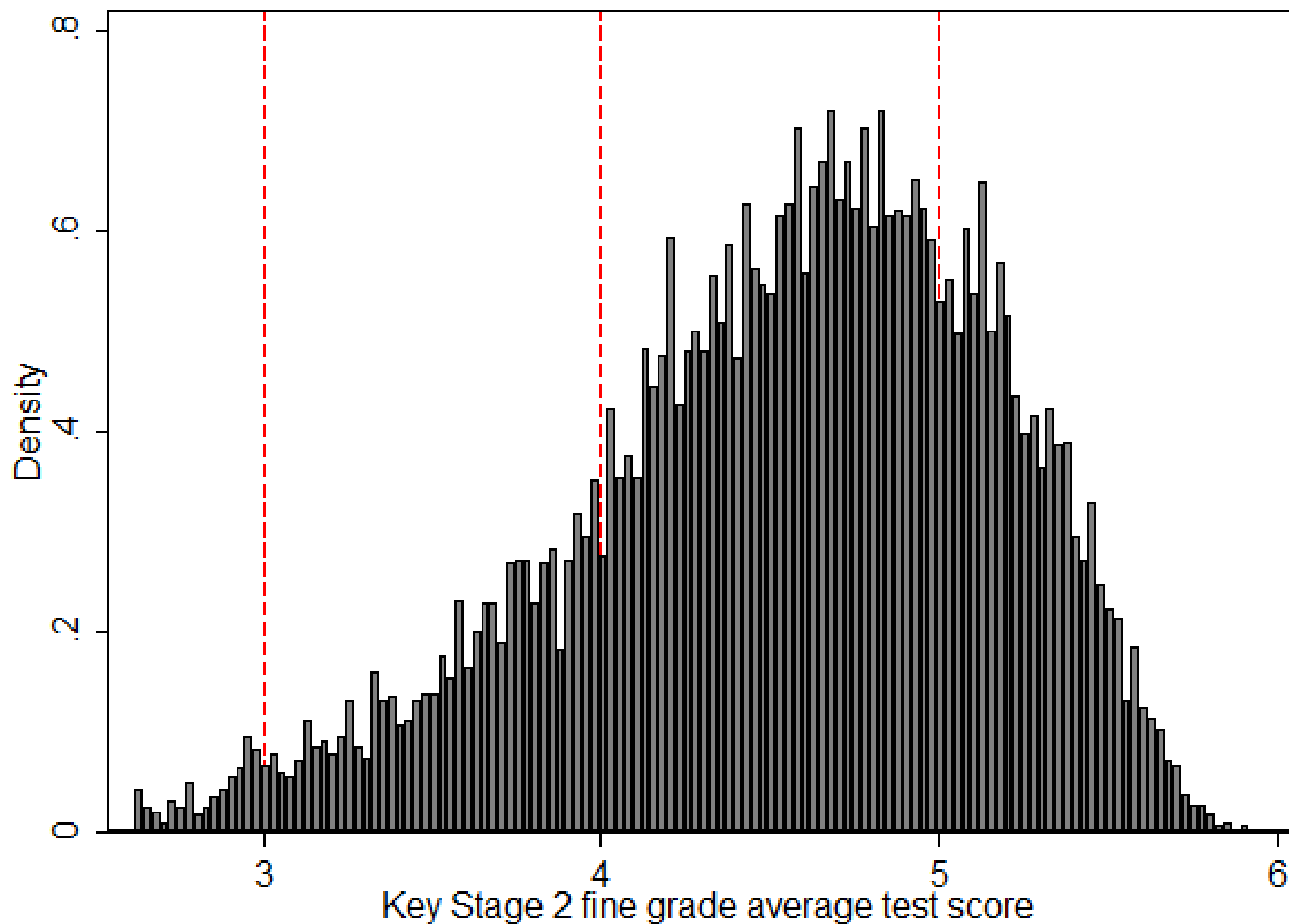# Do performance targets affect behaviour? Evidence from test scores in England
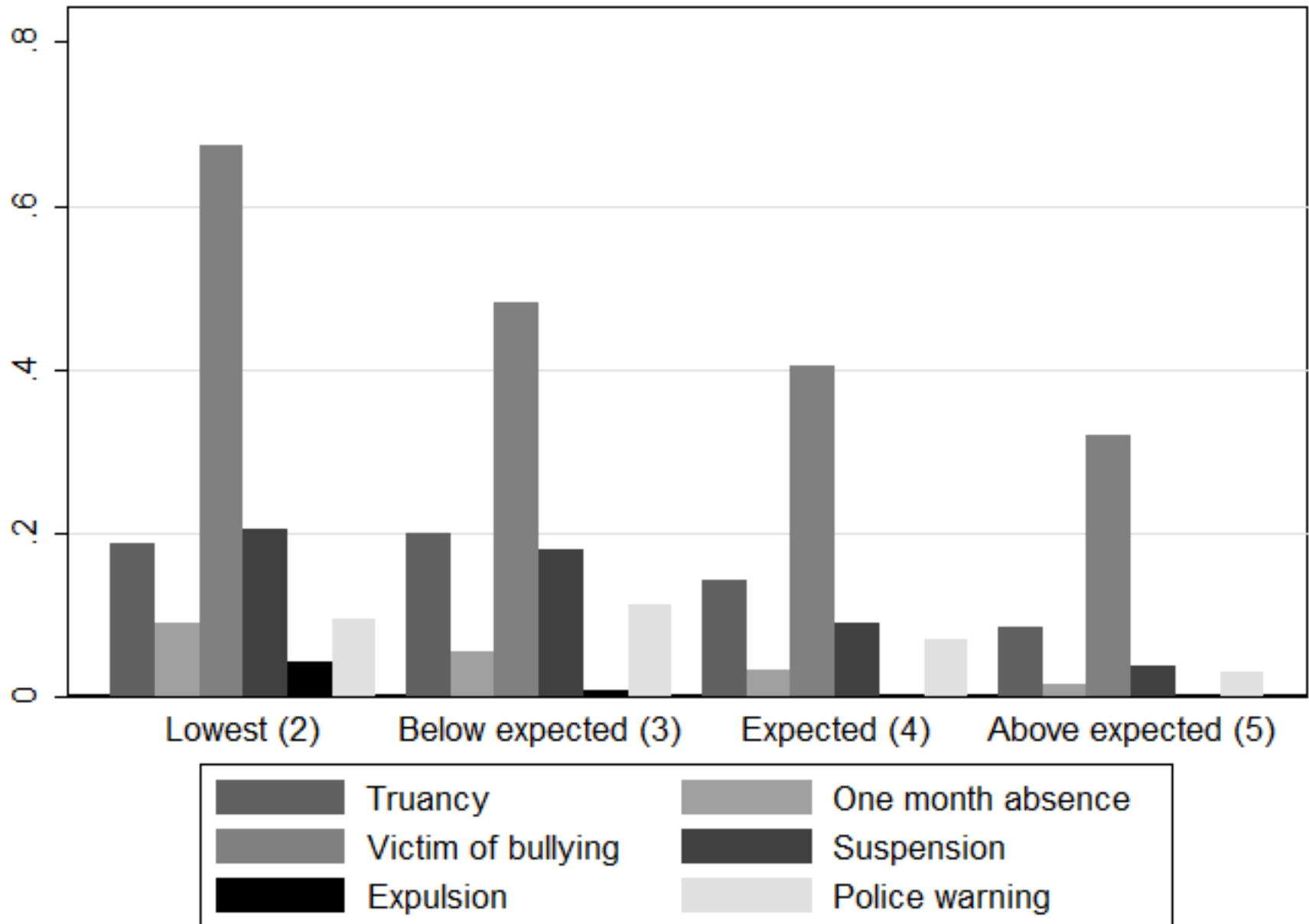
Marcello Sartarelli

INSTITUTE OF EDUCATION

January 20, 2010

Performance targets in test scores (NPD admin data)

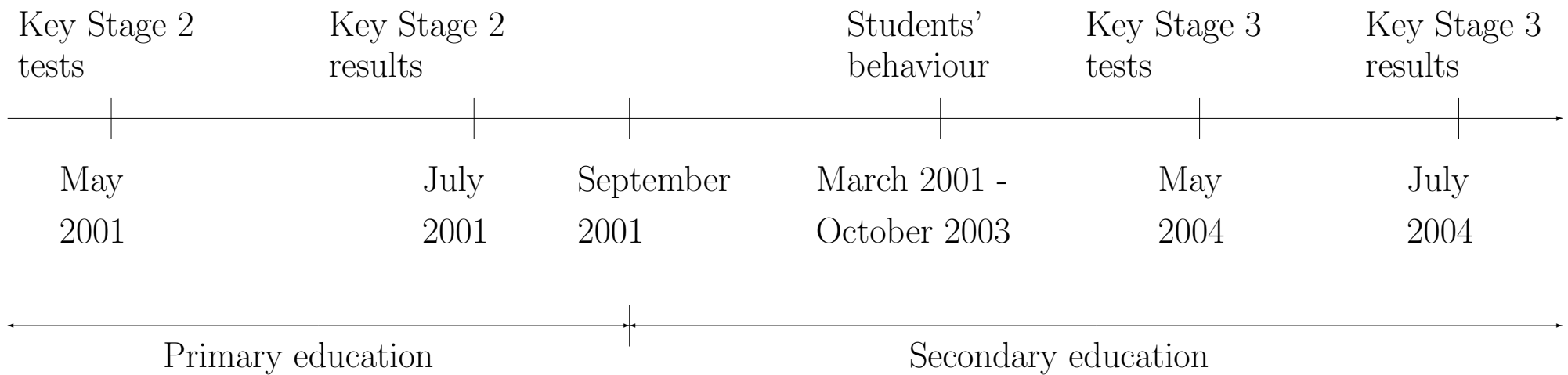# Civic behaviour by students (LSYPE survey data)

School curriculum in the UK

(1) Primary/ Secondary	(2) Age	(3) Stage	(4) Year	(5) Assessment	(6) Expected achievement level
	3-4	Early Years Foundation Stage (EYFS)			
	4-5	EYFS	Reception	Tests	6-9/13 elements
Primary School	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
Secondary School	11-12	Key Stage 3	7	Teacher assessments	
	12-13		8	Teacher assessments	
	13-14		9	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	5 A*-C or equivalent including English and Maths

Behavioural outcomes (LSYPE)

(1) Variable Name	(2) Question	(3) Time window w.r.t. survey (Years before)	(4) Mean
Truancy	In the last 12 months, have you ever played truant, that is missed school without permission, even if it was only for a half day or a single lesson?	Up to 1 year	0.14
One month absence	Can I check, in the last 12 months, has (name) been off school for a continuous period of 1 month or more, other than for school holidays?	Up to 1 year	0.03
Being bullied	The next question is about any bullying or other bad behaviour from other pupils at (his/her) school that you know have happened to (name) in the last 12 months. Have any of these things happened to (name) at school in the last 12 months? 1. Called names by other pupils at his/her school 2. Sent offensive or hurtful text messages or emails 3. Shut out from groups of other pupils or from joining in things 4. Made to give other pupils his or her money or belongings 5. Threatened by other pupils with being hit or kicked or with other violence 6. Actually being hit or kicked or attacked in any other way by other pupils 7. Any other sort of bullying 8. No, none of these things have happened in the last 12 months	Up to 1 year	0.43
Suspension	Has (name) been temporarily excluded, that is suspended, from a school for a time, in the past 3 years?	Up to 3 years	0.10
Expulsion	Has (name) been permanently excluded, that is expelled from school for good, in the past 3 years?	Up to 3 years	0.01
Police warning	Have the police got in touch with you (or your husband/wife/partner) about (name) because of something he/she had done in the last 3 years? 1. Yes , in last 3 years; 2. No; 3. Not in the last three years	Up to 3 years	0.07

Timing of events

Key Stage 2 tests — May 2001

Key Stage 2 results — July 2001

September 2001

Students' behaviour — March 2001 - October 2003

Key Stage 3 tests — May 2004

Key Stage 3 results — July 2004

Primary education

Secondary education

Disclosure of tests results

2010 end of key stage 2 pupil results

Qualifications and Curriculum Development Agency

Pupil's name		Class	

English

Teacher assessment results

Speaking and listening	Level	
Reading	Level	
Writing	Level	
Overall English result	Level	

Test results

Reading	Level	
Writing	Level	
Overall English result	Level	

Mathematics

Teacher assessment result	Level	
Test result	Level	

Science

Teacher assessment result	Level	

Level 3 and below represents achievement below the nationally expected standard for most 11-year-olds. Level 4 represents achievement at the nationally expected standard for most 11-year-olds. Levels 5 and 6 represent achievement above the nationally expected standard for most 11-year-olds.

External markers mark exams using a continuous scale [2.5,6]

Students obtain categorical results {3,4,5}

Motivation

Interest by policy in the determinants of

- achievement gaps and consequences in adulthood, e.g. No Child Left Behind (USA)

- well-being in young age and aulthood, e.g. Every Child Matters (UK)

Literature

- mixed evidence on behavioural effects of characteristics of the institutional setting in education (Reback (2010), Gaviria and Raphael (2001) and Dee (2004))

- positive non-market returns to education in adulthood (Grossman (2006) and Oreopoulos and Salvanes (2009))

Research design

B is a binary variable measuring behaviour. It can be interpreted as the observable proxy of whether a latent variable B^* about the importance of school is smaller than a threshold $\bar{B}^*$ or $B = I\{B^* < \bar{B}^*\}$

$$B = \alpha + \beta_{OLS}T + U_1 \tag{1}$$

β_{OLS} in equation (1) measures the change in the probability of behaviour, e.g. a student is truant, due to a unit increase in test score T

$$G_c = I\{T \geq \bar{T}_c\} \tag{2}$$
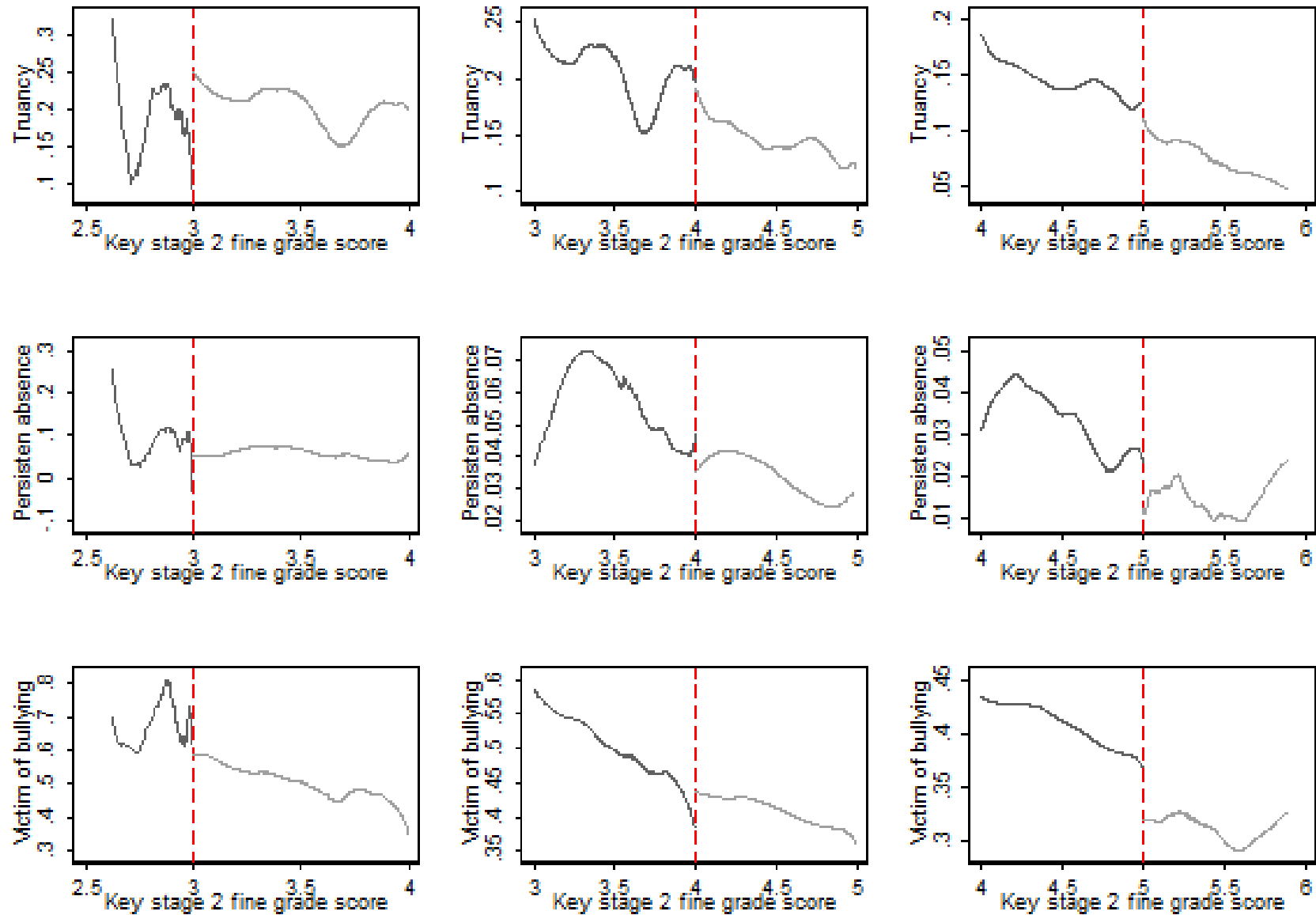$$B = f(T) + \beta_{RD}G_c + U_2 \tag{3}$$

Cutoffs in test scores $\bar{T}_c \in \{3, 4, 5\}$ determine whether a student meets a performance target G_c in equation (2)

β_{RD} in equation (3) measures the change in the probability of behaviour that meeting a target w.r.t. not meeting it leads to

Estimation

- Sharp discontinuities

- Estimate smooth polynomials in test scores separately for students to the left and right of a cutoff

- Choose the optimal bandwidth using the algorithm in Imbens and Kalyanaraman (2009)

- Use a window of size 2 and centered around each cutoff, e.g. cutoff 4 +/- 1

- Robustness checks for the validity of the design

RD plots: absence and being bullied

OLS and RD estimates

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	All sample				Females				Males			
	OLS	RD 2-3	RD 3-4	RD 4-5	OLS	RD 2-3	RD 3-4	RD 4-5	OLS	RD 2-3	RD 3-4	RD 4-5
Truancy	-.03	.08	-.05	-.01	-.03	-.07	-.02	.03	-.03	.12	-.06	-.06
	(.01)**	(.07)	(.04)	(.02)	(.02)	(.11)	(.04)	(.03)	(.02)**	(.08)	(.04)	(.02)***
Obs.	9858	1729	7119	8128	4871	800	3529	4070	4987	929	3590	4058
One month absence	-.02	-.03	-.004	-.02	-.02	-.02	-.007	-.01	-.02	-.03	.01	-.02
	(.007)***	(.04)	(.01)	(.01)*	(.01)**	(.06)	(.02)	(.02)	(.009)**	(.06)	(.02)	(.01)*
Obs.	9540	1634	6837	7905	4670	741	3360	3928	4870	893	3477	3977
Victim of bullying	-.09	-.15	.04	-.04	-.09	-.26	.06	-.07	-.08	-.12	.01	-.05
	(.02)***	(.08)*	(.03)	(.02)*	(.02)***	(.09)***	(.05)	(.04)*	(.02)***	(.12)	(.05)	(.04)
Obs.	9119	1586	6552	7532	4520	722	3255	3797	4599	864	3297	3735
Suspension	-.04	.006	.03	-.02	-.03	.28	.0004	.01	-.04	-.15	.04	-.05
	(.01)***	(.06)	(.02)	(.02)	(.01)***	(.12)**	(.03)	(.02)	(.02)**	(.10)	(.03)	(.03)
Obs.	9657	1676	6936	7980	4742	759	3424	3982	4915	917	3512	3998
Expulsion	-.002	-.004	-.0009	.004	-.003	.03	.006	.005	-.002	-.01	-.009	.003
	(.002)	(.04)	(.005)	(.002)*	(.002)	(.04)	(.004)	(.004)	(.004)	(.05)	(.009)	(.003)
Obs.	9674	1679	6948	7994	4747	762	3426	3984	4927	917	3522	4010
Police warning	-.03	.13	.06	-.02	-.008	-.21	.0001	.01	-.05	.14	.11	-.04
	(.01)**	(.08)*	(.03)*	(.02)	(.01)	(.07)***	(.04)	(.02)	(.02)***	(.09)	(.04)**	(.04)
Obs.	4458	768	3182	3690	2241	363	1606	1878	2217	405	1576	1812

- Effects are (not) significant at the expected target in (full sample) sub-samples by gender

- OLS (reduced-form) estimates are downward (upward) biased w.r.t. RD ones

Robustness check: pre-treatment values do not lead to jumps

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Achievement cutoff 3			Achievement cutoff 4			Achievement cutoff 5		
	Left	Right	P-value	Left	Right	P-value	Left	Right	P-value
Male	0.59	0.67	0.15	0.49	0.52	0.22	0.45	0.50	0.06
Teacher Assessment Test scores									
English lev. 2	0.46	0.56	0.12	0.03	0.02	0.88	.	.	.
English lev. 3	0.43	0.38	0.73	0.52	0.56	0.16	0.01	0.01	0.42
English lev. 4	0.02	-0.00	0.85	0.41	0.38	0.75	0.55	0.57	0.28
Maths lev. 2	0.31	0.50	0.02	0.01	0.00	0.83	.	.	.
Maths lev. 3	0.60	0.41	0.99	0.60	0.53	0.94	0.00	-0.00	0.88
Maths lev. 4	0.00	-0.00	0.53	0.35	0.42	0.05	0.48	0.56	0.04
Science lev. 2	0.25	0.22	0.67	0.00	-0.00	0.79	.	.	.
Science lev. 3	0.67	0.71	0.35	0.26	0.28	0.37	0.00	-0.00	0.75
Science lev. 4	0.00	0.04	0.10	0.69	0.67	0.66	0.41	0.43	0.27
School type at Key Stage 2									
VA school	0.06	0.11	0.16	0.15	0.14	0.69	0.20	0.24	0.08
VC school	0.10	0.08	0.61	0.09	0.08	0.59	0.09	0.10	0.34
Found.n school	-0.00	0.01	0.15	0.02	0.02	0.62	0.03	0.03	0.26
Ethnicity									
Black	0.05	0.11	0.07	0.10	0.08	0.81	0.06	0.06	0.53
Asian	0.31	0.24	0.80	0.19	0.22	0.15	0.13	0.13	0.53
Other	0.04	0.07	0.30	0.05	0.06	0.20	0.06	0.07	0.23
Socio-economic background									
SEN statement	0.14	0.21	0.11	0.03	0.02	0.77	0.00	0.00	0.22
SEN non-statemented	0.64	0.53	0.92	0.22	0.29	0.02	0.02	0.04	0.04
FSM	0.41	0.44	0.31	0.24	0.26	0.30	0.13	0.13	0.38
EAL	0.35	0.29	0.78	0.23	0.26	0.14	0.16	0.16	0.43
Main parent (MP)									
MP with a degree	0.01	0.02	0.35	0.04	0.06	0.13	0.12	0.18	0.01
MP higher education	0.11	0.02	0.96	0.08	0.11	0.16	0.17	0.15	0.84
MP GCSE	0.23	0.29	0.19	0.43	0.36	0.96	0.49	0.45	0.87
MP other qualification	0.16	0.20	0.26	0.15	0.13	0.81	0.07	0.10	0.07
MP's father with a degree	0.03	0.01	0.68	0.06	0.05	0.75	0.06	0.08	0.15

Robustness check: no spikes/gaming at cutoffs (McCrary (2008))
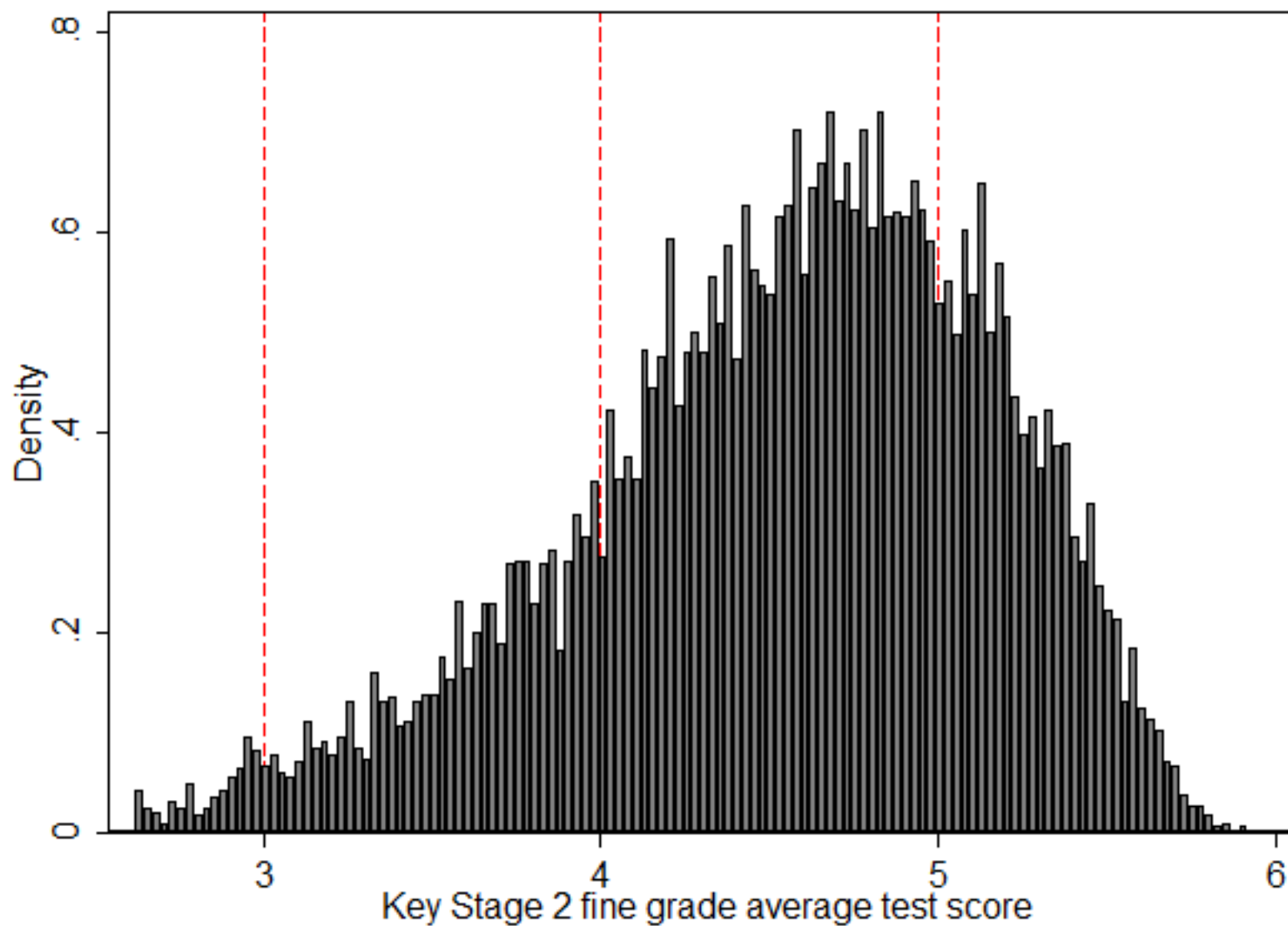
No rejection of the null hypothesis of no gaming

Discussion

- Heterogeneity in effects by ability, gender, type of outcome and parents' education suggest nature-nurture tradeoff (Lizzeri and Siniscalchi (2008))

- Valuable test to
  - assess behavioural effects of targets
  - inform education and public policies in the future (Urquiola and Verhoogen (2009))

- Future work:
  - i. motivation, effort and achievement (De Fraja *et al.* (2010))
  - ii. achievement and behaviour in secondary school and college
  - iii. statistical design to inform public policy

References

DE FRAJA, G., OLIVEIRA, T. and ZANCHI, L. (2010). Must try harder: Evaluating the role of effort in educational attainment. *The Review of Economics and Statistics*, **92** (3), 577–597.

DEE, T. S. (2004). Are there civic returns to education? *Journal of Public Economics*, **88** (9-10), 1697–1720.

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

GROSSMAN, M. (2006). Education and nonmarket outcomes. *Handbook of the Economics of Education*, vol. 1, *10*, Elsevier, pp. 577–633.

IMBENS, G. and KALYANARAMAN, K. (2009). *Optimal Bandwidth Choice for the Regression Discontinuity Estimator*. Working Paper 14726, National Bureau of Economic Research.

LIZZERI, A. and SINISCALCHI, M. (2008). Parental guidance and supervised learning. *The Quarterly Journal of Economics*, **123** (3), 1161–1195.

MCCRARY, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, **142** (2), 698–714.

OREOPOULOS, P. and SALVANES, K. G. (2009). *How large are returns to schooling? Hint: Money isn't everything*. Working Paper 15339, National Bureau of Economic Research.

REBACK, R. (2010). Schools' mental health services and young children's emotions, behavior, and learning. *Journal of Policy Analysis and Management*, **29** (4), 698–725.

URQUIOLA, M. and VERHOOGEN, E. (2009). Class-size caps, sorting, and the regression-discontinuity design. *American Economic Review*, **99** (1), 179–215.