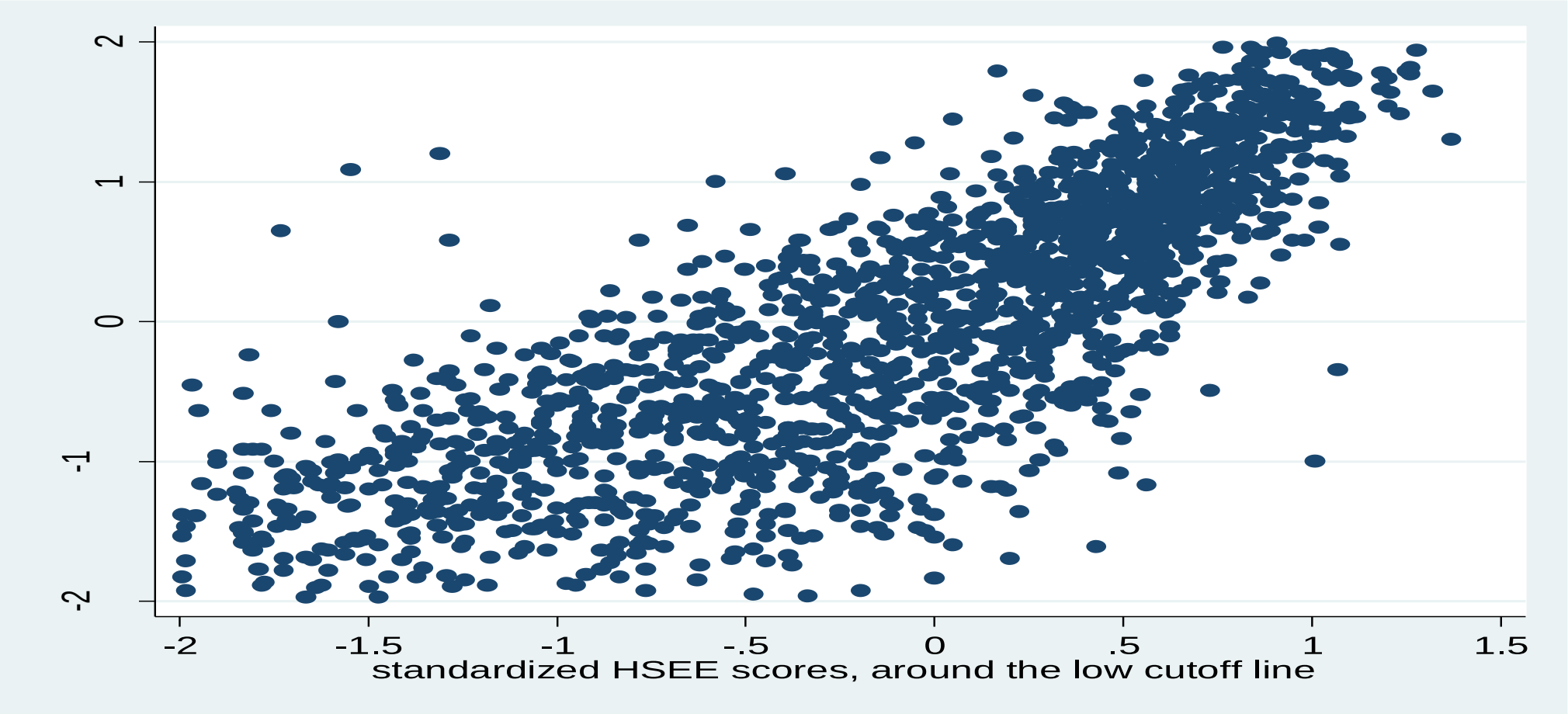
Lecture A2: Local Regression and Regression Discontinuity

Fan Regression (I):Local Polynomial Regression

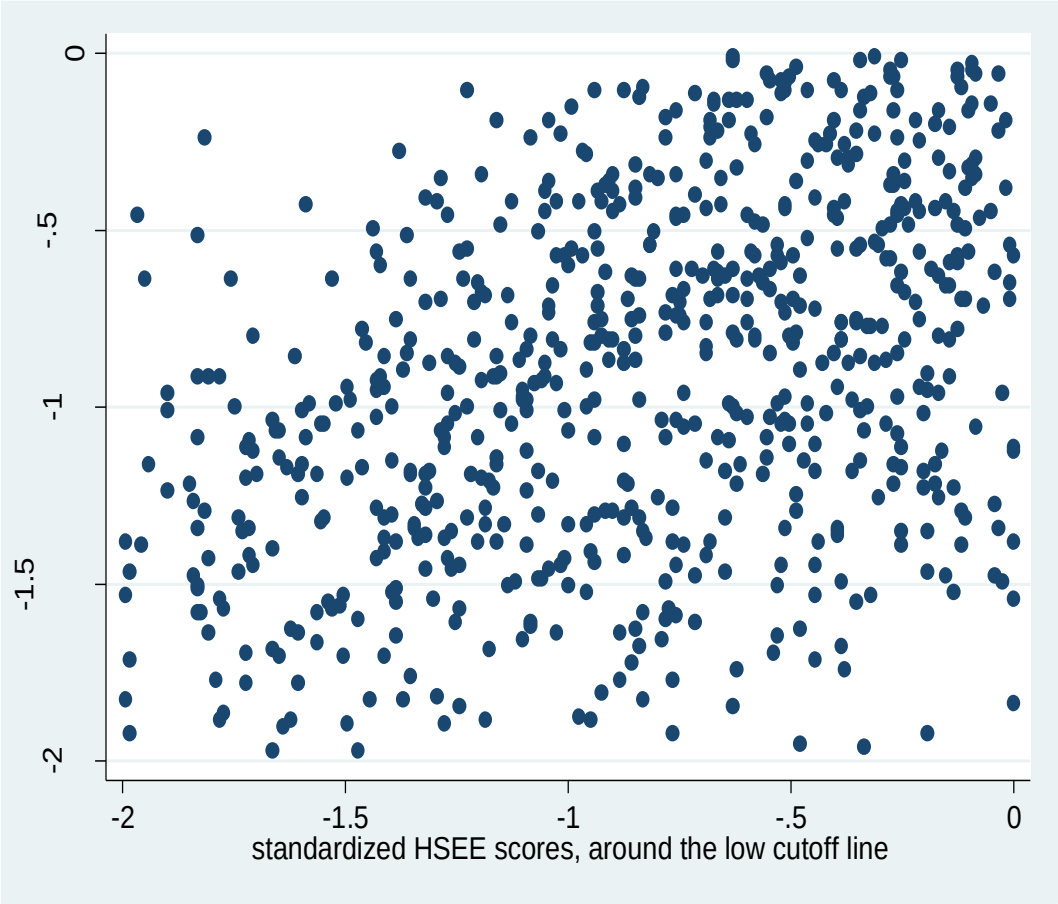
$$Y_i = m(X_i) + \epsilon_i$$

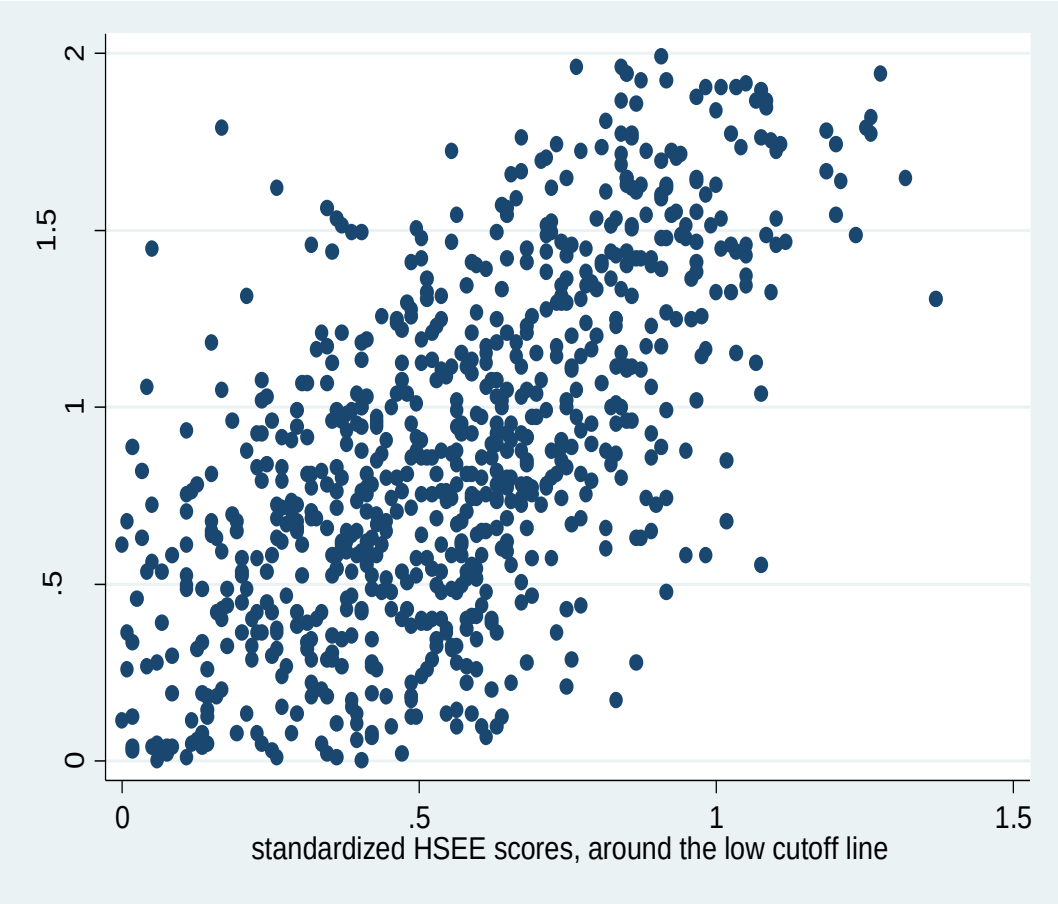
$$m(x) \approx m(x_0) + m^{(1)}(x_0)(x - x_0) + \frac{m^{(2)}(x_0)}{2!}(x - x_0)^2 + \cdots + \frac{m^{(p)}(x_0)}{p!}(x - x_0)^p$$

高考成绩与中考成绩

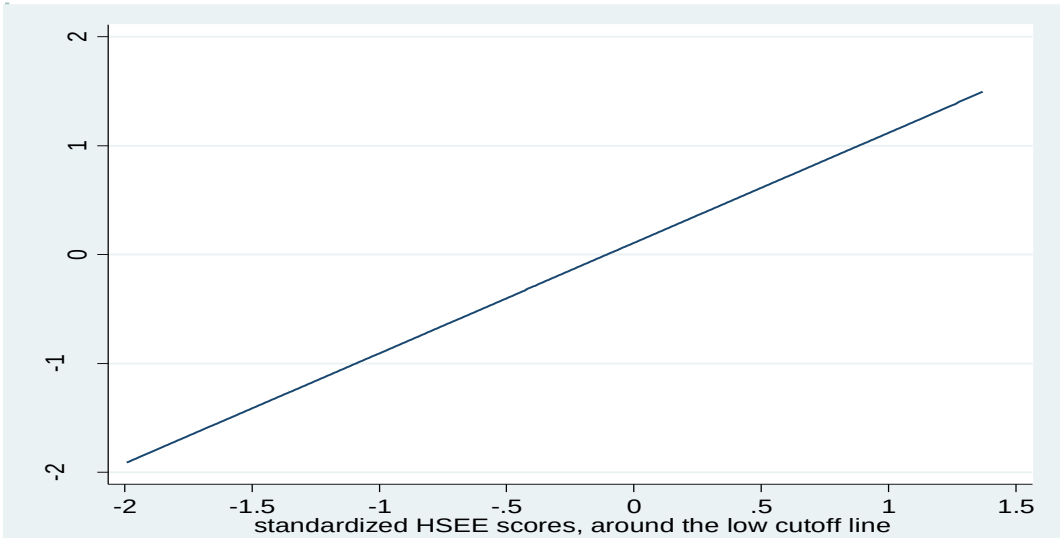


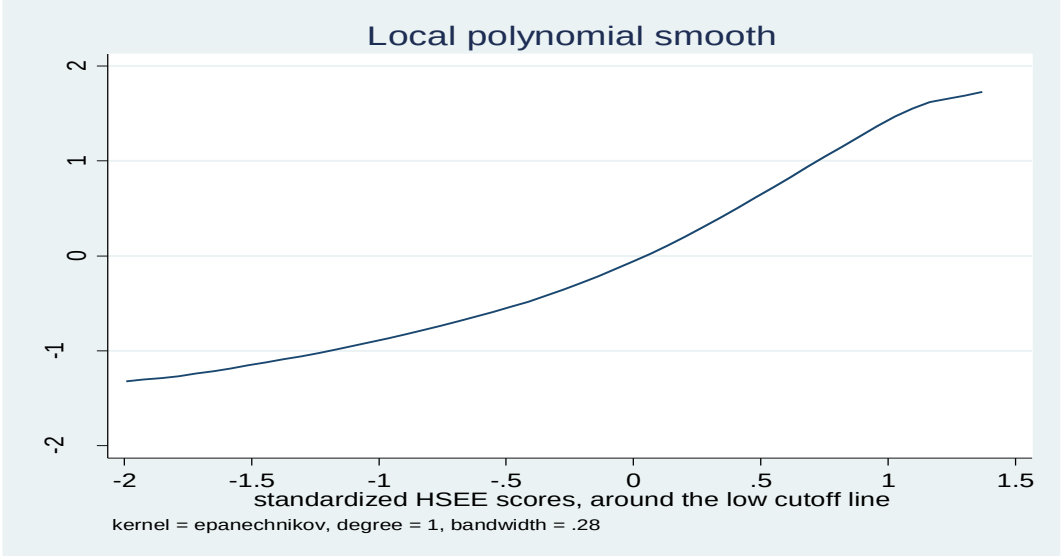
Source: Dee And Lan (2014)

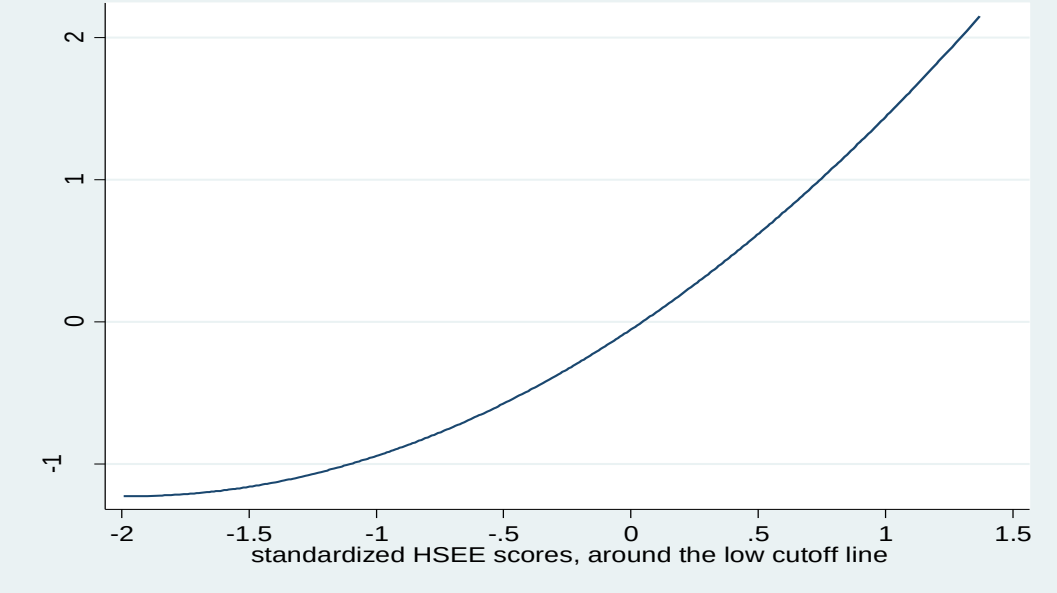

0
-.5
-1
-1.5
-2
-2
-1.5
-1
-.5
0
standardized HSEE scores, around the low cutoff line

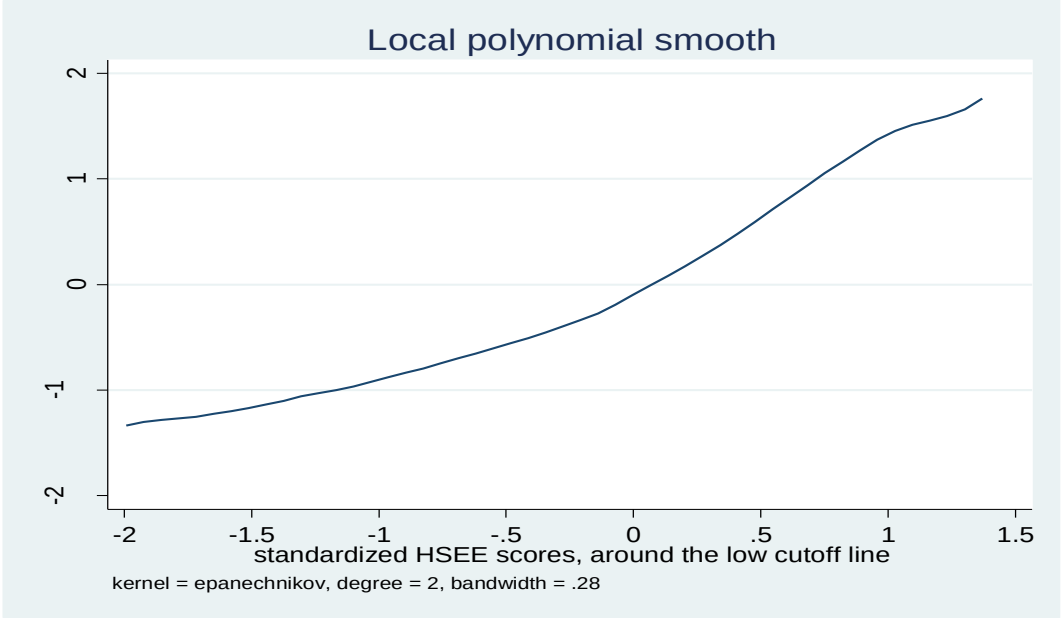

2
1.5
1
.5
0
0
.5
1
1.5
standardized HSEE scores, around the low cutoff line

OLS And Fan: Linear And Quadratic

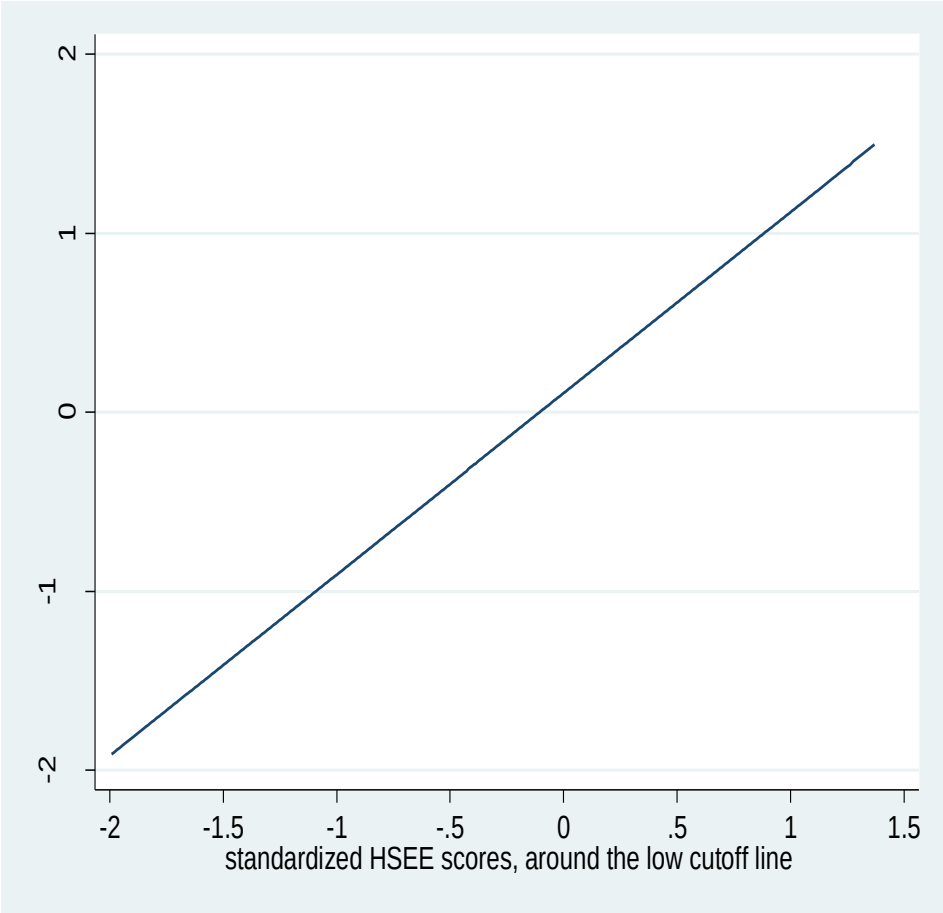


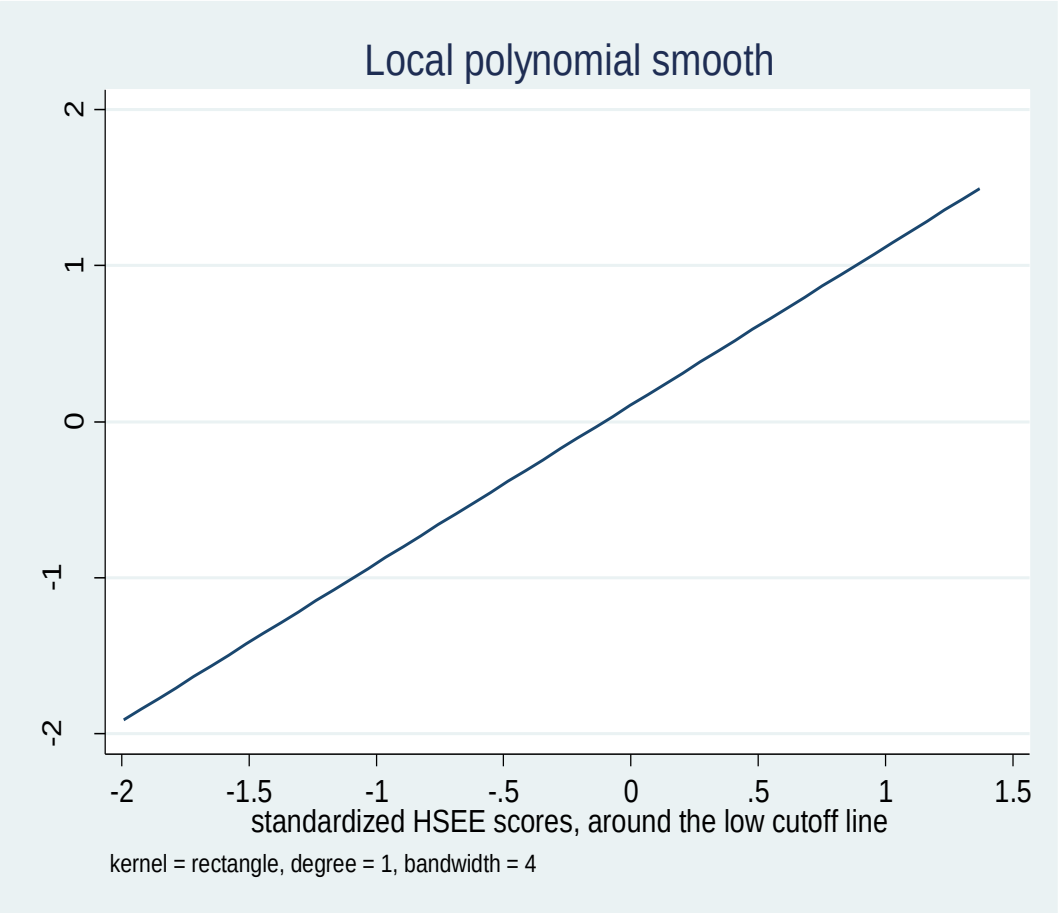






FAN=OLS If Bandwidth = Global, Weight = Equal





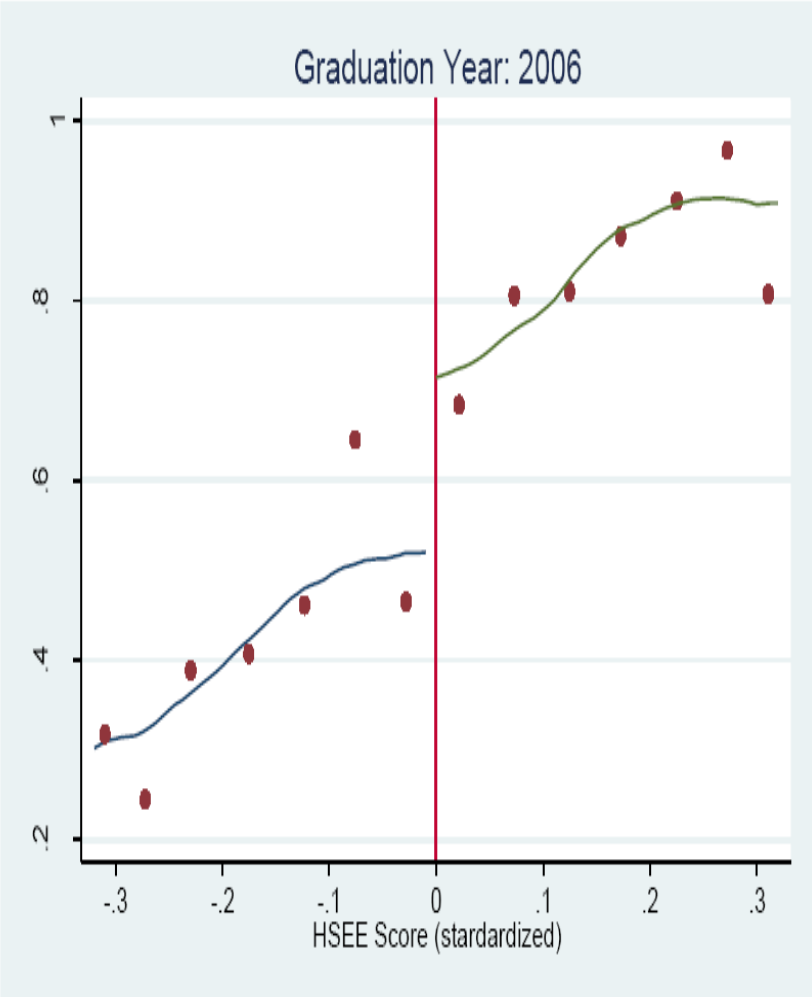
REGRESSION DISCONTINUITY: 中考，重点高中，高中表现

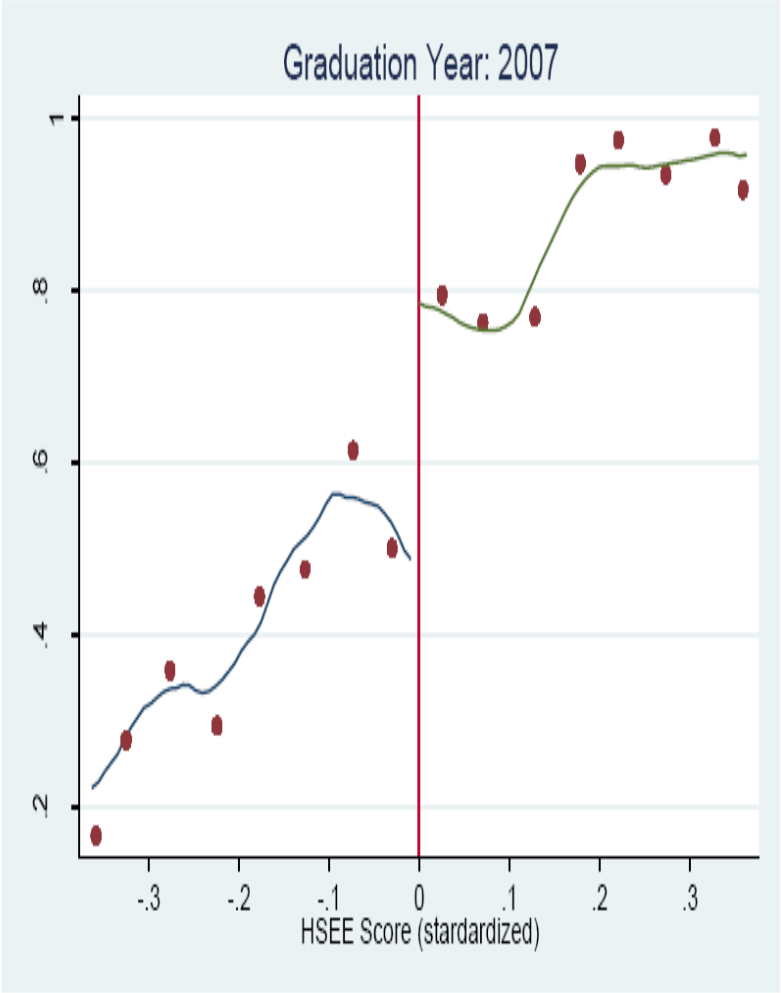
Sharp v.s. fuzzy

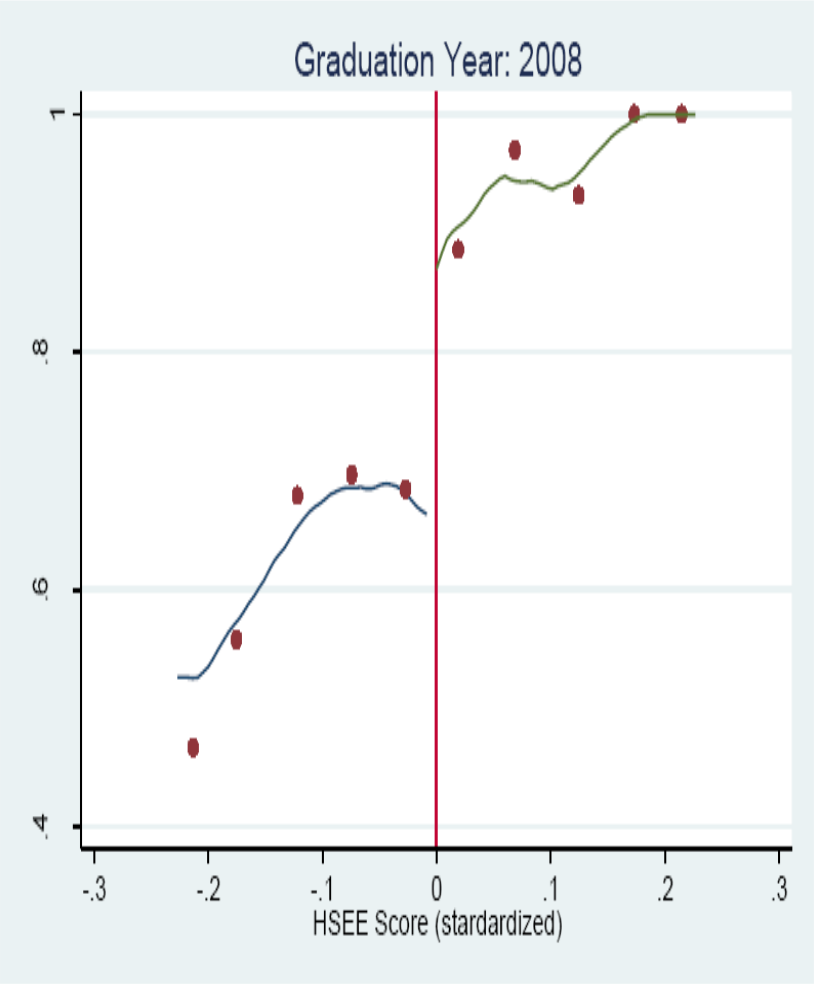
Randomness

Narrow interval around the cutoff: trade-off, fewer observations

中考成绩，重点高中择校线







Source: Dee And Lan (2014)

$$magnet_{it} = (1 - d_{it})f(s_{it}) + d_{it}f(s_{it}) + \beta_1 d_{it} + Year_t + X_{it} + e_{it} \quad (1)$$

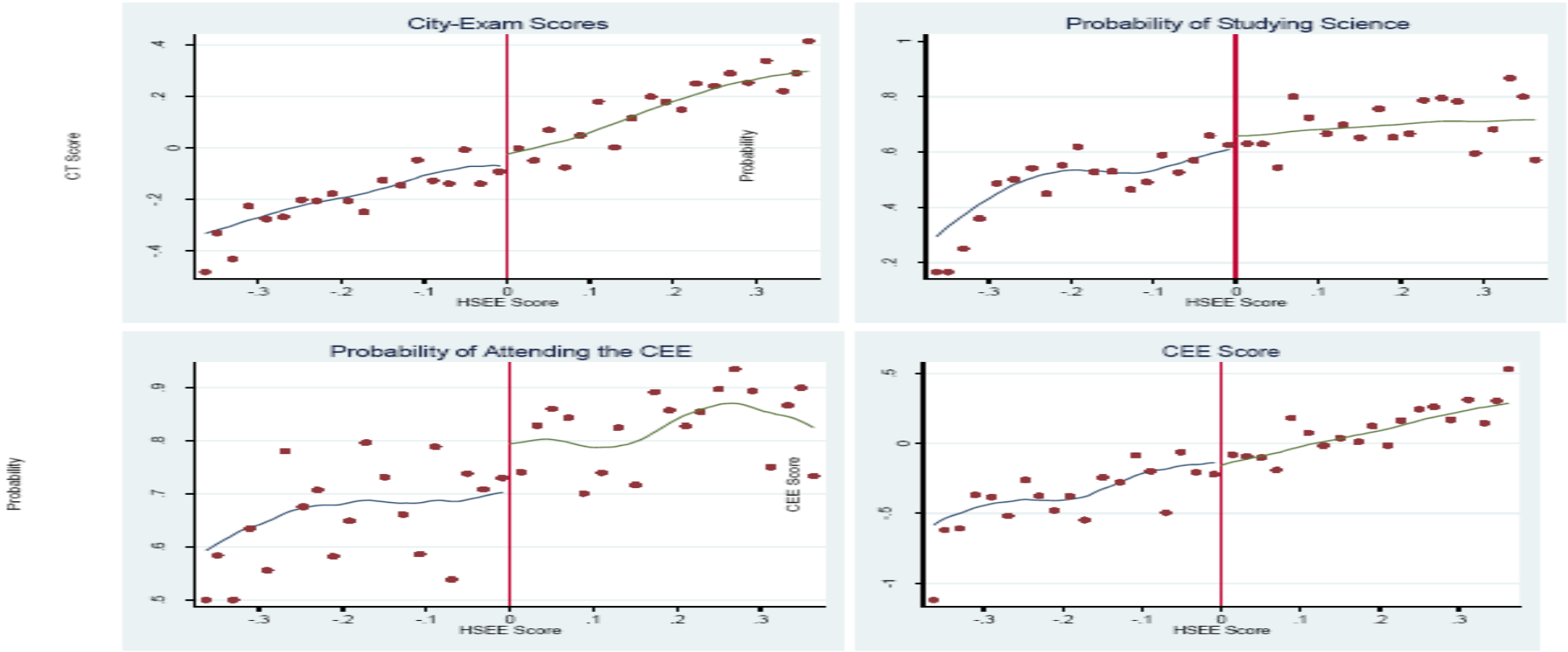
	linear control		2nd-order polynomial		3rd-order polynomial	
	(1)	(2)	(3)	(4)	(5)	(6)
Probability of attending a magnet school	0.134***	0.131***	0.161**	0.155**	0.266***	0.272***
	(0.042)	(0.041)	(0.063)	(0.063)	(0.086)	(0.085)
	1,690	1,690	1,690	1690	1,690	1,690

	[-0.05, 0.05]	[-0.1, 0.1]	[-0.15, 0.15]	[-0.2, 0.2]	local linear regression in the IK bandwidth
	(1)	(2)	(3)	(4)	(5)
Probability of attending a magnet school	0.314**	0.305***	0.211***	0.151***	0.282***
	(0.126)	(0.082)	(0.064)	(0.055)	(0.083)
	267	501	779	1,058	605

Source: Dee And Lan (2014)

高中表现

Figure 2 – High School Outcomes and HSEE Scores



Source: Dee And Lan (2014)

LEE TEST

Table 4 – Students' Characteristics around the Threshold of the Selection Fee

	Parametric Estimation			IK Estimation
	linear control	2nd-order polynomial (2)	3rd-order polynomial (3)	local linear regression (4)
Age	-0.068	-0.118	0.097	-0.012
	(0.073)	(0.107)	(0.139)	(0.103)
	1,690	1,690	1,690	1,002
Female	-0.029	-0.014	-0.091	-0.036
	(0.048)	(0.070)	(0.094)	(0.073)
	1,690	1,690	1,690	888
Minority Status	-0.011	-0.033	-0.028	-0.023
	(0.028)	(0.040)	(0.053)	(0.038)
	1,690	1,690	1,690	1,171

Robust standard errors in parentheses

Source: Dee And Lan (2014)

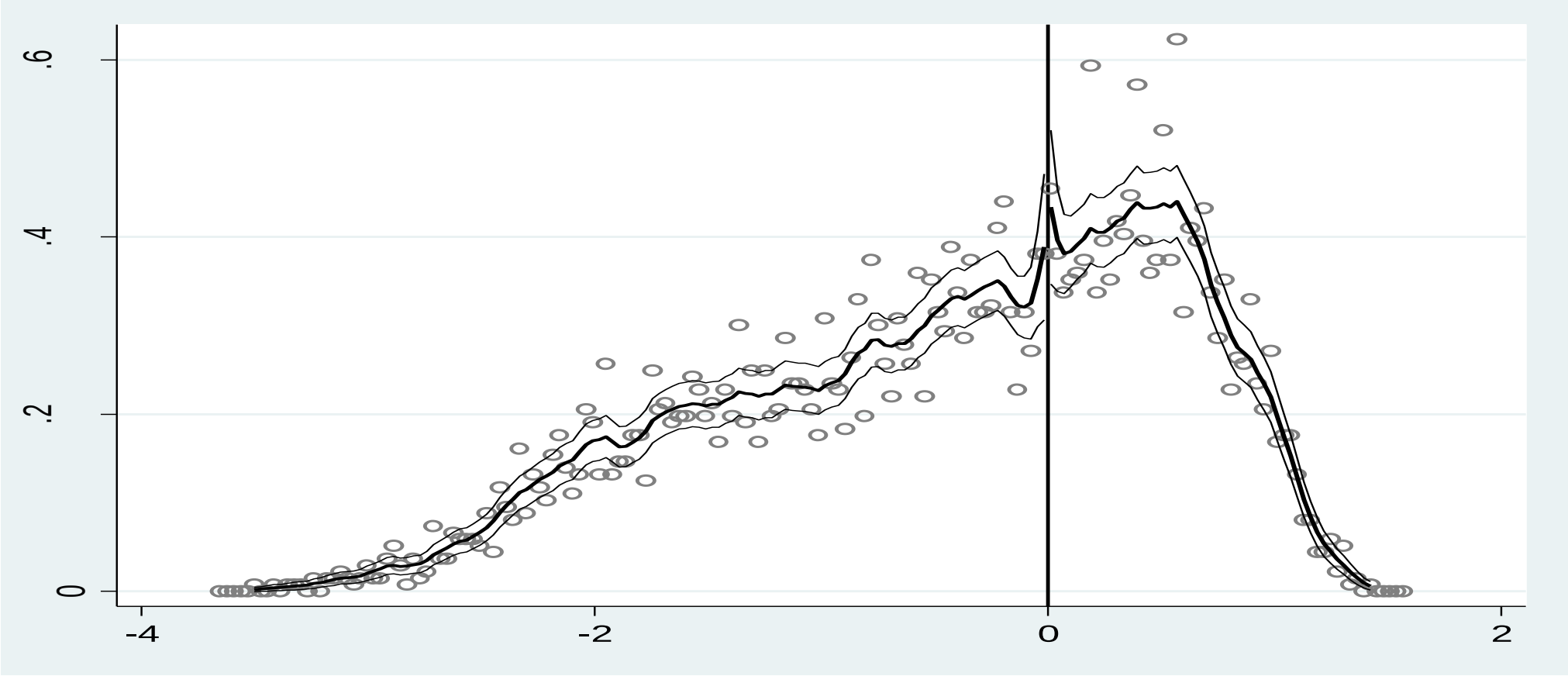
Placebo Test At Other Points

Table A4 – "Placebo" RD Estimates for the Probability of Attending a Magnet School at Alternative Thresholds

HSEE=-0.05	-0.104	HSEE=0.05	0.035
	(0.070)		(0.053)
HSEE=-0.10	0.036	HSEE=0.1	-0.137
	(0.064)		(0.078)
HSEE=-0.15	0.097	HSEE=0.15	0.059
	(0.076)		(0.041)
HSEE=-0.2	0.079	HSEE=0.2	-0.004
	(0.070)		(0.038)

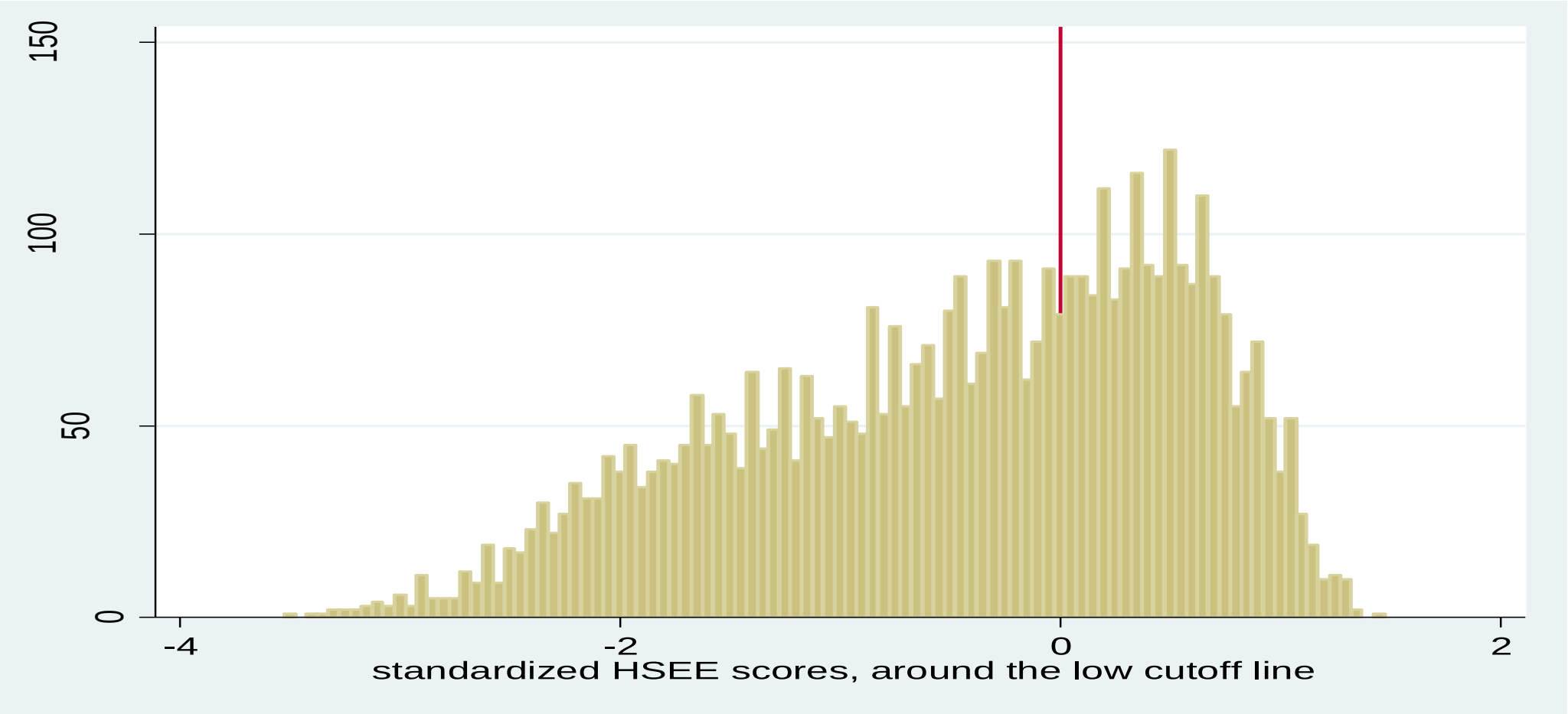
Source: Dee And Lan (2014)

McCrary Test (McCrary, 2008)



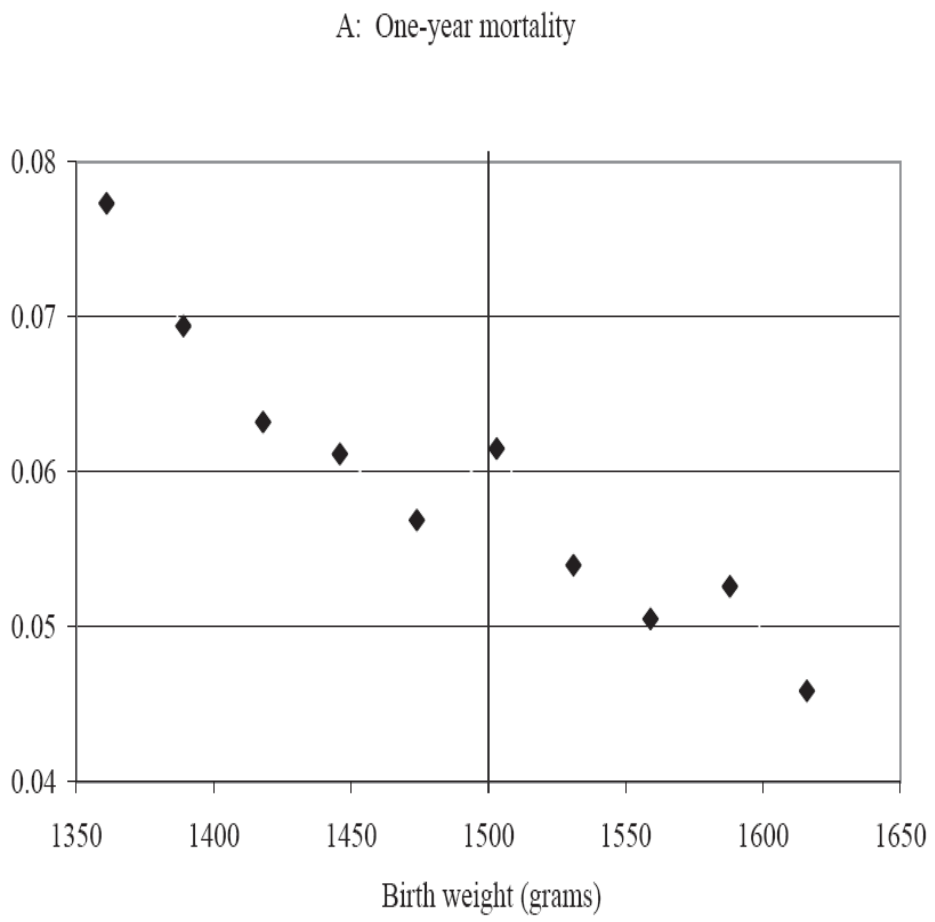
Source: Dee And Lan (2014)

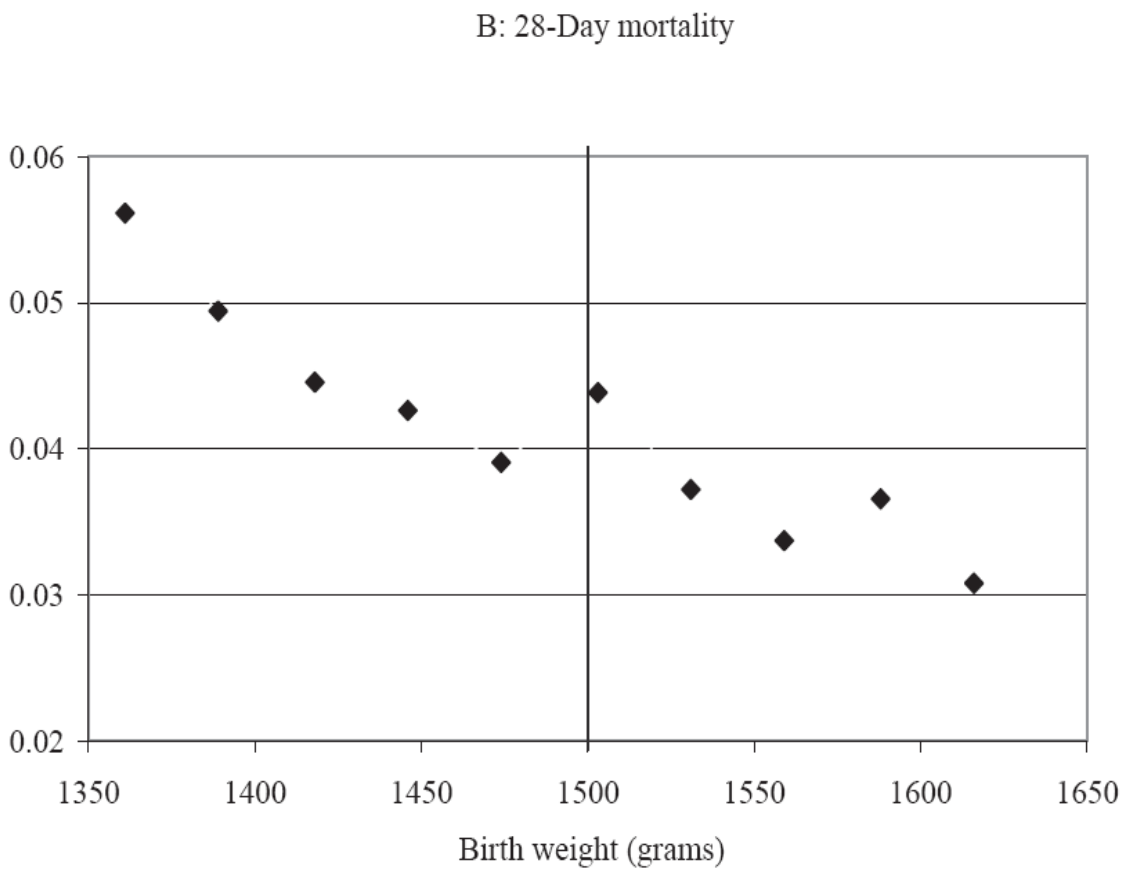
Data Heaping



Source: Dee And Lan (2014)

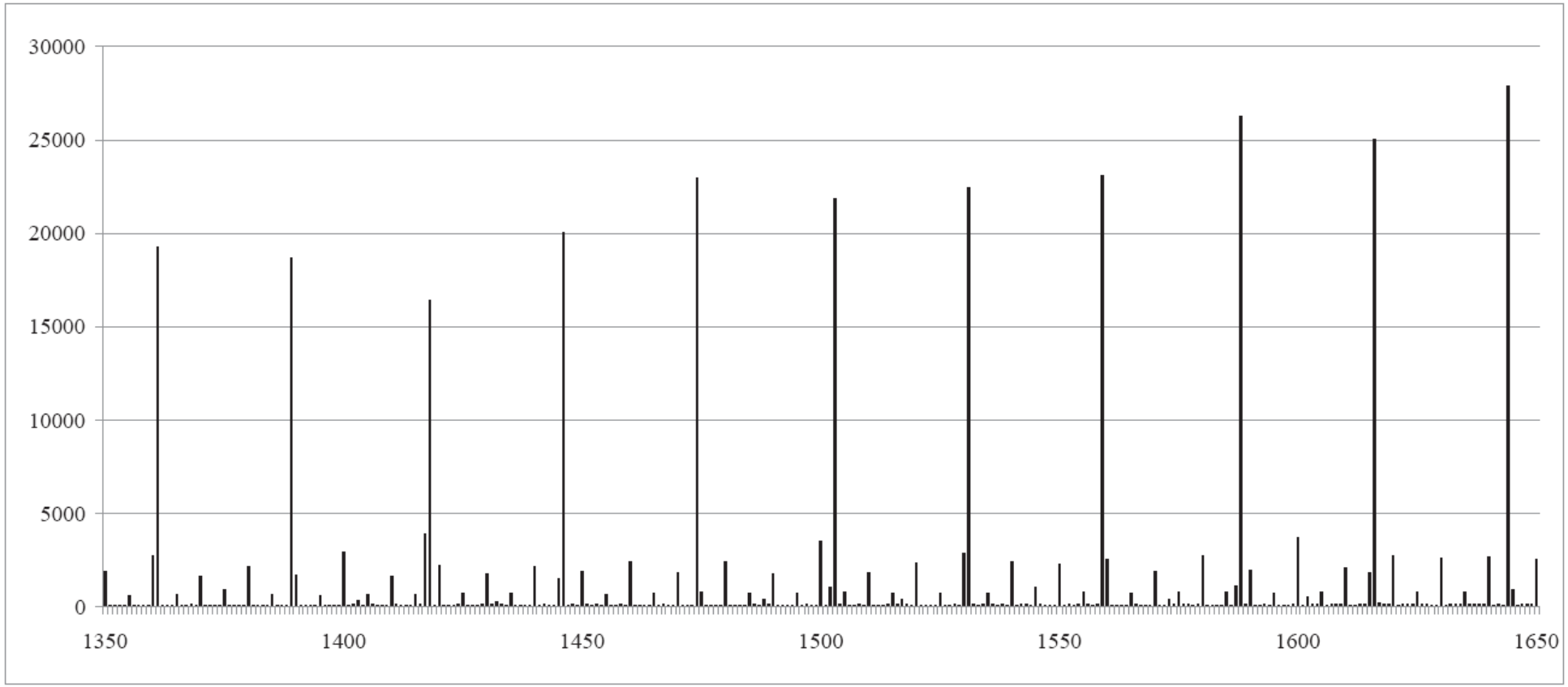
Mortality around 1500 grams in the U.S.: 1983-1991, 1995-2003





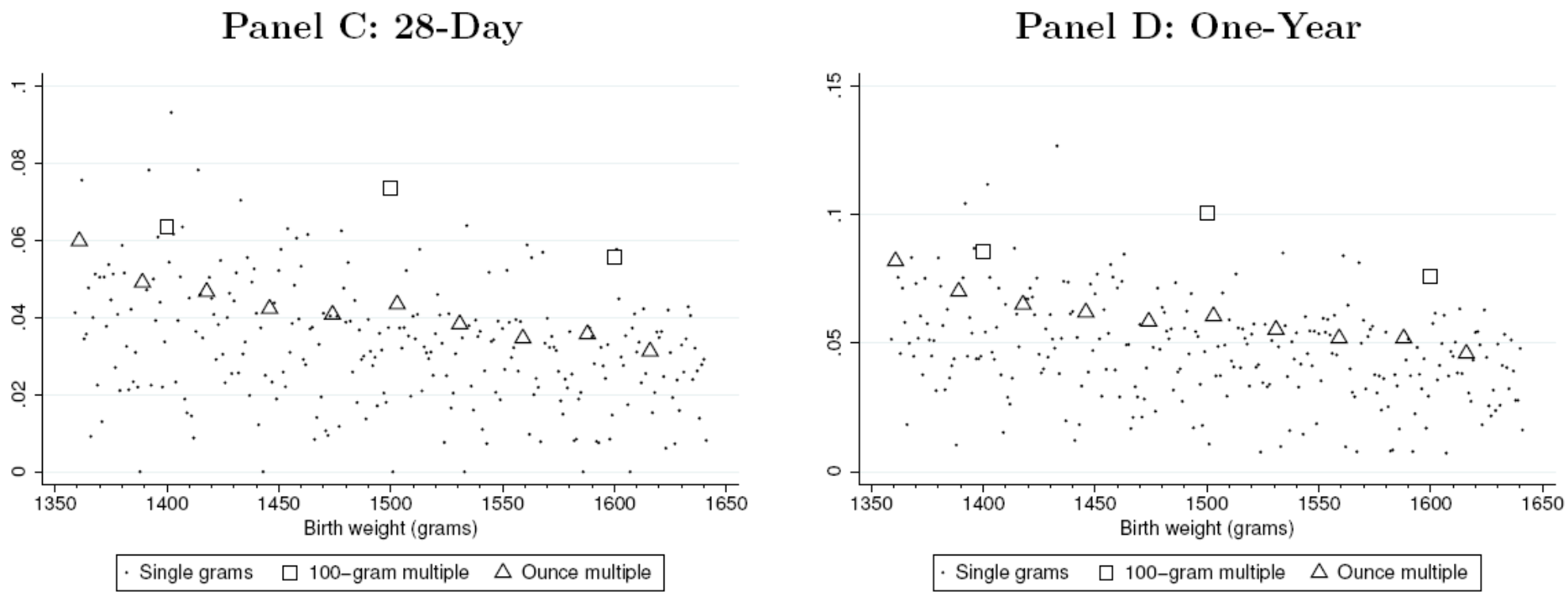
Source: Almond et. al. (2010)

Data Heaping: Frequency of Births by Grams in U.S. between 1350-1650 grams: 1983-1991, 1995-2003



Source: Almond et. al. (2010)

Non-random Heaping in Mortality



SOURCE: Barreca Et. Al. (2011)

Donut RD

Mortality Outcome	One-Year (1)	28-Day (2)	7-Day (3)	24-Hour (4)
Panel A: Our replication of ADKW's estimates				
Weight < 1,500 grams	-0.0071	-0.0071*	-0.0046	-0.0033
	(0.0041)	(0.0032)	(0.0028)	(0.0020)
Observations	202,078	202,078	202,078	202,078
Clusters	171	171	171	171
Panel B: Donut RD dropping those at 1,500-grams				
Weight < 1,500 grams	-0.0033*	-0.0042**	-0.0023	-0.0018
	(0.0014)	(0.0013)	(0.0013)	(0.0010)
Observations	198,534	198,534	198,534	198,534
Clusters	170	170	170	170
Panel C: Donut RD dropping those within 1 gram of 1,500-gram cutoff				
Weight < 1,500 grams	-0.0035*	-0.0043**	-0.0024	-0.00183
	(0.0014)	(0.0012)	(0.0013)	(0.0010)
Observations	198,334	198,334	198,334	198,334
Clusters	168	168	168	168
Panel D: Donut RD dropping those within 2 grams of 1,500-gram cutoff				
Weight < 1,500 grams	-0.0027*	-0.0037**	-0.0019	-0.0013
	(0.0014)	(0.0012)	(0.0012)	(0.0009)
Observations	197,135	197,135	197,135	197,135
Clusters	166	166	166	166
Panel E: Donut RD dropping those within 3 grams of 1,500-gram cutoff				
Weight < 1,500 grams	-0.0018	-0.0026	-0.0018	-0.0011
	(0.0019)	(0.0015)	(0.0015)	(0.0014)
Observations	175,108	175,108	175,108	175,108
Clusters	164	164	164	164

SOURCE: Barreca Et. Al. (2011)

References

Almond, D., J. J. Doyle, A. E. Kowalski, and H. Williams. "ESTIMATING MARGINAL RETURNS TO MEDICAL CARE: EVIDENCE FROM AT-RISK NEWBORNS." *The quarterly journal of economics* 125, no. 2 (2010): 591-634.

Barreca, Alan I., Melanie Guldi, Jason M. Lindo, and Glen R. Waddell. "Saving Babies? Revisiting the effect of very low birth weight classification." *The Quarterly Journal of Economics* 126, no. 4 (2011): 2117-2123.

Dee, Thomas, and Xiaohuan Lan (2014) "The Achievement and Course-Taking Effects of Magnet Schools: Regression-Discontinuity Evidence from Urban China", Mimeo, Fudan University

McCrary, Justin (2008): Manipulation of the running variable in the regression discontinuity design: A density test. Journal of Econometrics, 142(2): 698-714.