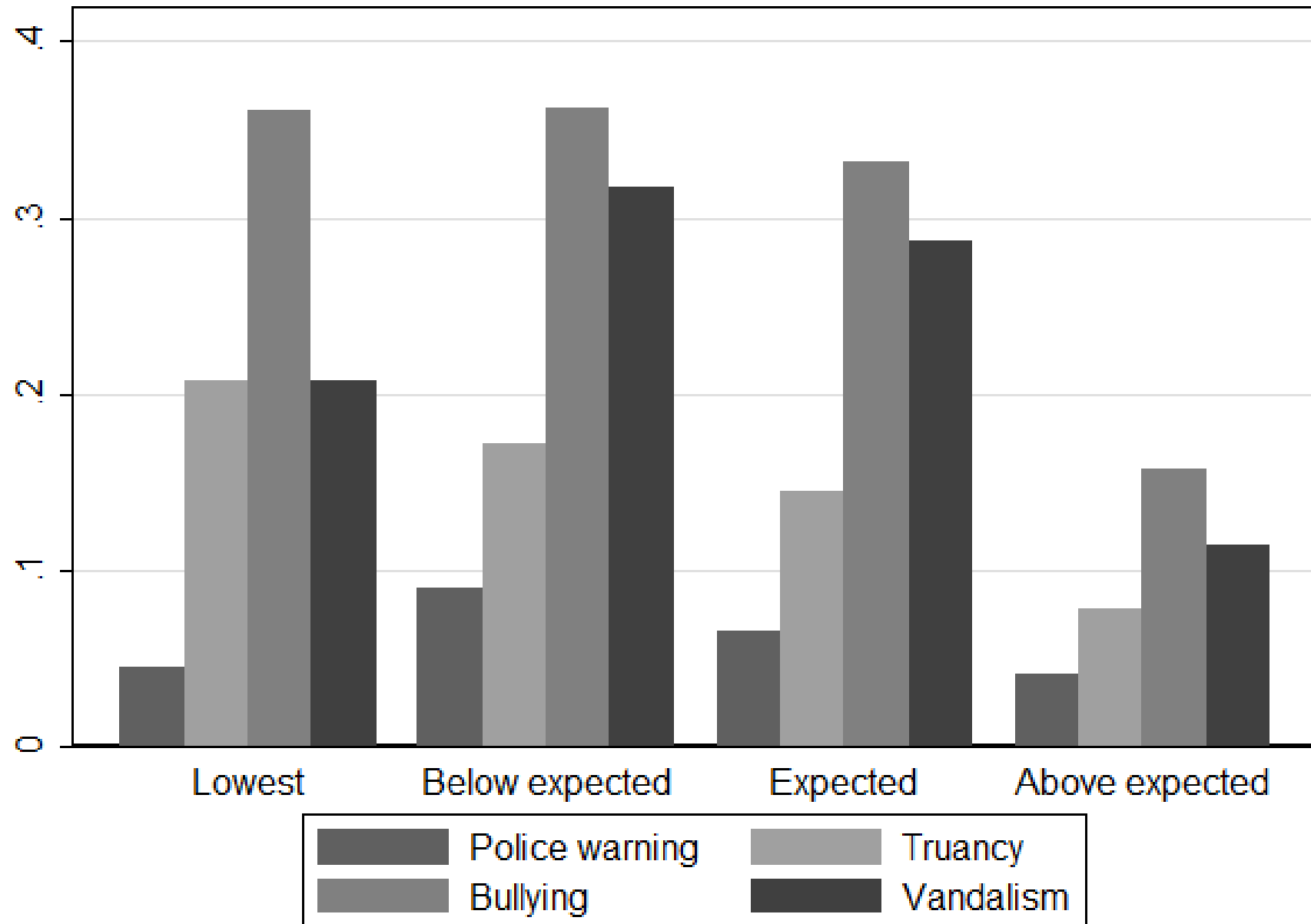# Does noisy feedback affect behaviour?
Evidence from discontinuities in test scores in England

Marcello Sartarelli

INSTITUTE OF EDUCATION

June 24, 2010

Behaviour of students by achievement level in 2001 (age 10-11, UK)

Feedback and behaviour of students

Using feedback (in education):

- helps to build human capital / signal ability through achievement

- fosters positive behaviour, e.g. lead a healthy and safe life (in students)

Sharp increase in interest in the role of education and students' behaviour:

- "No Child Left Behind" (14bln $) by the US Department of Education

- "Every child matters" (5-15bln £) by the UK Department of Children, School and Family (DCSF)

Contribution and findings

- Can noisy feedback affect on behaviour? What's the story?

- Substitution and income effects help to explain behaviour, e.g. police warnings and bullying, of students in secondary schools in England

News coverage on compulsory education in the UK

More parents teach their children at home, Guardian 2nd February 2005

- 20000 families took their children out of school in the past 12 months
- Bullying is the main concern of families opting out of mainstream education

Young pupils fuel record truancy, BCC 26th February 2009

- Unauthorised absence in primary schools was up from 0.52% to 0.57%
- 5% of all enrolments in secondary schools are classed as persistent absentees

Institutional setting: the British national school curriculum

(1) Age	(2) Stage	(3) Year	(4) Assessment	(5) Expected achievement level
3-4	Early Years			
4-5		Reception		
5-6	Key Stage 1	1		
6-7		2	Teacher assessments in English, Maths and Science (EMS)	2
7-8	Key Stage 2	3		
8-9		4		
9-10		5		
10-11		6	National and teacher assessments in EMS	4
11-12	Key Stage 3	7	Teacher assessments	
12-13		8	Teacher assessments	
13-14		9	National and teacher assessments in EMS and foundation subjects	5 or 6
14-15	Key Stage 4	10	Some children take GCSEs	
15-16		11	Most children take GCSEs or other national qualifications	

Toy model of effort, leisure, behaviour and feedback

Name	Quantity	Price
Effort	e	Ability a
Leisure	l	1 (unity)
Behaviour	b	Probability of punishment p

Toy model of effort, leisure, behaviour and feedback (cont.d)

$$\max_{e,b} u(y, l) \tag{1}$$

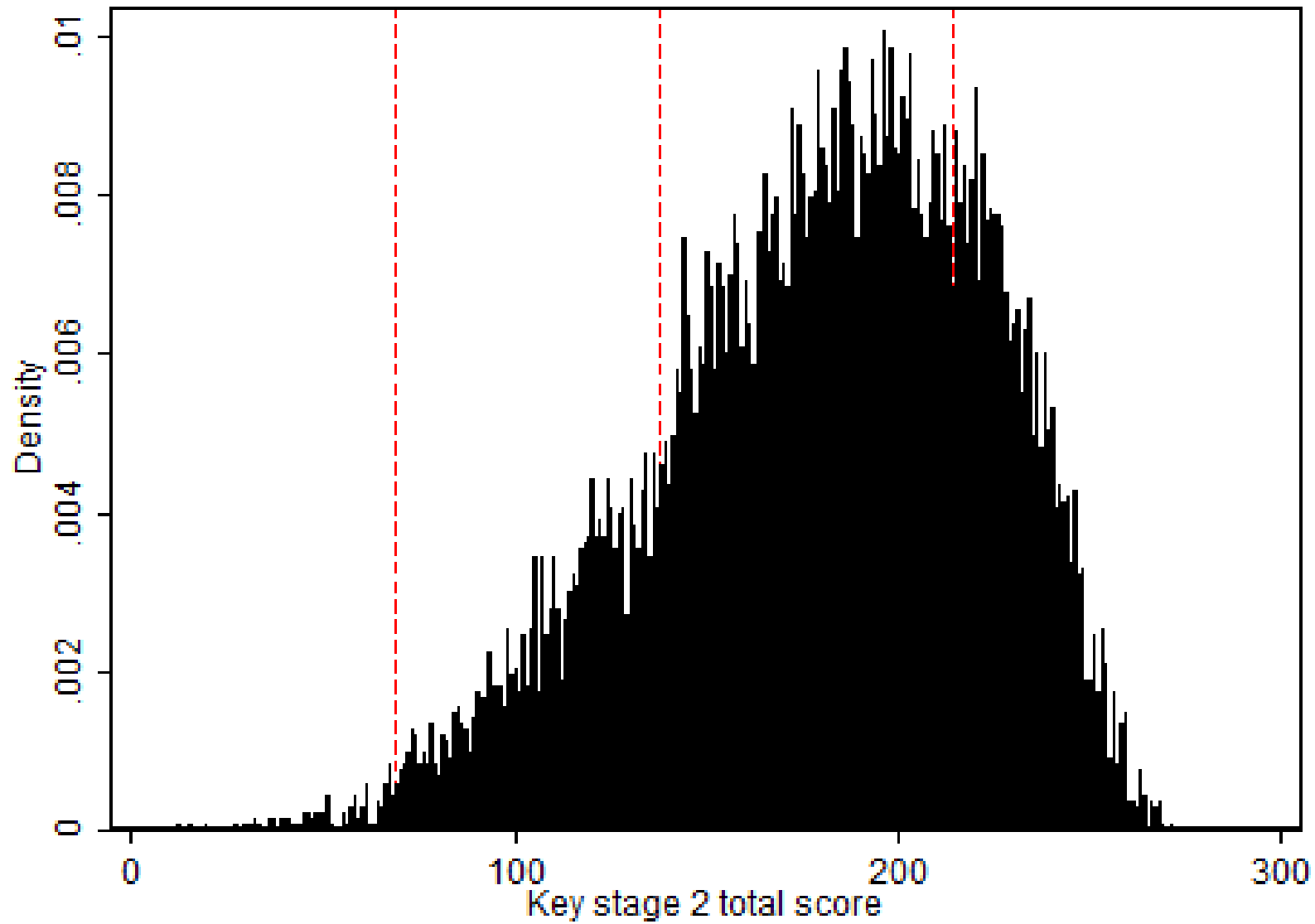$$\text{s.t. } y = f(ea, pb - (1-p)b) \text{ and } e + l + b \leq T$$

$$f'_e u'_e - u'_l = 0, \ MRS_{l,e} = f'_e \tag{2}$$

$$f'_b u'_b - u'_l = 0, \ MRS_{l,b} = f'_b \tag{3}$$

$$...MRS_{b,e} = \frac{f'_e}{f'_b} \tag{4}$$

$$\text{then } \frac{db}{da} = \text{substitution effect } + \text{ income effect} \overset{\leq}{\underset{>}{}} 0? \tag{5}$$

Distribution of test scores and cutoffs determining achievement level

Construction of total score, synthetic cutoffs & McCrary's test

	2-3	3-4	4-5
cutoffs English	29	50	70
t-test	5.73	18.79	23.28
cutoffs Maths	22	49	79
t-test	6.03	9.72	10.99
cutoffs Science	18	39	65
t-test	3.23	10.53	19.51
cutoffs total score	69	138	214
t-test	2.48	4.87	4.82

Institutional setting: grading of Key Stage 2 tests

Teachers are not involved in grading their students' test scripts

KS tests are marked anonymously by national curriculum tests markers using a scale 0-100+

Students and parents get the
- discrete teacher assessment levels that the child has achieved

- average achievement level

  − for all the children in a child's age group in the same school
  − in the previous year in England

Grading systems and tables are hardly manipulable as they are periodically reviewed jointly by the Qualifications and Curriculum Authority

Summary statistics by total achievement level

Achievement level	English ≤ 3	**English** $= 4$	English > 4	Maths ≤ 3	**Maths** $= 4$	Maths > 4	Science ≤ 3	**Science** $= 4$	Science > 4
2	0.34	0.00	0.66	0.17	0.00	0.83	0.77	0.02	0.22
3	0.74	0.21	0.05	0.86	0.06	0.08	0.45	0.54	0.00
4	0.11	0.73	0.15	0.15	0.76	0.08	0.01	0.78	0.21
5	0.04	0.02	0.93	0.08	0.03	0.89	0.08	0.08	0.85

Summary statistics

	Mean	StdDev
Police warning in wave 1	0.06	.
Bullying episosed in wave 1	0.30	.
Ever truant in wave 1	0.13	.
Ever smoked cigarettes in wave 1	0.08	.
Ever tried alcohol in wave 1	0.39	.
Ever smoked cannabis in wave 1	0.07	.
Vandalism episodes in wave 1	0.25	.
White	0.59	.
Black	0.07	.
Asian	0.17	.
Other	0.06	.
SEN statement	0.03	.
SEN non-statemented	0.14	.
Free school meals	0.18	.
English additional language	0.28	.
Voluntary aided schools	0.16	.
Voluntary controlled schools	0.08	.
Other schools	0.02	.
Community schools	0.64	.
Key stage 2 total score	178.91	43.34
Key stage 2 English score	59.37	14.23
Key stage 2 Maths score	61.72	20.92
Key stage 2 Science score	56.58	13.20

Timescale of Key Stage 2 tests and behaviour of students

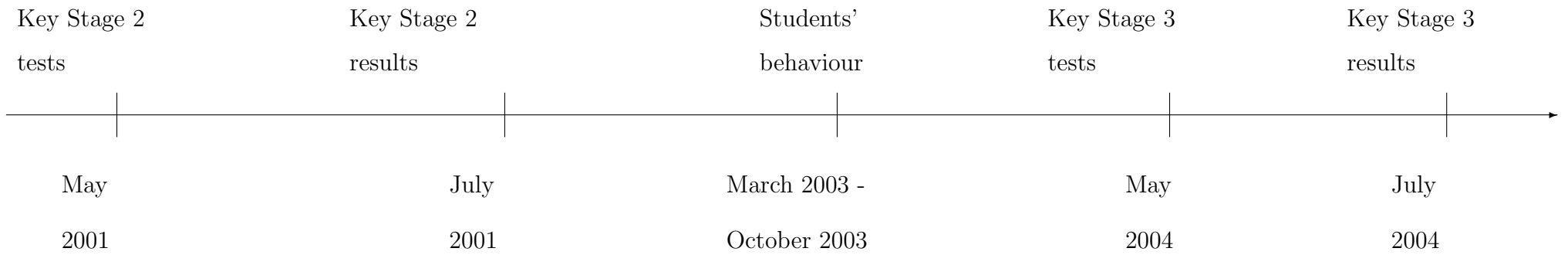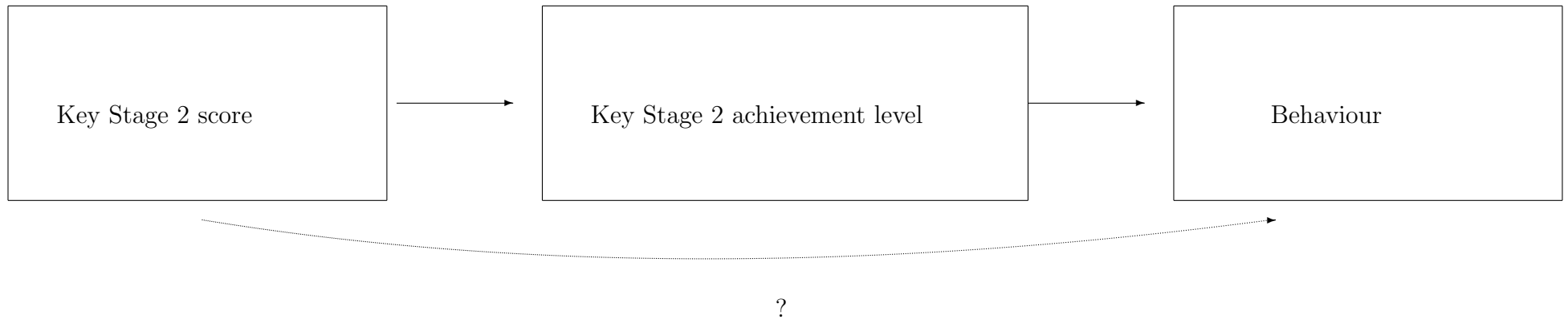| Key Stage 2 tests	Key Stage 2 results	Students' behaviour	Key Stage 3 tests	Key Stage 3 results
May 2001	July 2001	March 2003 - October 2003	May 2004	July 2004

Illustration of the research design

Research design

$$B_i^* = \alpha + \beta A_i + U_i$$

$$B = \begin{cases} 1 \text{ if } B^* \geq 0 \\ 0 \text{ otherwise} \end{cases} \tag{6}$$

$$B_i = \alpha + \beta A_i + U_i \tag{7}$$

B^* is latent continuous measure of behaviour

B equals 1 in the event of e.g. police warning if $B^* \geq 0$ and 0 otherwise

A is binary achievement level, e.g. pass/fail; unobservables may correlate with A and B, e.g. parental guidance

S is a continuous test score underlying A and $\bar{S}$ is an administrative cutoff in S

$\beta < (>)\ \ 0$: does substitution income effect dominate? between what activities?

Results: dependent variable police warning

	OLS		RD			
cutoff 3-4	-0.029***	-0.016*	0.019*	0.023**	0.016	0.020
S.d.	(0.0061)	(0.0066)	(0.0087)	(0.0088)	(0.015)	(0.019)
AIC	349.3	33.0	-305.3	-546.0	-552.3	-549.3
cutoff 4-5	-0.023***	-0.024***	-0.011	-0.015*	0.0091	0.0060
	(0.0043)	(0.0043)	(0.0061)	(0.0062)	(0.010)	(0.013)
AIC	350.0	12.7	-302.6	-543.1	-551.1	-547.5
Covariates	N	Y	N	Y	Y	Y
Higher order	N	N	N		2nd	3rd
N	14183	14183	13298	13298	13298	13298

Results: dependent variable bullying

	OLS			RD		
cutoff 3-4	-0.042***	-0.043***	0.026	0.033*	0.017	-0.030
	(0.011)	(0.011)	(0.016)	(0.016)	(0.025)	(0.033)
AIC	18772.9	17829.4	17442.1	16595.0	16595.4	16589.9
cutoff 4-5	-0.040***	-0.059***	-0.038**	-0.045***	-0.036	-0.044
	(0.0090)	(0.0085)	(0.013)	(0.013)	(0.021)	(0.027)
AIC	18767.6	17797.2	17436.2	16586.9	16588.7	16591.6
Covariates	N	Y	N	Y	Y	Y
Higher order	N	N	N		2nd	3rd
N	14183	14183	13298	13298	13298	13298

Results: dependent variable truancy

	OLS		RD			
cutoff 3-4	-0.040***	-0.016	0.028*	0.032**	-0.012	-0.0041
	(0.0087)	(0.0094)	(0.012)	(0.013)	(0.022)	(0.028)
AIC	10169.5	9981.5	9108.8	8990.0	8985.0	8988.4
cutoff 4-5	-0.040***	-0.035***	-0.034***	-0.040***	-0.021	-0.015
	(0.0065)	(0.0066)	(0.0090)	(0.0092)	(0.015)	(0.019)
AIC	10155.9	9955.3	9101.1	8979.7	8982.8	8986.3
Covariates	N	Y	N	Y	Y	Y
Higher order	N	N	N		2nd	3rd
N	14183	14183	13298	13298	13298	13298

Results: dependent variable vandalism

	OLS			RD		
cutoff 3-4	-0.045***	-0.027*	0.030*	0.032*	-0.054*	-0.045
	(0.011)	(0.012)	(0.015)	(0.015)	(0.023)	(0.033)
AIC	17327.7	17035.3	16155.5	15908.1	15871.4	15874.9
cutoff 4-5	-0.063***	-0.056***	-0.059***	-0.064***	-0.024	-0.037
	(0.0083)	(0.0083)	(0.012)	(0.012)	(0.020)	(0.026)
AIC	10155.9	9955.3	9101.1	8979.7	8982.8	8986.3
Covariates	N	Y	N	Y	Y	Y
Higher order	N	N	N		2nd	3rd
N	14183	14183	13298	13298	13298	13298

Robustness checks: pre-treatment covariates, cutoff 3-4

	OLS coeff.	S.e.
Black	.09	(.10)
Asian	-.10	(.13)
Other	.06	(.05)
SEN statement	.01	(.05)
SEN non-statemented	.32	(.13)**
Free school meals	.01	(.13)
English additional language	-.09	(.13)
Voluntary aided schools	.05	(.11)
Voluntary controlled schools	0.08	(.10)
Community school	.02	(.02)

Robustness checks: gaming around cutoffs
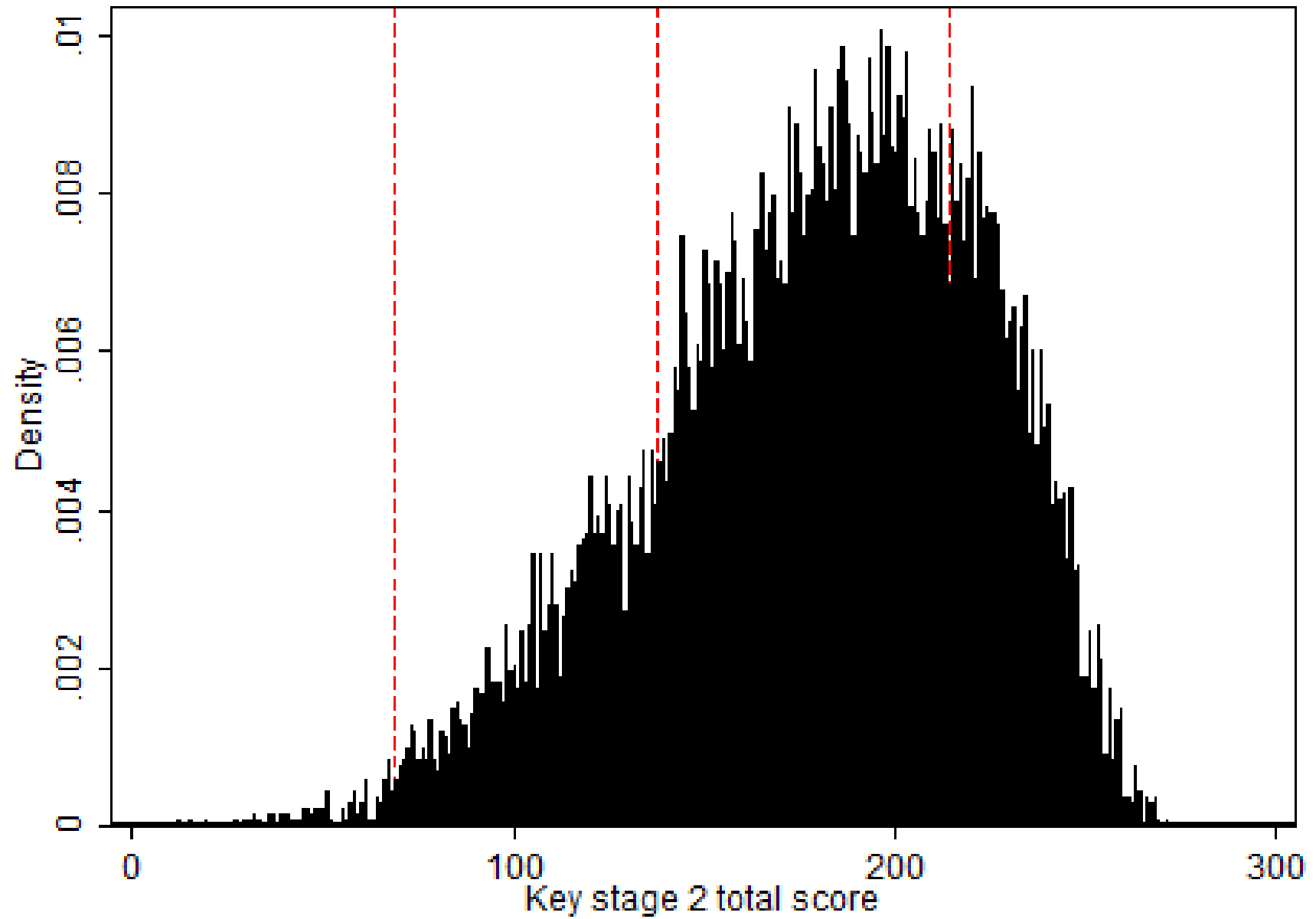
Squaring interpretation of estimates with theory
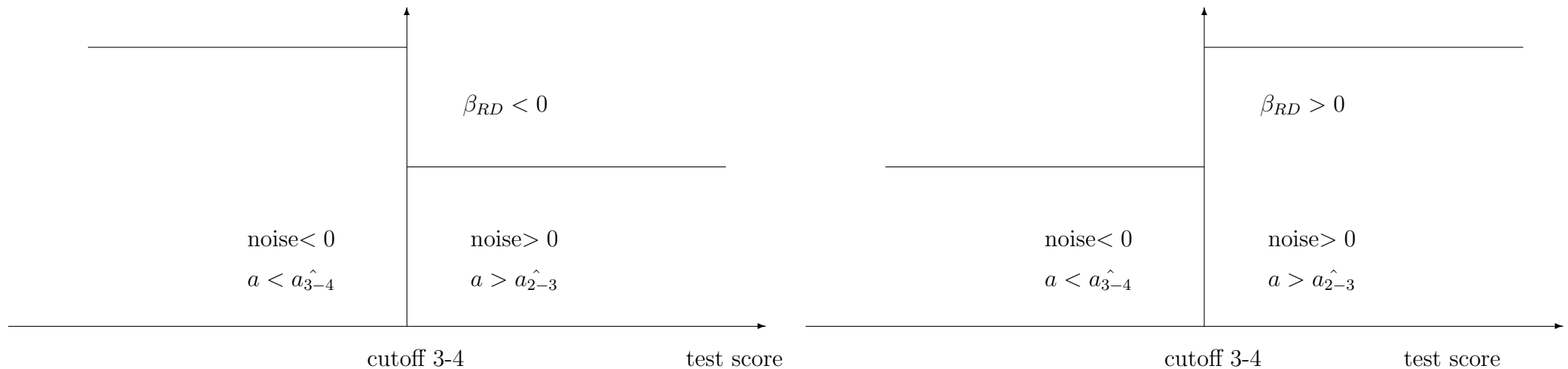
$\beta_{RD} < 0$
income effect dominates if noise<0
substitution effect dominates if noise>0
Under the ceteris paribus assumption that

$\beta_{RD} > 0$
substitution effect dominates if noise<0
income effect dominates if noise>0
leisure is unchanged

Discussion

- OLS is mostly downward biased w.r.t. RD

- No effect at policy-relevant score cutoff 3-4 can be due to gaming and/or substitution and income effects netting off

- Negative effect at cutoff 4-5 suggests that substitution (income) effect dominates for students with $a > (<)\bar{a}$

- I extend the results in Azmat and Iriberri (2009) and Bandiera et al. (2009) by identifying the effect of feedback beyond test scores and reconciling with theory

- Novel application on feedback in test scores and behaviour to the work on

   − non-market outcomes in Grossman (2005) a Gaviria and Raphael (2001)

   − confidence and motivation in Benabou and Tirole (2002, 2003)

Next steps

- Longitudinal data on test scores and behaviour

- Effect on achievement in the school leaving age exam (GCSE) and beyond: A-level, enrollment at college and degree choice

- Timing of recording of information in the survey data and administrative data

- Multiple cutoffs: incentives and RD

References

Azmat, G. and N. Iriberri (2009) *The Importance of Relative Performance Feedback Information: Evidence from a Natural Experiment using High School Students*, CEP Discussion Paper 0915

Bandiera, O., V. Larcinese and I. Rasul (2009) *Blissful Ignorance? Evidence From a Natural Experiment on The Effect of Individual Feedback on Performance*, UCL mimeo

Benabou, R. and J. Tirole (2002) *Self-Confidence and Personal Motivation*, Quarterly Journal of Economics 117 (3), 871-915

Benabou, R. and J. Tirole (2003) *Intrinsic and Extrinsic Motivation*, Review of Economic Studies 70, 489-520

Gaviria, A. and S. Raphael, (2001) *School-based peer effects and juvenile behavior*, Review of Economics and Statistics, 83 (2), 257-268

Grossman, M. (2005) *Education and Non Market Outcomes*, Chapter 10 in the *Handbook of the Economics of Education*, E. Hanushek and F. Welch (eds.), North Holland, Amsterdam

Lee, D. and D. Card (2008) *Regression Discontinuity Inference with Specification Error*, Journal of Econometrics 142 (2), 655-674

Lee, D. and T. Lemieux (2009), *Regression Discontinuity Designs in Economics*, NBER Working Paper 14723

Lizzeri, A. and M. Siniscalchi (2008) *Parental Guidance and Supervised Learning*, Quarterly Journal of Economics 123, 1161-1195

McCrary, J. (2008) *Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test*, Journal of Econometrics 142, 698-714

Oreopoulos, P. and K. Salvanes (2009) *How large are returns to schooling? Hint: Money isn't everything*, NBER Working Paper 15339

Urquiola, M. and E. Verhoogen (2009) *Class-Size Caps, Sorting, and the Regression-Discontinuity Design*, American Economic Review 99 (1), 179-215