

Does a good mark affect the well-being of students?
Evidence from discontinuities in test scores in England

Marcello Sartarelli
INSTITUTE OF EDUCATION

5 June 2009

Education and well-being of school children

Joint objective of compulsory education:

- help students to build human capital / signal their ability through achievement
- teach students norms to foster well-being, e.g. lead a healthy and safe life

Sharp increase in interest in the role of education in shaping individuals' well-being:

- "No Child Left Behind" by the US Department of Education
- "Every child matters" by the UK Department of Children, School and Family

No consensus on the channels through which education influences children's well-being

Contribution and findings

- interpret and identify the effect of achievement on the well-being of students
- estimate a negative effect of achievement on the probability that the police stops the child using data on secondary school children in England

Summary evidence on well-being of school children

Whether police have got in touch with the family because of children actions	Whether ever vandalised public property		Total
	No	Yes	
No	.8645	.0659	.9304
Yes	.0502	.0194	.0696
Total	.9147	.0853	1

Source: [Longitudinal Survey of Young People in England Wave 2](#), 12344 observations

Parents report that **7%** of children behaved in such a way to prompt the police to contact or visit; under-reporting is a potential source of bias

Children report that **9%** vandalised public property; estimates vary by gender and socio-economic background

Among the children whose parents get a police contact or visit:

- few report to have vandalised public property
- the remaining contacts by the police deal with such other self-reported problems as shoplifting, smashing public property and fights

Achievement levels and test marks in the 2004 Key Stage 3 maths tests, tier 3-5

(1) Achievement level	(2) Test mark range	(3) Percentage of students
2	23-28	2.97
3	29-63	36.31
4	64-98	47.79
5	99+	12.95

Source: National Pupil Database, 12344 observations

36% of children score level 3 in the Key Stage 3 maths while DCSF expects them to reach level 4 or greater at that stage of their education

Achievement and broader outcomes of education are of great concern in the UK due to mixed evidence on the disengagement of school children

Assignment of students across tiers in Key Stage 3 Maths

(1) KS3 Maths test tier	(2) Percentage of students	(3) Percentage of students below level 4 \forall tier	(4) Percentage of families the police contacted
3-5	21.98	39.2	10.9
4-6	33.71	0.56	7.9
5-7	27.95	0	4.3
6-8	16.33	0	2.1

Source: [Longitudinal Survey of Young People in England Wave 2](#) and National Pupil Database, 12344 observations

The paper focuses on students in Key Stage 3 Maths tier 3-5 among whom 39.2% under-achieves

From tier 4-6 on

- there are no under-achievers
- police contacts decrease

Discretionary assignment by teachers of students to different *tiers* in Key Stage tests motivates the focus on one tier; Miranda and Sartarelli (2009) deal with inter-tiers variation

Institutional framework: the British national school curriculum

Age	Stage	Year	Assessment
3-4	Early Years Foundation Stage		
4-5		Reception	
5-6	Key Stage 1	1	
6-7		2	Teacher assessments in English, maths and science
7-8	Key Stage 2	3	
8-9		4	
9-10		5	
10-11		6	National tests and teacher assessments in English, maths and science
11-12	Key Stage 3	7	Ongoing teacher assessments
12-13		8	Ongoing teacher assessments
13-14		9	National and teacher assessments English, maths and science and the other foundation subjects
14-15	Key Stage 4	10	Some children take GCSEs
15-16		11	Most children take GCSEs or other national qualifications

The UK school curriculum for children from 5 years of age to the compulsory school leaving age at 16 is divided into four Key Stages (KS) from 1 to 4

Secondary school goes from Key Stage 3 to 4 with children aged 11 to 16

Institutional framework: grading of test scripts

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

- test marks for KS2 only and others on demand
- 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:

- DCSF
- Qualifications and Curriculum Authority (QCA)

Partial literature review

Gibbons and Silva (2008), enjoyment at school and achievement

Benabou and Tirole (2003), motivation and achievement

Flink et al. (1990), students-teachers interaction and cognitive dissonance

Lizzeri and Siniscalchi (2008), sheltering - learning from experience tradeoff

Akerlof and Dickens (1982) risk and cognitive dissonance

Grossman (2005), education and non-market outcomes

Heckman et al (2006), (non-)cognitive abilities, labor market outcomes and social behaviour

Stylised model of achievement and well-being

Let parents and children jointly choose time to spend in supervised learning t_s and leisure t_l to maximise

- a concave and C^2 utility u function arising from two activities/goods:
 - supervised learning $s(t_s, a)$ by schools/parents with varying intensity of incentives, competition and achievement captured by a
 - leisure or free-of-supervision activity $l(t_l)$
- subject to a time constraint $t_s + t_l \leq 1$

$$\begin{aligned} \max_{t_s, t_l} u(t_s, t_l, a) &= \underbrace{s(t_s, a)}_{\text{supervised activity}} + \underbrace{l(t_l)}_{\text{leisure activity}} \\ \text{s.t.} \quad t_s + t_l &= 1 \end{aligned}$$

On average parents would like extra t_s while children extra t_l

Comparative statics of the model equilibrium

$$t_l^* \in \arg \max_{t_l} u(t_l, a) = \underbrace{s(1 - t_l, a)}_{\text{supervised activity}} + \underbrace{l(t_l)}_{\text{leisure activity}}$$

$$\frac{\partial t_l^*}{\partial a} \begin{matrix} \leq \\ > \end{matrix} 0$$

The change in the optimal allocation t_l^* of leisure time by parents and children is a proxy for a change in well-being, defined by a measure of outcome evaluated by some other party as in Heckman (2008)

Benabou and Tirole (2003) reconcile the contrasting predictions on the effect of a shock to the utility function in:

- economics where incentives matter $\frac{\partial t_l^*}{\partial a} < 0$, see Lazear (1990)
- psychology where incentives may undermine motivation and $\frac{\partial t_l^*}{\partial a} > 0$, see Flink et al. (1990)

Whether $\frac{\partial t_l^*}{\partial a} \begin{matrix} \leq \\ > \end{matrix} 0$ is ultimately an empirical matter

Research design

$$T^* = \alpha + \beta A + U \quad (1)$$

$$T = \begin{cases} 1 & \text{if } T^* \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

$$A = f(Z) \quad (2)$$

$$\begin{aligned} \lim_{Z \downarrow \bar{Z}} E[T|Z] - \lim_{Z \uparrow \bar{Z}} E[T|Z] &= \beta + \lim_{Z \downarrow \bar{Z}} E[U|Z] - \lim_{Z \uparrow \bar{Z}} E[U|Z] \\ &= \beta \text{ if } \lim_{Z \downarrow \bar{Z}} E[U|Z] = \lim_{Z \uparrow \bar{Z}} E[U|Z] \end{aligned} \quad (3)$$

T^* is latent/unobservable leisure time. T equals 1 in the event of a police contact triggered by a value $T^* \geq 0$ above which well-being deteriorates and 0 otherwise

A is qualitative achievement, e.g. bad or good; unobservables may correlate with A and T , e.g. parental guidance

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

β is the average effect of A on T under the identifying assumption $\lim_{Z \downarrow \bar{Z}} E[U|Z] = \lim_{Z \uparrow \bar{Z}} E[U|Z]$ that errors are continuous at \bar{Z}

Illustration of the research design using an administrative discontinuity

(1) Achievement level	(2) Test mark range	(3) Percentage of students
2	23-28	2.97
3	29-63	36.31
4	64-98	47.79
5	99+	12.95

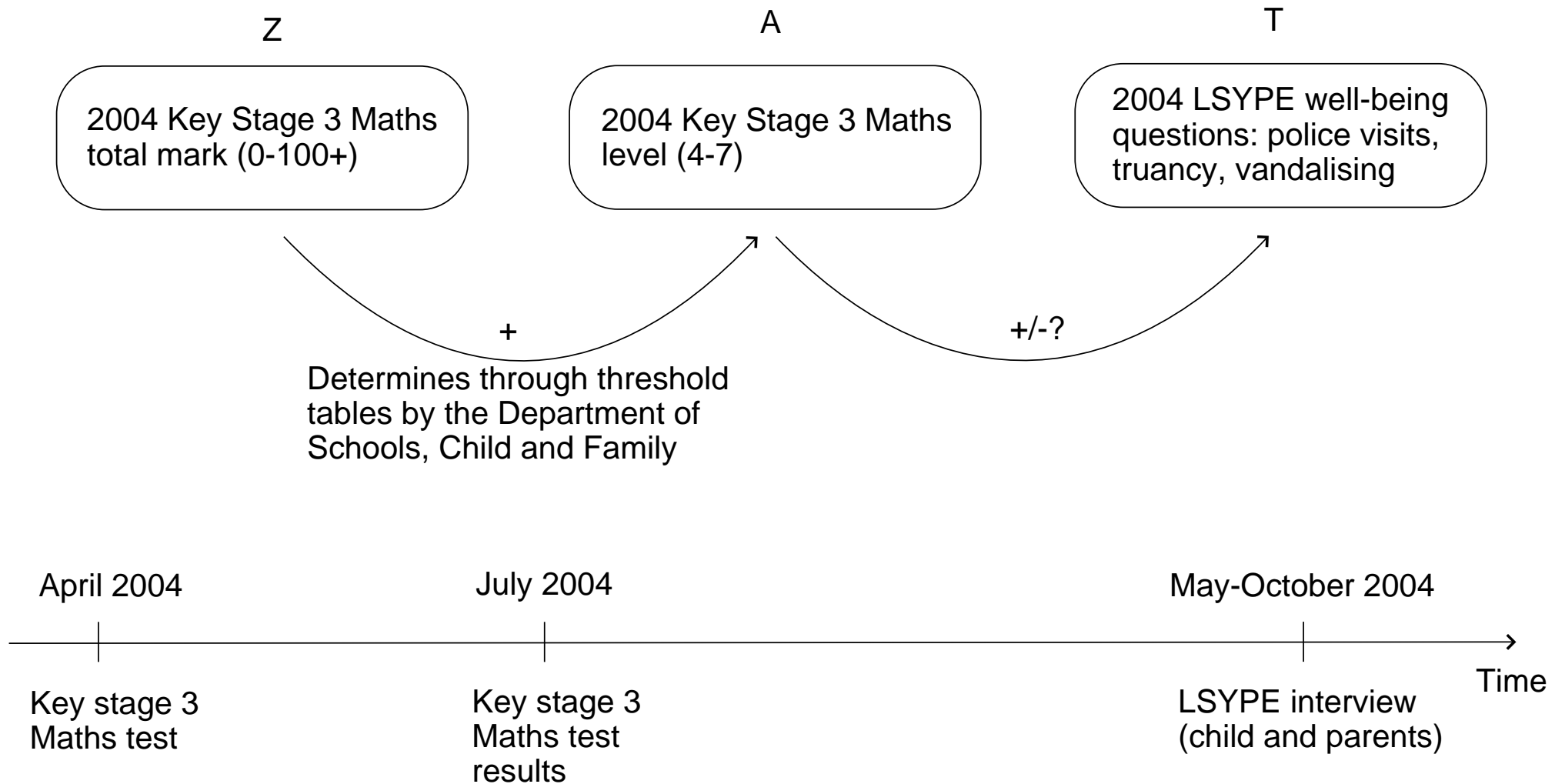
Source: UK Qualifications and Curriculum Authority

An administrative rule mapping test marks in column (2) to achievement levels in column (1):

- is the assignment to treatment rule
- conforms to a *sharp* regression discontinuity design (RDD)

Two students scoring 63 and 64 achieve respectively levels around the cutoff 3, bad and 4, good, by chance, absent manipulation of the score or the level

Illustration of the research design

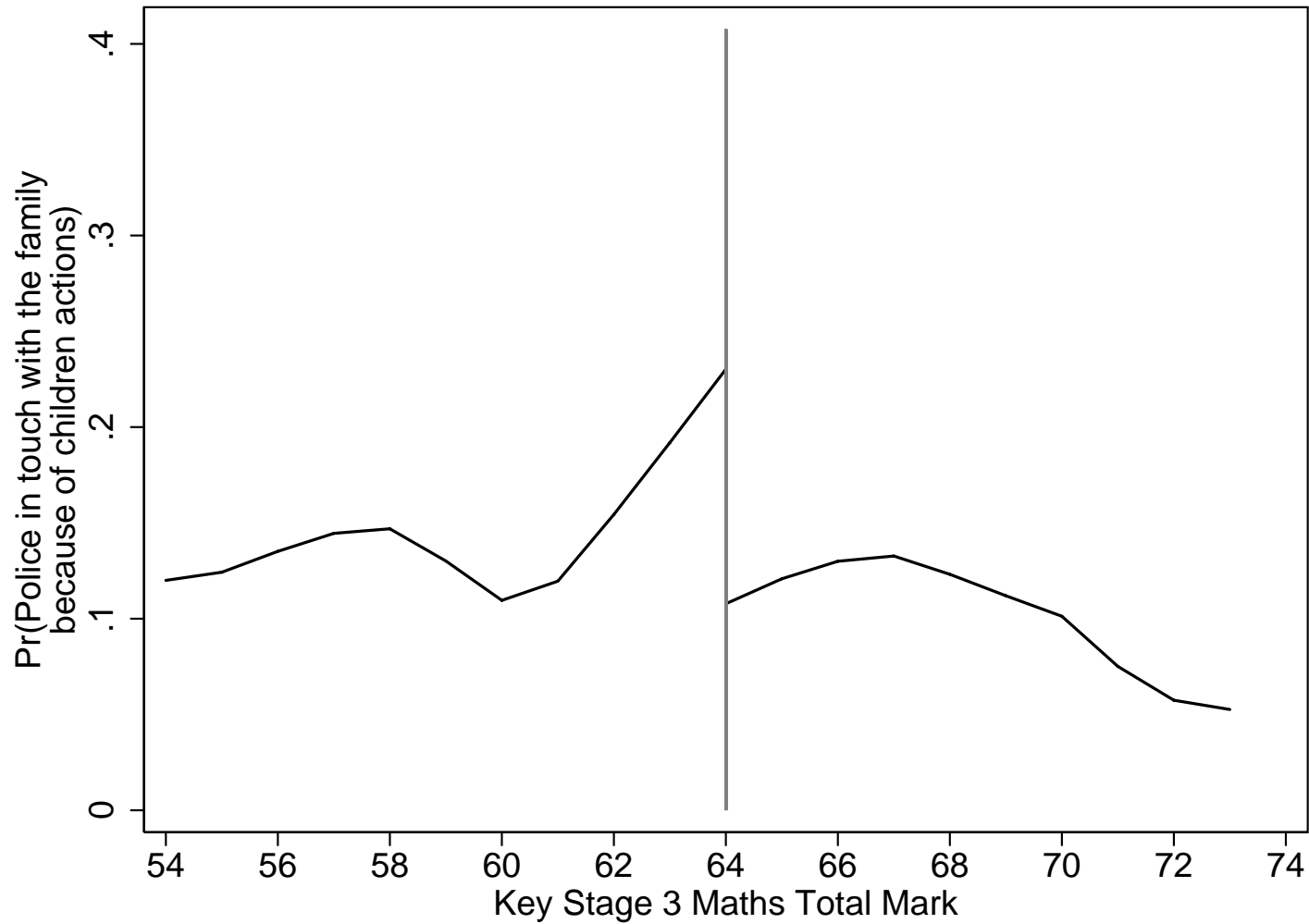


RDD estimation strategy in 3 steps

I carry out as in Imbens and Lemieux (2007):

- 1) graphical analysis by estimating local polynomials $f(Z)$ of well-being separately below/above the achievement cutoff \bar{Z}
- 2) 2SLS-IV estimates of the effect of a good mark on well-being following the characterisation of TSLS as RDD in Hahn et al. (2001)
- 3) robustness checks to rule out that estimates are potentially biased due to:
 - unobservables that affect both pre-treatment variables and the outcome variable
 - sorting around the achievement cutoff \bar{Z}

Estimation step 1: graphical analysis



An increase KS3 maths test mark from left to right of the levels 3-4 cutoff decreases the probability of police contacts by 14 percentage points

Estimation step 1: graphical analysis (cont.d)

(1)	(2)	(3)	(4)
Distance in KS3 maths marks from the cutoff	Estimate	Standard Error	Observations
2	-.144	.113	112
3	-.140	.118	166
4	-.135	.112	223
5	-.134	.113	271

If one re-estimates the effect of the good grade on well-being by using a window that is centered at the cutoff and that has a varying size as in column (1), estimates:

- do not change considerably
- are statistically significant at 10%

Estimation step 2: comparison between RDD and OLS estimates

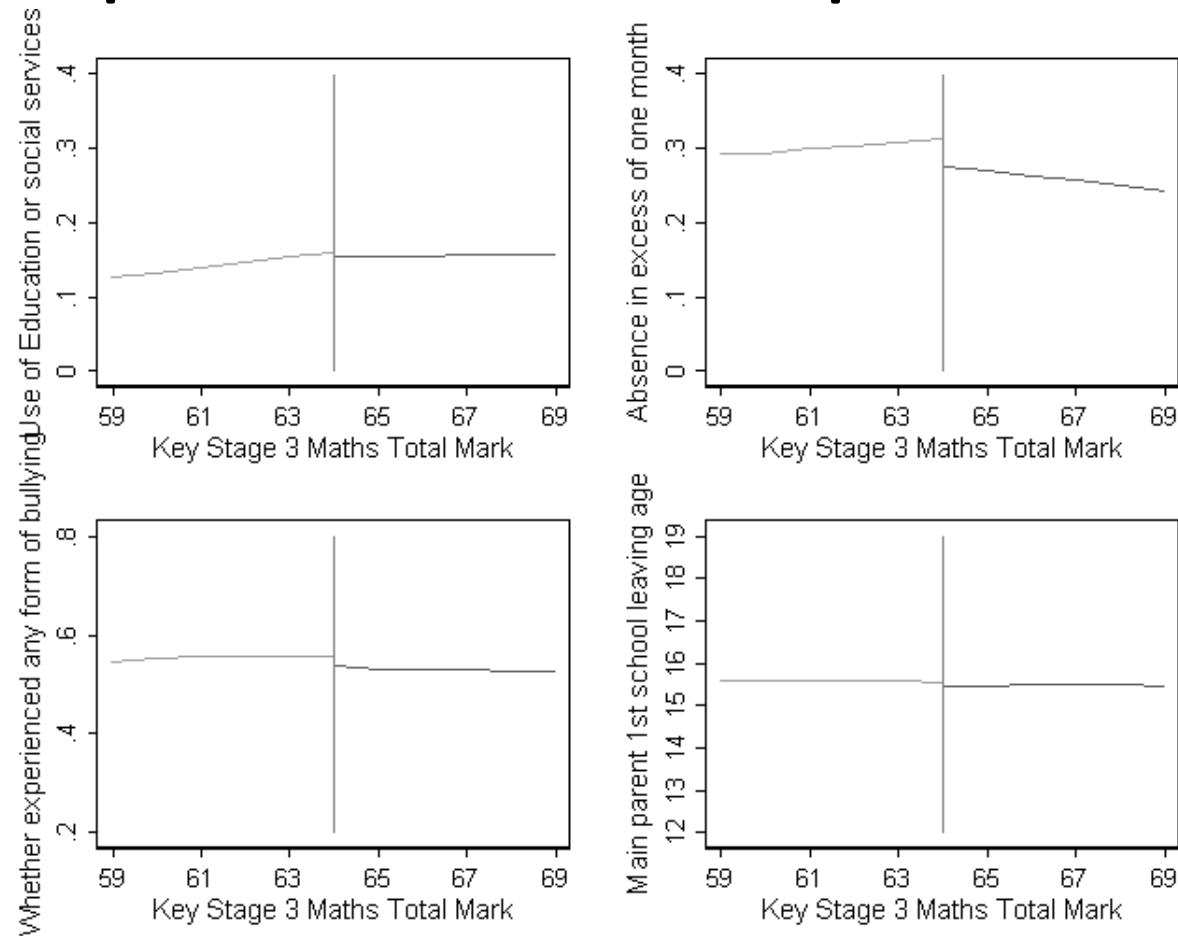
	RDD	OLS		
		Covariates	Achievement quadratic	Achievement quadratic and covariates
Estimate	-.044**	-.036	.003	.013
Standard error	.0176	.047	.031	.071

Estimates using the 2SLS specification in Hahn et al. (2001) say that a jump from a bad grade on the left of the cutoff to a good one on the right decreases the probability of a police contact by 4 percentage points

Under the weak and testable RDD assumptions OLS estimates:

- are upward biased
- yield a spurious zero correlation between achievement and well-being that is likely driven by unobservables

Estimation step 3: unobservables and pre-treatment variables

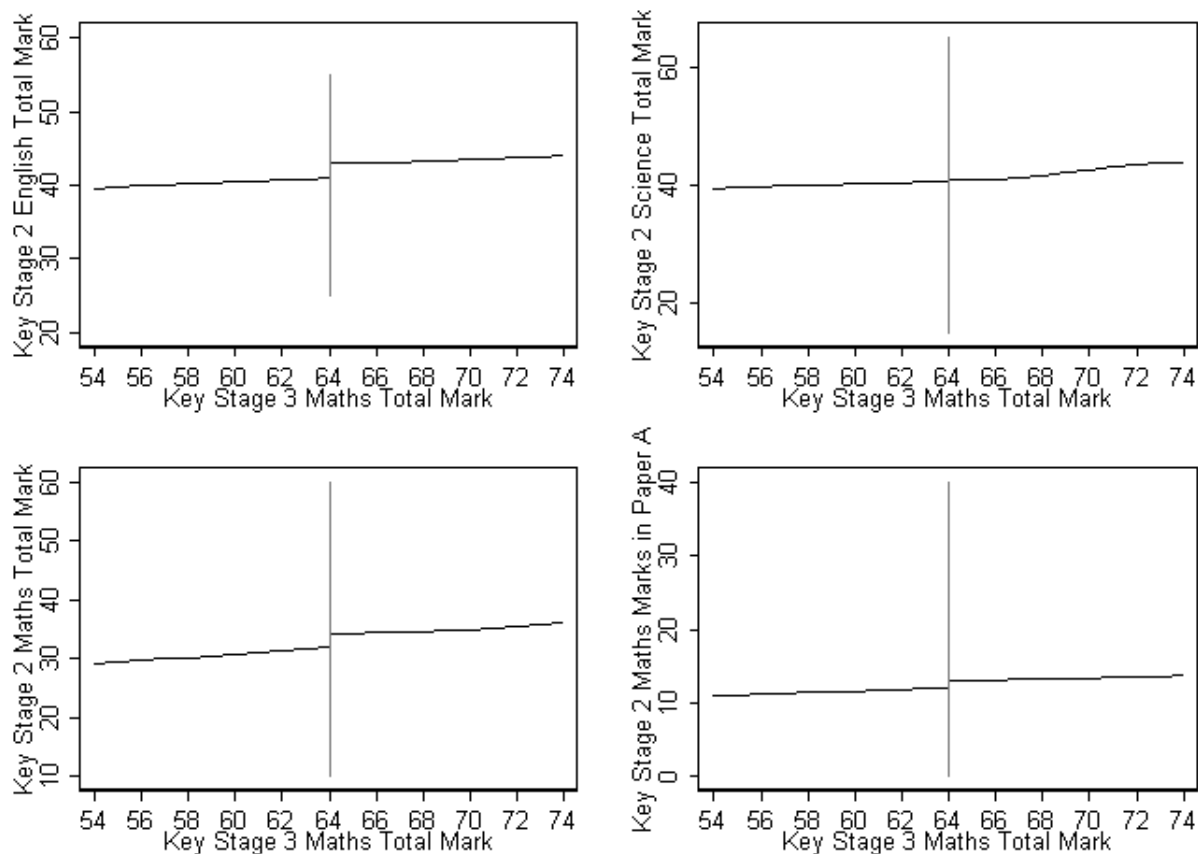


The research design fails if unobservables affect:

- pre-treatment covariates that are determined e.g. the year before the KS3 test
- the KS3 maths test mark around the cutoff at 64

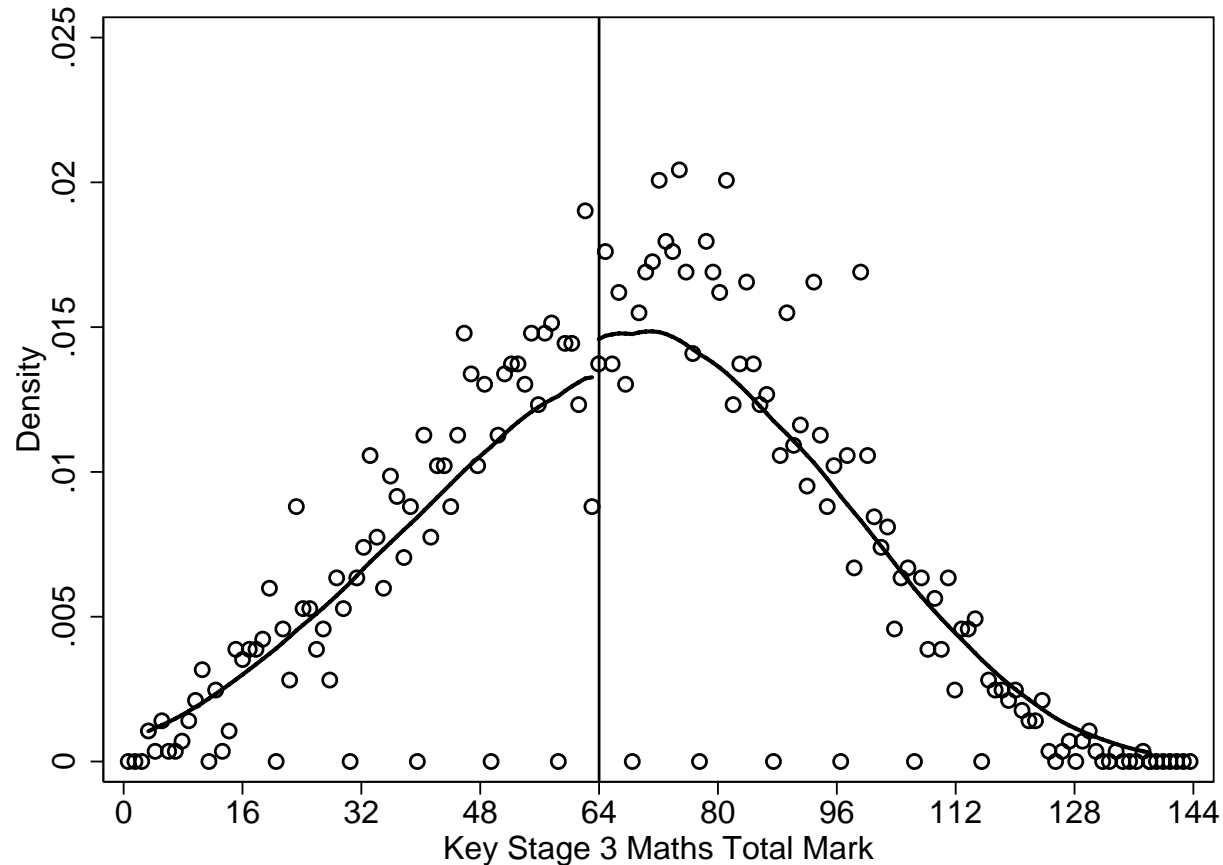
No jumps in pre-treatment variables, e.g. parental guidance, rule out unobservable

Estimation step 3: unobservables and pre-treatment variables (cont.d)



Small jumps in pre-treatment variables around the achievement cutoff support the identifying assumption $\lim_{Z \downarrow \bar{Z}} E[U|Z] = \lim_{Z \uparrow \bar{Z}} E[U|Z]$

Estimation step 3: sorting around the cutoff (cont.d)



Performing the McCrary (2008) test of the continuity of KS3 maths test mark distribution at the mark cutoff at 64:

- no suspicious masses in the empirical density suggest no sorting at the cutoff
- the *null hypothesis of no sorting at the cutoff* is not rejected at 5% significance level

Summary of preliminary findings

Evidence on bully, truant and disengaged school children has increased the concern of policy makers on their well-being and on the role that education plays in fostering it

I interpret the event of a police contact with parents due to the childrens' behaviour as excess unobservable leisure time arising from lack of supervision for a child that can impact on her well-being

I estimate the effect of a good mark on well-being and I find that it decreases the probability of a police contact/visit by 4 percentage points. This suggests that:

- a good mark or positive achievement label fosters well-being
- policy makers put under scrutiny the relative contribution of education to teaching core subjects, e.g. maths, as well as to the well-being of children

Next steps

Sartarelli (2009) looks at the long term effect on the behaviour of adults of:

- secondary school exit exam results
- manipulation of achievement at school

Miranda and Sartarelli (2009) test predictions on motivation and achievement by modelling with a principal-agent framework the:

- discretionary assignment by teachers of students to different tiers at the same Key Stage test
- monitoring activity by the Office for Standards in Education, the "watchdog" of schools and local services in the UK

References

- Akerlof, G. and Dickens, W. (1982) *The Economic Consequences of Cognitive Dissonance*, American Economic Review 72 (3), 307-319
- Benabou, R. and J. Tirole (2003) *Intrinsic and Extrinsic Motivation*, Review of Economic Studies 70, 489-520
- Flink, C., Boggiano, A. and Barrett, M. (1990) *Controlling teaching strategies: Undermining children's self-determination and performance* Journal of Personality and Social Psychology 59, 916-924
- Gibbons, S. and O. Silva (2008) *School Quality, Child Wellbeing and Parents' Satisfaction*, [London School of Economics mimeo](#)
- 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
- Hahn, J., P. Todd and W. Van der Klaauw (2001) *Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design*, Econometrica 69, 201-09
- Heckman, J. (2008) *Econometric Causality*, [NBER Working Paper 13934](#)
- Heckman, James J., J. Stixrud, and S. Urzua (2006) *The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior*, Journal of Labor Economics 24, 41182
- Imbens, G. and T. Lemieux (2008) *Regression discontinuity designs: A guide to practice*, Journal of Econometrics 142, 615-635
- Lazear, E. (2000), *Performance, Pay and Productivity*, American Economic Review 90, 1346-1361
- Lizzeri, A. and Siniscalchi, M. (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
- Miranda, A. and Sartarelli, M. (2009) *Agency theory in schools: do incentives affect performance? Evidence from the assignment to tests in schools*, Institute of Education mimeo
- Sartarelli, M. (2009) *A note on incentives and the manipulation of test scores. Evidence from secondary school tests in the UK*, Institute of Education mimeo