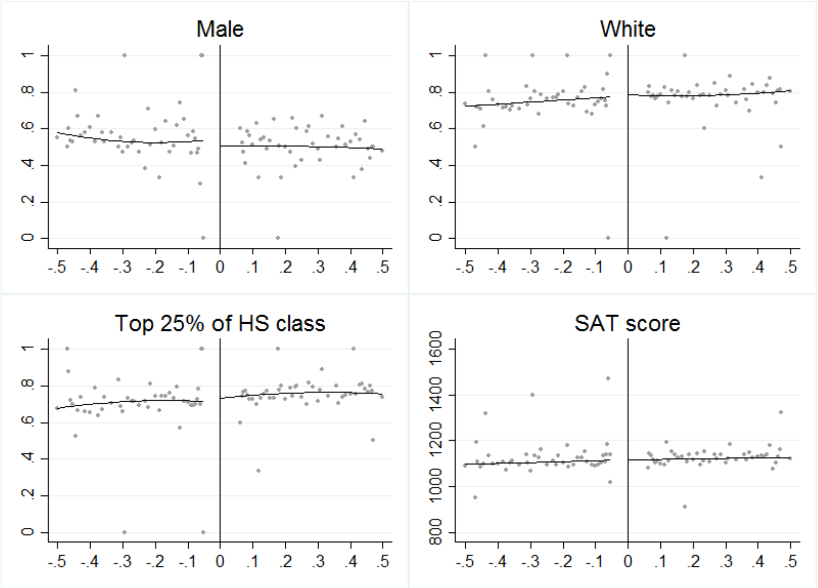


Figure: Probation and the first semester GPA (centered at 2.0)

Validity of the RD Design



Validity of the RD Design

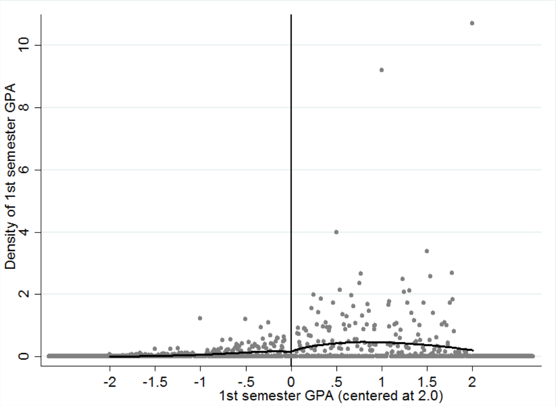


Figure: Density of the running variable (first semester GPA)

RD Validity Tests

Table 2 RD Validity Tests

I: RD effects of Academic Probation on Covariates			
Male	0.032 (0.045)	Top 25% of HS Class	-0.040 (0.036)
White	0.005 (0.038)	HS NHS member	-0.006 (0.033)
SAT score	0.158 (12.87)	Feeder school	0.025 (0.025)
II: Discontinuity in the Density of Running Variable			
	0.115 (0.600)		0.047 (0.041)

Note: In Panel I, the CCT bias-corrected estimates along with robust standard errors are reported. In Panel II, the first column reports the estimated discontinuity in logarithm of the empirical density of the running variable (with a bin width 0.01); the second column reports the estimated discontinuity by the nonparametric density estimator of Cattaneo, Jansson, and Ma (2016).

College Completion and Final GPA

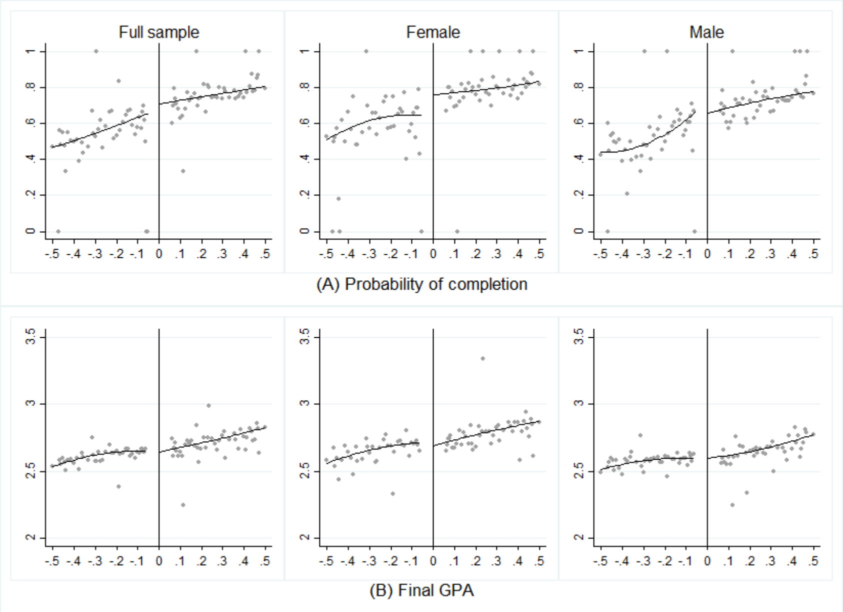
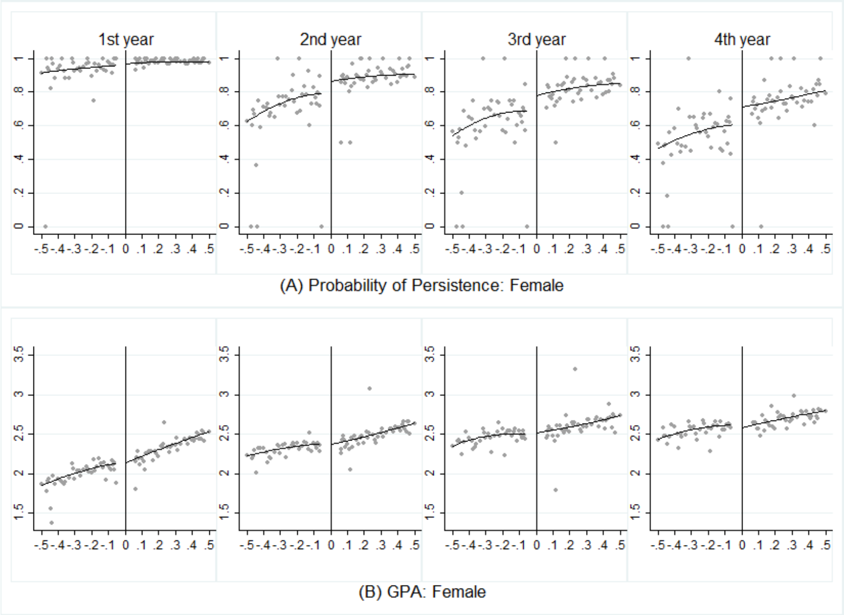


Table 3 Effects of Academic Probation on College Completion and Final GPAs

	Full sample		Female		Male	
	I: RDD with sample selection					
1):Pr($S_0 = 1$)	0.824	(0.018)***	0.834	(0.023)***	0.820	(0.025)***
2):Pr($S_1 = 1$)	0.773	(0.039)***	0.659	(0.064)***	0.875	(0.080)***
Extensive margin: 2)-1)	-0.051	(0.037)	-0.182	(0.068)***	0.057	(0.090)
3): E($Y_0 \mid S_0 = 1$)	2.727	(0.016)***	2.768	(0.020)***	2.686	(0.022)***
4): E($Y_1 \mid S_1 = 1$)	2.771	(0.026)***	2.837	(0.039)***	2.716	(0.036)***
Intensive margin: 4)-3)	0.045	(0.036)	0.069	(0.050)	0.030	(0.050)
	II: Standard RDD					
	0.029	(0.032)	0.049	(0.040)	0.047	(0.058)
	III: Bounds for always participating compliers					
Lower bound 2	0.045	(0.036)	0.069	(0.050)	0.030	(0.052)
Upper bound 2	0.209	(0.098)**	0.148	(0.121)	0.030	(0.102)
90% CI 2	[-0.002	0.336]	[-0.002	0.318]	[-0.055	0.198]
N	64,310		32,952		31,358	

Note: All estimates are conditional on compliers at the 1st semester GPA equal to 2.0; Estimation of the extensive and intensive margins, and the bounds follows the description in Sections 2 and 3.1, respectively. The CCT bias-corrected robust inference is used; In Panel III, 1 refers to the bounds under the monotonic sample selection assumption, while 2 refers to the bounds assuming additionally mean dominance, particularly $E\left(Y_0 \mid S_0 = 1, S_1 = 0, C, R = r_0\right) \geq E\left(Y_0 \mid S_0 = 1, S_1 = 1, C, R = r_0\right)$; Bootstrapped standard errors are in the parentheses; Imbens and Manski's (2004) CIs are reported; *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

Female College Persistence and GPA



Male College Persistence and GPA

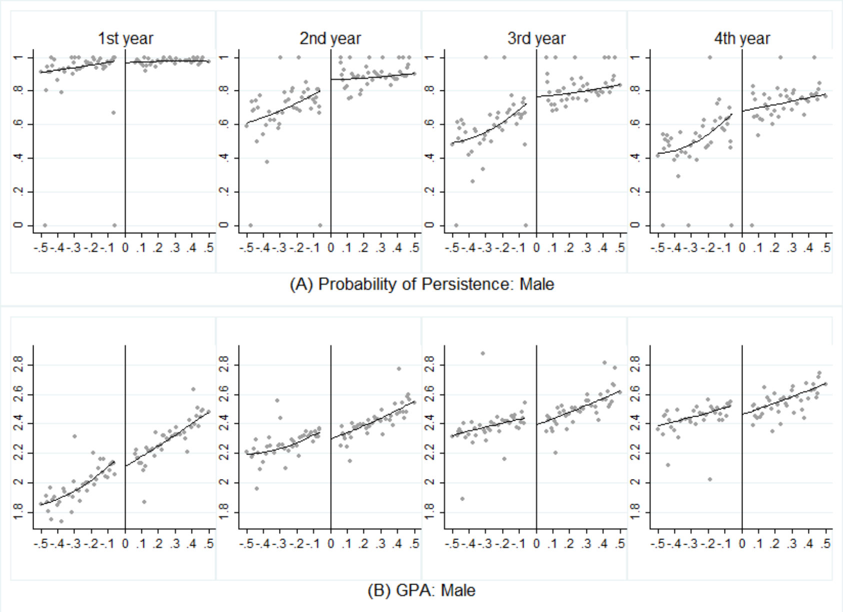


Table 4 Effects on College Persistence and GPAs (Women)

	1st year	2nd year	3rd year	4th year
	I: RDD with sample selection			
(1):Pr($S_0 = 1$)	0.974	0.898	0.848	0.801
	(0.005)***	(0.019)***	(0.023)***	(0.029)***
(2):Pr($S_1 = 1$)	0.956	0.779	0.686	0.641
	(0.015)***	(0.049)***	(0.065)***	(0.070)***
Extensive margin: (2)-(1)	-0.011	-0.117	-0.162	-0.167
	(0.017)	(0.052)**	(0.071)**	(0.073)**
(3): E($Y_0 \| S_0 = 1$)	2.149	2.480	2.608	2.683
	(0.011)***	(0.017)***	(0.018)***	(0.023)***
(4): E($Y_1 \| S_1 = 1$)	2.125	2.529	2.636	2.775
	(0.027)***	(0.035)***	(0.040)***	(0.056)***
Intensive margin: (4)-(3)	-0.024	0.049	0.029	0.092
	(0.042)	(0.050)	(0.070)	(0.073)
	II: Standard RDD			
	0.033	0.039	0.012	0.106
	(0.023)	(0.039)	(0.042)	(0.042)**
	III: Bounds for always participating compliers			
lower bound 2	-0.024	0.049	0.029	0.092
	(0.042)	(0.050)	(0.070)	(0.073)
Upper bound 2	0.080	0.066	0.229	0.108
	(0.069)	(0.147)	(0.183)	(0.104)
90% CI 2	[-0.079 0.169]	[-0.031 0.300]	[-0.063 0.421]	[-0.008 0.272]
N	51,374	51,115	48,128	40,921

Note: All estimates are conditional on compliers at the 1st semester GPA equal to 2.0; Esti-

Table 5 Effects on College Persistence and GPAs (Men)

	1st year	2nd year	3rd year	4th year
	I: RDD with sample selection			
(1):$\Pr(S_0 = 1)$	0.973	0.925	0.872	0.809
	(0.005)***	(0.017)***	(0.022)***	(0.027)***
(2):$\Pr(S_1 = 1)$	0.975	0.856	0.872	0.847
	(0.015)***	(0.047)***	(0.056)***	(0.066)***
Extensive margin: (2)-(1)	0.000	-0.063	-0.008	0.036
	(0.017)	(0.050)	(0.065)	(0.078)
(3): $E(Y_0 \mid S_0 = 1)$	2.108	2.416	2.517	2.597
	(0.014)***	(0.018)***	(0.021)***	(0.026)***
(4): $E(Y_1 \mid S_1 = 1)$	2.192	2.523	2.611	2.710
	(0.024)***	(0.039)***	(0.038)***	(0.043)***
Intensive margin: (4)-(3)	0.084	0.107	0.094	0.113
	(0.027)***	(0.054)**	(0.054)*	(0.072)
	II: Standard RDD			
	0.078	0.103	0.098	0.142
	(0.025)***	(0.046)**	(0.049)**	(0.064)**
	III: Bounds for always participating compliers			
Lower bound 2	0.084	0.107	0.094	0.113
	(0.027)	(0.054)	(0.054)	(0.072)
Upper bound 2	0.084	0.188	0.094	0.113
	(0.088)	(0.138)	(0.142)	(0.194)
90% CI 2	[0.039 0.230]	[0.030 0.384]	[0.006 0.327]	[-0.006 0.429]
N	51,374	51,115	48,128	40,921

Note: All estimates are conditional on compliers at the 1st semester GPA equal to 2.0; Esti-

Who drop out?

Lower ability women drop out? No

Table 6 Mean Characteristics of Subgroups of Compliers

	Always participants	Quitters
White	0.781 (0.067)***	0.893 (0.374)**
SAT score	1,093 (14.19)***	1,112 (106.4)***
Top 25% of HS class	0.774 (0.068)***	0.948 (0.460)**
HS NHS	0.268 (0.059)***	0.346 (0.450)
Feeder school	0.172 (0.055)***	0.005 (0.419)

Note: Estimates are based on the sample of women; NHS means National Honors Society member; The CCT bias-corrected estimates are reported; Bootstrapped standard errors are in the parentheses; *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

Estimated quitter characteristics

Estimated quitter characteristics suggest

1. Monotonic sample selection is plausible.
2. Quitters would perform no worse than always participating compliers.

Assume the above, bounds on the always participating compliers (bounds 2 in the tables) are

$$\Delta_{ms'}^{UB} = \mathbb{E}_1 - \frac{1}{1-q}\int_{-\infty}^{Q_0(1-q)} y dF_{Y_0^*|S_0=1,R=r_0,C}(y)$$

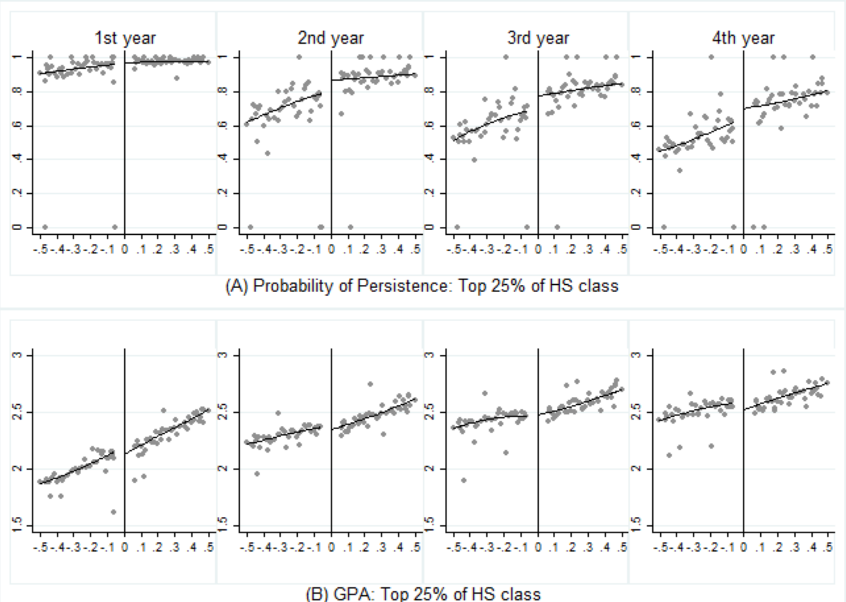
$$\Delta_{ms'}^{LB} = \mathbb{E}_1 - \mathbb{E}_0.$$

Academic Probation and College Performance: Findings

Striking gender differences in responses:

- Women are significantly more likely to drop out once placed on probation.
    - No evidence on low-ability women dropping out (evidence consistent with a discouragement effect).
- Little impacts on men's dropout rate.
- Men cope with this negative signal by temporarily improving their performance to avoid being suspended.

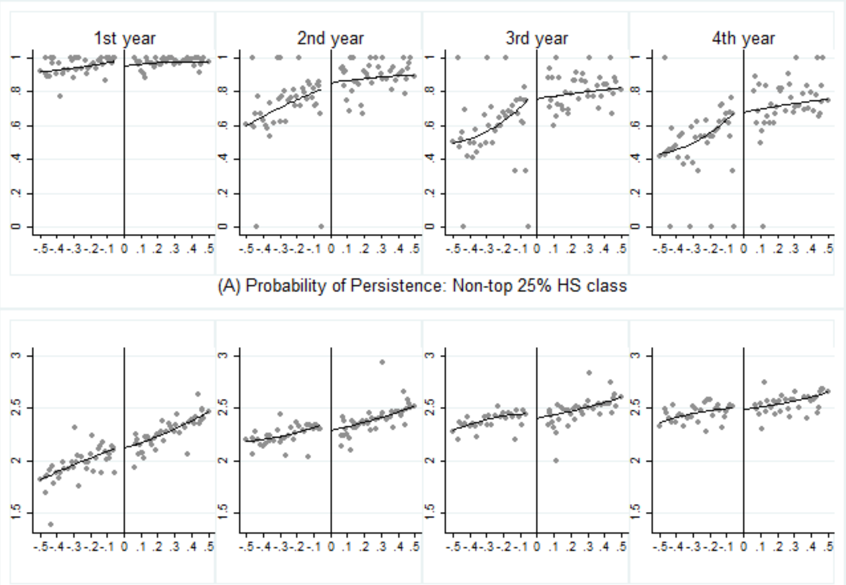
College Persistence and GPA: Top 25% of HS class



(A) Probability of Persistence: Top 25% of HS class

(B) GPA: Top 25% of HS class

College Persistence and GPA: Non-top 25% of HS class



(A) Probability of Persistence: Non-top 25% HS class

(B) GPA: Non-top 25% of HS class

Table 1A Effects on College Persistence and GPAs (Top 25% of HS Class)

	1st year	2nd year	3rd year	4th year
	I: RDD with sample selection			
(1):Pr($S_0 = 1$)	0.975	0.914	0.863	0.808
	(0.003)***	(0.015)***	(0.018)***	(0.023)***
(2):Pr($S_1 = 1$)	0.976	0.774	0.696	0.700
	(0.014)***	(0.042)***	(0.055)***	(0.054)***
Extensive margin: (2)-(1)	-0.019	-0.116	-0.157	-0.087
	(0.015)	(0.041)***	(0.057)***	(0.049)*
(3): E($Y_0 \| S_0 = 1$)	2.134	2.455	2.576	2.642
	(0.009)***	(0.014)***	(0.016)***	(0.020)***
(4): E($Y_1 \| S_1 = 1$)	2.178	2.526	2.627	2.762
	(0.018)***	(0.028)***	(0.030)***	(0.037)***
Intensive margin: (4)-(3)	0.044	0.071	0.050	0.120
	(0.027)*	(0.035)**	(0.041)	(0.050)**
	II: Standard RDD			
	0.033	0.039	0.012	0.106
	(0.027)	(0.039)	(0.042)	(0.042)**
	III: Bounds for always participating compliers			
Lower bound 2	0.044	0.071	0.050	0.120
	(0.027)*	(0.035)**	(0.041)	(0.050)**
Upper bound 2	0.119	0.279	0.232	0.308
	(0.055)	(0.128)	(0.126)	(0.107)
90% CI 2	[0.009 0.190]	[0.026 0.444]	[-0.003 0.395]	[0.056 0.445]
N	51,374	51,115	48,128	40,921

Note: All estimates are conditional on compliers at the 1st semester GPA equal to 2.0; Esti-

Table 2A Effects on College Persistence and GPAs (Non-top 25% of HS Class)

	1st year	2nd year	3rd year	4th year
	I: RDD with sample selection			
(1):$\Pr(S_0 = 1)$	0.960	0.896	0.840	0.787
	(0.010)***	(0.027)***	(0.034)***	(0.046)***
(2):$\Pr(S_1 = 1)$	0.993	0.872	0.895	0.865
	(0.018)***	(0.064)***	(0.079)***	(0.109)***
Extensive margin: (2)-(1)	0.035	-0.024	0.047	0.070
	(0.022)	(0.068)	(0.088)	(0.113)
(3): $E(Y_0 \mid S_0 = 1)$	2.102	2.428	2.530	2.646
	(0.018)***	(0.027)***	(0.032)***	(0.044)***
(4): $E(Y_1 \mid S_1 = 1)$	2.138	2.547	2.625	2.695
	(0.032)***	(0.052)***	(0.060)***	(0.080)***
Intensive margin: (4)-(3)	0.036	0.119	0.094	0.049
	(0.042)	(0.070)*	(0.071)	(0.095)
	II: Standard RDD			
	0.043	0.118	0.102	0.077
	(0.038)	(0.057)**	(0.073)	(0.078)
	III: Bounds for always participating compliers			
Lower bound 2	0.036	0.119	0.094	0.049
	(0.042)	(0.070)*	(0.071)	(0.095)
Upper bound 2	0.036	0.300	0.094	0.049
	(0.082)	(0.199)	(0.159)	(0.193)
90% CI 2	[-0.033 0.171]	[0.024 0.569]	[-0.022 0.355]	[-0.108 0.366]
N	12,868	12,795	11,947	10,607

Note: All estimates are conditional on compliers at the 1st semester GPA equal to 2.0; Esti-

Conclusion

- This paper extends the standard RD design to deal with sample selection.
- Identification relies on smoothness conditions; empirically plausible and readily testable.
- Main results:
  - Point identification of the extensive and intensive margins of the RD treatment effect.
  - Sharp bounds for the treatment effect among the always participating compliers.
  - Point identification of characteristics of each subgroup of compliers (e.g., always participating compliers vs. quitters/new participants)
- Applies these results and shows striking gender difference in responses to academic probation along the extensive vs. intensive margin.

Extensions

Possible extension/future work

- S as a mechanism variable:
  - T can affect Y through S, so can extend the current framework to separate 1) the mechanism effect of S on Y holding T fixed from 2) the net effect of T on Y, holding S fixed.
  - e.g., T =college educ., Y =earnings, S =blue- vs. white-collar occupation

- Current Theorem 1 can be extended directly to the case where S is categorial;
  - So to identify $E\left[1\left(S_t = k\right)|R = r_0, C\right]$ and $E\left[g\left(Y_t^*\right)|S_t = k, R = r_0, C\right]$, for $k = 1, ..., K$, where $1\left(S_t = k\right)$ indicates selection into category k.